Abstract

Ab Initio Studies of Ferroelectric Thin Films

Mehmet Doğan

2017

Epitaxial interfaces between metal oxides and semiconductors have been of significant research interest due to their potential use in electronic device applications. Thin films of metal oxides can display many functional physical properties, an important example of which is ferroelectricity. Ferroelectric thin metal oxide films grown on semiconductors can enable non-volatile transistors, where the state of the device is encoded in the polarization state of the oxide which determines the electronic transport properties of the semiconductor. This thesis presents theoretical studies of a number of metal oxide on semiconductor systems using first principles electronic structure methods. We have studied the $BaTiO_3/Ge$ interface as a candidate of a ferroelectric oxide/semiconductor system. In one set of studies of this interface, we have shown how cross-interfacial structural couplings can create atomic-scale structural motifs in the metal oxide that do not exist in any of its bulk phases. Separately, we have found that multiple polarization states in the $BaTiO_3$ film are possible and, in principle, that one can switch between them by the application of an external electric field. Unfortunately, the overall direction of the polarization is pinned by the interface chemistry in this system. In order to modify the interface chemistry to promote ferroelectricity, we have proposed the usage of a buffer layer between the oxide and the semiconductor, such as a monolayer of zirconia. We have explored the possible stable configurations of single monolayers of ZrO_2 on Si and found that multiple polarization states are indeed stabilized. We have found that ferroelectric switching between these two structures would lead to modifications of the Si electronic band properties in a manner comparable to available experimental results. We have developed a discrete lattice model to predict domain behavior in these monolayer films at finite temperatures. In a final set of works, we have conducted a study of thin films of doped hafnia which have recently shown ferroelectric behavior. We have focused on strain effects in doped HfO_2 to explain some of the experimental observations from a structural point of view. Our findings provide an understanding for the stabilization of ferroelectricity in hafnia based thin films.

Ab Initio Studies of Ferroelectric Thin Films

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

in Candidacy for the Degree of

Doctor of Philosophy

by

Mehmet Doğan

Dissertation Director: Sohrab Ismail-Beigi

May 2017

Contents

List of Figures

List of Tables

Acknowledgements

First, I would like to express my special appreciation and thanks to my advisor, Sohrab Ismail-Beigi. In addition to being a great scientist, he is also an admirable mentor. I deeply appreciate the time he spent with me to teach me physics, math, chemistry, coding and communicating science, to count just a few. He has always listened to me and directed me in the right path. When I joined the group in 2012, I had not taken a single class in solid state physics. Through his mentorship, Sohrab has made it possible for me to become a materials physicist who is able to contribute to scientific knowledge.

I would like to thank my Thesis Committee members, Victor Henrich, Corey O'Hern and Fred Walker for their scientific support through my PhD candidacy, and also for their comments and suggestions for the improvement of this thesis.

I would like to thank my experimental collaborators, Divine Kumah, Fred Walker, Nanbo Gong and T. P. Ma, for providing me with a rich opportunities to do interesting science that is also experimentally relevant. I have learned a lot from them and also gained a thorough understanding of my research field through my interactions with them. I would also like to thank my previous advisor, Karyn Le Hur, with whom I had the opportunity to work for one year.

The members of the Ismail-Beigi Research Group have contributed immensely to my personal and professional time at Yale. The group has been a source of friendships as well as good advice and collaboration. I would like to thank my colleagues Stephen

Eltinge, Alexandru Georgescu, Jie Jiang, Arvin Kakehani, Minjung Kim, Xin Liang, Andrei Malashevich and Subhasish Mandal.

My time at Yale was made enjoyable in large part by the many friends that became a part of my life. I would like to thank Burak Kaya with whom I enjoyed discussing math and music. I would also like to thank Ferzan Tapramaz, with whom I spent many nights drinking Turkish tea discussing everything from games to linguistics. Finally, I would like to thank Judith Hoeller, Naim Karaçaylı and Josh Stanley, who became three of my very close friends later in my PhD.

My wife, Ceyda, and I had the opportunity to have wonderful neighbors with whom we shared many memories. I would like thank our former neighbor, Jane Dougall, and her beloved cat, Boo. Jane baked us desserts to cheer us up when Ceyda and I were occupied with deadlines. I admired Boo for his love of life. In his mid-twenties and after an accident that left him with no tail and an injured leg, he was still full of joy. I would also like to thank my former landlady, Nissa Simon, who helped me navigate my life during my first year in the United States.

I am especially grateful for my loving family. My mother has always been a great role model of dedication and resilience, and my father has been an endless source of unconditional support. I would like to thank my brother, Yunus, for always being an instant message away to cheer me up.

Last, but certainly not least, I would like to thank my wife, Ceyda Durmaz Doğan. She has been part of every step of this long and challenging journey. Without her unconditional love, I would not have been able to be a happy and productive PhD candidate. She has been with me through the happiest and the most difficult moments, laughing, crying, working, playing, planning and discussing. I am especially grateful for her support through the last bit of this journey. While I spent countless hours on writing and making figures, she patiently and tirelessly took care of me.

Chapter 1

Introduction

Crystalline metal oxides have been one of the primary ingredients of nanodevices since the beginning of the age of nanotechnology in the 1980s [2]. These materials exhibit almost every possible solid state phenomenon. The possibilities multiply when thin films of crystalline metal oxides interface with other crystalline materials. The recent advances in fabrication techniques such as growth by molecular beam epitaxy (MBE) allow for heterostructures with atomically abrupt interfaces so that the structural and electronic properties of the two interfacing materials couple [3]. In this Thesis, we primarily focus on crystalline metal oxide/semiconductor interfaces with the aim of achieving ferroelectricity in these oxide films.

Thin films of crystalline metal oxides grown on semiconductors are a potential basis of future transistors. A typical metal-oxide-semiconductor field-effect transistor (MOSFET) chiefly consists of a semiconductor substrate and a thin insulating layer (see Figure 1.1). Traditionally Si is used as the semiconductor and amorphous SiO_2 is used as the insulator. MOSFETs have been shrinking in size over the decades, halving roughly every 18 months (known as Moore's law) [3]. In order for a transistor to be reliable and able to operate at room temperature, the capacitance of the insulator must be kept roughly constant as the size decreases [4]. As the the area of the

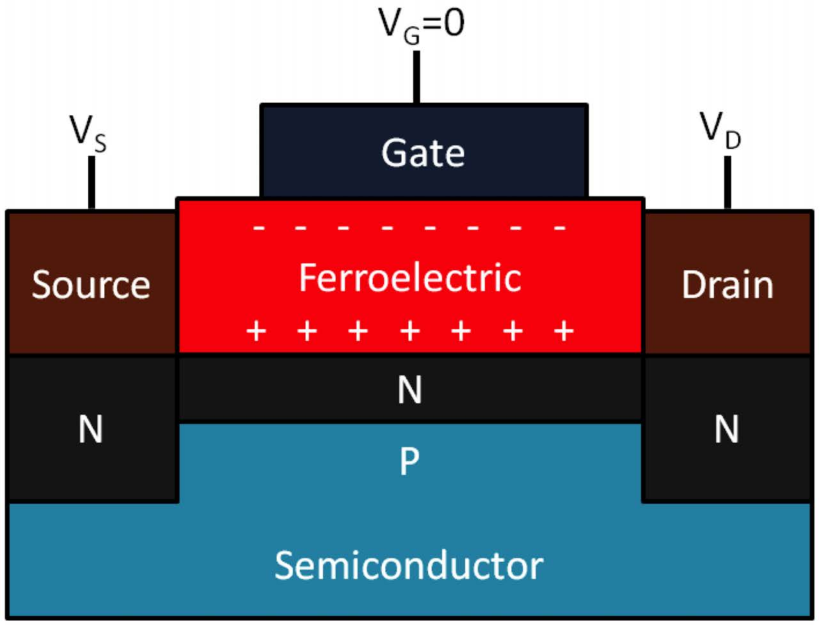


Figure 1.1: A typical MOSFET with a ferroelectric as the insulating layer.

transistor decreases, the insulating layer must also get thinner in order to keep the capacitance the same, according to the equation

$$C = \frac{\kappa \epsilon_0 A}{t}, \tag{1.1}$$

where C is the capacitance, κ is the relative permittivity (dielectric constant), ϵ_0 is the permittivity of free space, A is the area and t is the thickness of the insulating layer. The thickness of the SiO_2 layer has become as small as 1 nm, which gives rise to gate induced leakage currents due to quantum tunneling of electrons, and thus renders the layer non-insulating [5]. Using materials with higher dielectric constants than SiO_2, or "high-k materials", is one potential solution to this issue. If κ is high enough, it is possible to thicken the layer in order to prevent leakage currents while keeping the capacitance fixed. Some crystalline metal oxides are good candidates for high-k materials, having high dielectric constants at room temperature. They are also more easily studied both theoretically and experimentally compared to amorphous oxides. Recently $SrTiO_3$, with a room-temperature static dielectric constant of 300, was successfully grown on Si [6]. The capacitance of this 15 nm-thick $SrTiO_3$ region was equivalent to that of 1 nm-thick SiO_2. This means that quantum tunneling of electrons across the top and the bottom of the film is greatly suppressed, while retaining the same surface area and electrical properties. Another promising high-k material candidate is $BaTiO_3$, which has a room-temperature static dielectric constant of 5000 [7].

In addition, being ferroelectric in the bulk, $BaTiO_3$ might help realize non-volatile transistors. A non-volatile transistor would not require constant application of a gate voltage to stay in the "on" state; the polarization of the ferroelectric would remain "on" even after the applied voltage is turned off (see Figure 1.1). If this feature can be achieved, it would reduce the role of the gate voltage to a momentary switch, immensely lowering the power consumption of the transistor. The hypothetical device with Ge as the semiconductor and $BaTiO_3$ as the insulator was patented in 1957, but the attempts to actualize it by glueing the BTO onto Ge failed due to the uncoupling of the two regions [6].

Epitaxial growth of crystalline metal oxides on semiconductors is best achieved through layer-by-layer molecular beam epitaxy (MBE) [3]. The most important benefit of this method is the ability to modify the growth conditions to achieve thermodynamic stability for each layer. On silicon, the greatest challenge to the growth has been the spontaneous formation of SiO_2 at the interface, which requires careful manipulation of temperature and oxygen pressure. Epitaxial $BaTiO_3$ on Ge was first achieved with a 25 nm-thick film of $BaTiO_3$ [6]. It has been theoretically estimated that a BTO film would not be ferroelectric in the direction that is normal to the interfacial plane, if it were thinner than 7 nm [8]. The latest stage in the experimental growth, being conducted at Yale by our collaborators, consists of BTO slabs of 5 and 11 atomic layers, corresponding to approx. 1 nm and 2 nm, respectively. In the early parts of this work, we study the epitaxial interface between BTO and Ge and compare our findings with the experimental results.

In a related but separate course of research, we investigate the structural properties of ZrO_2 monolayers on Si(001). Our collaborators at Yale have grown amorphous but atomically abrupt ZrO_2 films with the thickness of a single monolayer on Si(001). C-V measurements indicate that these films may be ferroelectric. Our results indicate that there is a multiplicity of (meta)stable configurations with varying out-of-plane polarizations, which corroborates the experimental claim of ferroelectricity. In order to predict the behavior of these films at finite temperatures, we set up a discrete lattice Hamiltonian, which we then treat using Monte Carlo simulations. Lastly, we investigate the configurations of ZrO_x monolayers on Si varying oxygen content.

For a final set of studies, we turn our attention to doped HfO_2. HfO_2 is widely used as a gate oxide in transistor devices today, and the recent discovery of ferroelectricity in HfO_2-based thin films has sparked great research interest [9, 10]. It is well established that the ferroelectric properties of these films greatly depend on film thickness, doping element and concentration, and annealing conditions. Empirical data obtained in the last few years have provoked the nanodevice community to design and develop devices based on these films. However, an atomistic understanding of how ferroelectricity emerges in these films under certain conditions is incomplete. In this Thesis, we attempt to explain some of the empirical findings using bulk and thin film calculations.

This Thesis is organized as follows. In Chapter 2, we introduce density functional theory, which is the main theoretical framework used throughout this work, along with other methods that we have employed. In Chapter 3, we present our published study on the BTO thin films epitaxially grown on Ge [11] which summarizes the experimental results and discusses the structural motifs in the oxide in terms of the eigenmodes of the bulk force constant matrix. In Chapter 4, we report our full theoretical study of these interfaces, focusing on structural and electronic properties and thermodynamic stability of these thin films. In Chapter 5, we present our work on ZrO_2 monolayers on Si, which we find to be a system with a rich landscape of configurations. In Chapter 6, we discuss the origins of ferroelectricity in doped HfO_2 films. Finally, in Chapter 7, we summarize each of the projects reported in this Thesis, and discuss future research.

Chapter 2

Methods

2.1 Density functional theory

The main tool we use throughout this Thesis is density functional theory (DFT) which we shall briefly outline in this section. Developed in 1960s by Hohenberg, Kohn and Sham [12, 13], the chief achievement of DFT is to reduce the full many-body problem to a system of self-consistent equations with functions of a single spatial variable, thereby making the problem much more tractable.

In the Born-Oppenheimer approximation, the non-relativistic Schrödinger equation for N electrons is

$$H\Psi\left(\mathbf{r_1},...,\mathbf{r_N}\right)=E\Psi\left(\mathbf{r_1},...,\mathbf{r_N}\right). \tag{2.1}$$

In natural units ($\hbar=1$, $e=1$, $m_e=1$), H can be written as

$$H=T+V_{ee}+V_{ei}, \tag{2.2}$$

where

$$T=-\frac{1}{2}\sum_{j=1}^{N}\nabla_j^2 \tag{2.3}$$

is the kinetic energy operator,

$$V_{ee} = \frac{1}{2} \sum_{j \neq k}^{N} \frac{1}{|\mathbf{r}_j - \mathbf{r}_k|} \tag{2.4}$$

is the electron-electron interaction, and

$$V_{ei} = \sum_{j=1}^{N} v\left(\mathbf{r}_j\right) \tag{2.5}$$

is the electron-ion interaction with

$$v\left(\mathbf{r}\right) = -\sum_{J} \frac{Z_J}{|\mathbf{r} - \mathbf{R}_J|}. \tag{2.6}$$

Solving for the full N-body wavefunction is not feasible, nor is it needed since we are not interested in tracking all the correlations of all electrons simultaneously. Typically we are interested in the ground state energy E_0, band energies and gaps, and the electron density $n(\mathbf{r})$. E_0 is given by

$$\begin{aligned} E_0 &= \left\langle \Psi_0 \left| T + V_{ee} + V_{ei} \right| \Psi_0 \right\rangle \\ &= \left\langle \Psi_0 \left| T + V_{ee} \right| \Psi_0 \right\rangle + \int d^3\mathbf{r}\, n\left(\mathbf{r}\right) v\left(\mathbf{r}\right), \end{aligned} \tag{2.7}$$

where Ψ_0 is the ground-state wavefunction and $n(\mathbf{r})$ is the ground-state electron density. DFT was born from the observation that the first term does not depend on ionic configuration, and given the ionic configuration evaluation of the second term only requires the electron density $n\left(\mathbf{r}\right)$. Additionally, the ionic configuration enters the evaluation of E_0 only through $v\left(\mathbf{r}\right)$. Thus given $v\left(\mathbf{r}\right)$, we have H, we obtain Ψ_0,E_0 and thus $n\left(\mathbf{r}\right)$. Schematically

$$v\left(\mathbf{r}\right) \rightarrow H \rightarrow \Psi_0, E_0 \rightarrow n\left(\mathbf{r}\right).$$

The first theorem of DFT establishes that there is a one-to-one mapping between $v(\mathbf{r})$ and $n(\mathbf{r})$, i.e. the arrows above are two-way [12]. Thus we can treat Ψ_0 as a functional of density $n(\mathbf{r})$. Defining

$$F[n] \equiv \langle \Psi_0[n] | T + V_{ee} | \Psi_0[n] \rangle , \tag{2.8}$$

we have

$$E_0[n] = F[n] + \int d^3\mathbf{r}\, n(\mathbf{r})\, v(\mathbf{r}) . \tag{2.9}$$

According to the second theorem of DFT, $E_0[n]$ is the minimum of the functional defined as $\mathcal{E}_v[\tilde{n}] \equiv F[\tilde{n}] + \int d^3\mathbf{r}\, \tilde{n}(\mathbf{r})\, v(\mathbf{r})$, and is attained at $\tilde{n} = n$ [12]. The first two theorems of DFT establish that we can express all quantities of interest as functions of $n(\mathbf{r})$, which is a single-variable function, unlike the N-body wavefunction.

In order to compute the functional $F[n]$ for a given $n(\mathbf{r})$, let us write it as follows:

$$\begin{aligned} F[n] &= \langle \Psi_0[n] | T + V_{ee} | \Psi_0[n] \rangle \\ &= T_s[n] + E_{\mathrm{H}}[n] + E_{xc}[n] \\ &= T_s[n] + \frac{1}{2}\int d^3\mathbf{r} \int d^3\mathbf{r}' \frac{n(\mathbf{r})\, n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n] , \end{aligned} \tag{2.10}$$

where $T_s[n]$ is the kinetic energy of a non-interacting electron system with density $n(\mathbf{r})$, $E_{\mathrm{H}}[n]$ (Hartree energy) is the classical electrostatic repulsion between electrons with density $n(\mathbf{r})$, and $E_{xc}[n]$ is the exchange-correlation energy which describes the quantum behavior that is not included in the Hartree term. Exchange (x) refers to the fact that no two electrons with equal spins can simultaneously occupy the same state. Correlations (c) capture the fact that electrons repel each other; so the probability of finding two electrons at positions $\mathbf{r}$ and $\mathbf{r}'$ is less than $n(\mathbf{r})\, n(\mathbf{r}')$ if $\|\mathbf{r} - \mathbf{r}'\|$ is small.

The third theorem of DFT describes how the ground state $n(\mathbf{r})$ is found, which is equivalent to finding the N lowest-energy states of the following set of self-consistent equations, known as Kohn–Sham equations [13]:

$$\begin{aligned}
\left[-\frac{\nabla^2}{2}+v_{\text{eff}}\left(\mathbf{r}\right)\right]\psi_j\left(\mathbf{r}\right) &= \epsilon_j\psi_j\left(\mathbf{r}\right), && (2.11)\\
v_{\text{eff}}\left(\mathbf{r}\right) &= v\left(\mathbf{r}\right)+\int d^3\mathbf{r}'\frac{n\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}+v_{xc}\left(\mathbf{r}\right), && (2.12)\\
v_{xc}\left(\mathbf{r}\right) &\equiv \frac{\delta E_{xc}\left[n\right]}{\delta n\left(\mathbf{r}\right)}, && (2.13)\\
n\left(\mathbf{r}\right) &= \sum_{j=1}^{N}\left|\psi_j\left(\mathbf{r}\right)\right|^2. && (2.14)
\end{aligned}$$

The three terms in the effective potential $v_{\text{eff}}\left(\mathbf{r}\right)$ are the ionic potential, the average electrostatic (Hartree) potential and the exchange-correlation potential $v_{xc}\left(\mathbf{r}\right)$, which includes all the many-body interactions of the electrons as a single-variable function. Once these equations are self-consistently solved to yield the set of wavefunctions $\left\{\tilde{\psi}_j\left(\mathbf{r}\right)\right\}$, we compute the ground-state electron density as

$$n_0\left(\mathbf{r}\right)=\sum_{j=1}^{N}\left|\tilde{\psi}_j\left(\mathbf{r}\right)\right|^2 \tag{2.15}$$

and the ground-state total energy as

$$E_0=-\frac{1}{2}\sum_{j=1}^{N}\left\langle\tilde{\psi}_j\left|\nabla^2\right|\tilde{\psi}_j\right\rangle+\int d^3\mathbf{r}\,n_0\left(\mathbf{r}\right)v\left(\mathbf{r}\right)+\frac{1}{2}\int d^3\mathbf{r}d^3\mathbf{r}'\frac{n_0\left(\mathbf{r}\right)n_0\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}+E_{xc}\left[n_0\right]. \tag{2.16}$$

The core of DFT summarized above serves to transfer the problem from a multi-variable framework to a single-variable framework, without resorting to approximations. However, expressing the exchange-correlation energy functional requires approximations.

2.2 Exchange-correlation energy functional

The most common method to approximate the exchange-correlation energy is the local density approximation (LDA) [13, 14], where space is divided up into volume elements inside which there is a homogeneous electron gas at a certain density:

$$E_{xc}^{\mathrm{LDA}}\left[n\right] = \int d^3\mathbf{r}\, n\left(\mathbf{r}\right) \varepsilon_{xc}\left(n\left(\mathbf{r}\right)\right), \tag{2.17}$$

where$\varepsilon_{xc}\left(n\right)$ is the exchange-correlation energy of a homogeneous electron gas at density n, which has been computed numerically to high accuracy (and analytically in certain limits) [15, 14].

Approximations of $E_{xc}\left[n\right]$ are a well-studied subject and different approximations exist with different merits. LDA is not expected to work well if $n\left(r\right)$ varies rapidly, if there are regions where $n\left(r\right)$ is very low, or when the electrons are strongly correlated. Still it has been observed that LDA is able to calculate the following quantities quite accurately: Cohesive and binding energies, ground-state energy differences of similar configurations, lattice parameters and atomic positions [16].

In this work we mainly use a well-established improvement on LDA, called generalized gradient approximation (GGA) [17, 18]:

$$E_{xc}^{\mathrm{GGA}} = \int d^3\mathbf{r}\, n\left(\mathbf{r}\right) \varepsilon_{xc}^{\mathrm{GGA}}\left(n\left(\mathbf{r}\right), \left|\nabla n\left(\mathbf{r}\right)\right|\right), \tag{2.18}$$

where $\varepsilon_{xc}^{\mathrm{GGA}}\left(n, \left|\nabla n\right|\right)$ is the exchange-correlation energy of an electron gas with density n and density gradient ∇n. The expression for $\varepsilon_{xc}^{\mathrm{GGA}}$ is not unique, and improving on it is a topic of ongoing research. By including information about the gradient of n, this form of the functional can describe exchange-correlation effects better. GGA is known to improve the accuracy of binding energies and energy barriers compared to LDA [18].

2.3 Plane waves and pseudopotentials

Since the systems we study are periodic in space, we solve the Kohn–Sham equations iteratively in unit cells of our choosing, with the help of Bloch's theorem, and using plane waves as the basis:

$$n\left(r\right) = \sum_{n,\mathbf{k}} \left|\psi_{n\mathbf{k}}\left(\mathbf{r}\right)\right|^2, \tag{2.19}$$

$$\psi_{n\mathbf{k}}\left(\mathbf{r}\right) = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\Omega} \sum_{\mathbf{G}} c_{n\mathbf{k}}\left(\mathbf{G}\right) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{2.20}$$

where Ω is the volume of the unit cell, and $c_{n\mathbf{k}}\left(\mathbf{G}\right)$ are plane wave expansion coefficients. Because $c_{n\mathbf{k}}\left(\mathbf{G}\right)$ should go to zero as $|\mathbf{G}| \to \infty$, we can truncate the expansion at a pre-defined cutoff such that the sum is done over all $\mathbf{G}$ with $|\mathbf{G}| \leq G_{\text{cut}}$. The accuracy of the approximation monotonically increases with increasing G_{cut}. In our calculations G_{cut} is specified through E_{cut}, where $E_{\text{cut}} = {}^{G_{\text{cut}}^2}\!/_2$, and E_{cut} is treated as a convergence parameter.

The advantages of using a plane wave basis are that (i) the kinetic energy has a very simple expression in this basis, (ii) the periodicity of the crystal is included in the plane wave expansion itself, and (iii) systematic convergence of the total energy to a desired accuracy is guaranteed by increasing E_{cut}. Because this expansion is essentially a Fourier expansion, capturing rapidly varying regions of the wavefunction requires a large number of plane waves. However, these regions are typically in the vicinity of atomic nuclei, and in general do not need to be exactly reproduced since they are not critical for bonding. To describe these regions in a smooth way to save computational cost, we use pseudopotentials [19]. In calculations with pseudopotentials, core electrons are excluded from the treatment, and the potential on a valence electron is modified such that it fits the true potential outside a cutoff radius r_{cut} but is smoother inside the cutoff radius r_{cut}. Because chemical processes occur between nuclei rather than on the nuclei, we expect this approximation to accurately describe solids or molecules.

2.4 Calculation of forces

Given the atomic positions, we solve the Kohn–Sham equations with the help of plane waves and pseudopotentials to find the electronic ground state. However, in order to find the optimal atomic positions and/or lattice parameters, or to compute phonons, we need to compute the derivatives of the total energy with respect to ionic positions, i.e. the forces on ions. The force on ion J is

$$\mathbf{F}_J = -\frac{d}{d\mathbf{R}_J}\left(E_0 + E_{ii}\right), \tag{2.21}$$

where E_0 is given by equation (2.16) and E_{ii} is the sum of the electrostatic interaction energy between ion-ion pairs. The second term can be written as

$$E_{ii} = \frac{1}{2}\sum_{I\neq K}\frac{Z_J Z_I}{\left|\mathbf{R}_I - \mathbf{R}_K\right|} \tag{2.22}$$

and its derivative is straightforward to compute. The derivative of the first term is more involved, but can be simplified by expressing E_0 as a function of $\{\mathbf{R}_J\}$ and $\left\{\tilde{\psi}_j\right\}$. Because E_0 is *stationary* with respect to the wavefunctions $\left\{\tilde{\psi}_j\right\}$, its derivative with respect to $\mathbf{R}_J$ is equal to its *explicit* derivative with respect to $\mathbf{R}_J$ (Hellman–Feynman theorem [20]). Since the only term in equation (2.16) with explicit dependence on ionic positions is the second term which expresses the electron-ion interaction, the derivative becomes

$$\frac{dE_0}{d\mathbf{R}_J} = \frac{\partial E_0}{\partial\mathbf{R}_J} = \int d^3\mathbf{r}\, n_0\left(\mathbf{r}\right)\frac{\partial v\left(\mathbf{r}\right)}{\partial\mathbf{R}_J}. \tag{2.23}$$

Therefore the force on ion J is

$$\mathbf{F}_J = -\int d^3\mathbf{r}\, n_0\left(\mathbf{r}\right) \frac{\partial v\left(\mathbf{r}\right)}{\partial \mathbf{R}_J} - \frac{1}{2}\frac{\partial}{\partial \mathbf{R}_J} \sum_{I \neq K} \frac{Z_J Z_I}{\left|\mathbf{R}_I - \mathbf{R}_K\right|}, \tag{2.24}$$

which can be computed for a given ionic configuration and electron density without significant computational effort.

2.5 Nudged elastic bands method

On many occasions in this Thesis, we will be interested in transitions between related (meta)stable configurations. In transition state theory, the rate at which a transition between two configurations happens is given by

$$R = \nu e^{-\frac{E_b}{k_\mathrm{B} T}}, \tag{2.25}$$

where ν is the attempt frequency (usually of the order of a typical vibrational frequency), E_b is the energy barrier, k_B is Boltzmann's constant and T is the temperature [21, 22, 23].

The nudged elastic bands (NEB) method [24, 25] is a method to find the minimum energy path (MEP) and thus the energy barrier (E_b) between a given pair of configurations, enabling us to estimate the likelihood of a transition between two states at a given temperature. For a system with N particles, finding the MEP is equivalent to finding a path between two points in a $3N$ dimensional configuration space such that maximum value of the energy on the path is minimum. Between two (meta)stable states, i.e., energy basins in the configuration space, this is equivalent to finding a saddle point between the two basins. A saddle point in this $3N$ dimensional space is a minimum of energy along $3N - 1$ dimensions, and a maximum of energy along 1 dimension, namely the reaction coordinate.

In Figure 2.1, a simplified 2 dimensional energy landscape is presented with two basins. An NEB calculation takes these two states as an input and first generates intermediate "images" connecting these two states. This initial set of images is shown as circles on the straight dashed line in the right panel of the figure. Then these images are relaxed so that the transition path relaxes to the MEP. Because relaxing the images without any constraints would lead them to relax into a basin, the images are connected with fictitious springs with spring constant k. The images are relaxed in the direction perpendicular to the transition path according to the force written as

$$\mathbf{F}_{\perp}^{(i)} = -\nabla E\left(\mathbf{R}_i\right)|_{\perp} = -\nabla E\left(\mathbf{R}_i\right) + \nabla E\left(\mathbf{R}_i\right) \cdot \hat{\tau}_i, \tag{2.26}$$

where $E\left(\mathbf{R}_i\right)$ is the energy of the image i at position $\mathbf{R}_i$ in the configuration space and $\hat{\tau}_i$ is the normalized local tangent to the transition path. Along the transition path, the images are kept separate through the fictitious spring forces that can be written as

$$\mathbf{F}_{\|}^{(i)} = k\left(|\mathbf{R}_{i+1} - \mathbf{R}_i| - |\mathbf{R}_i - \mathbf{R}_{i-1}|\right)\hat{\tau}_i, \tag{2.27}$$

where k is the spring constant. The total force on an image during an NEB step is then

$$\mathbf{F}^{(i)} = \mathbf{F}_{\perp}^{(i)} + \mathbf{F}_{\|}^{(i)}. \tag{2.28}$$

This process relaxes the transition path onto the MEP; however, to locate the saddle point on the MEP to find the energy barrier E_b, one of the images is chosen to move along the MEP until its energy is maximized. This image, generally taken to be the image with highest energy after the initial relaxation onto the MEP, is called the "climbing image" [24, 25]. In Figure 2.1, the climbing image is shown as a green circle.

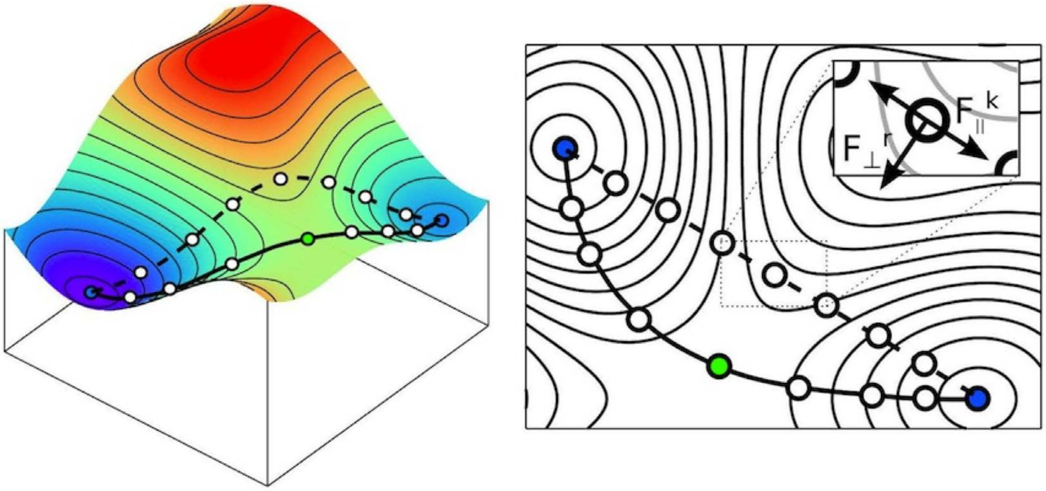


Figure 2.1: A simple two dimensional energy landscape with two (meta)stable states (blue circles) and the minimum energy transition path that can be calculated via the NEB method. For an image on the initial path shown as a straight dashed line in the right panel, the real perpendicular force $\mathbf{F}_\perp$ and the fictitious spring forces $\mathbf{F}_\parallel$ are shown schematically in the inset. After the NEB calculation is complete, the green circle sits on the saddle point and gives the energy barrier E_b. [The figure is copied with the owner's permission from http://umet.univ-lille1.fr/Projets/RheoMan/en/to-learn-more-about/nudged-elastic-band.php]

2.6 Simulation methods for lattice models

2.6.1 Monte Carlo algorithms for statistical physics

In this Thesis, we will describe the finite temperature behavior of some systems (see Chapter 5) using Monte Carlo simulations. A thermodynamic system is described by its partition function

$$Z = \sum_{\{s\}} e^{-\frac{E_s}{k_{\mathrm{B}}T}}, \tag{2.29}$$

where the sum runs over all possible states of the system, E_s is the energy of state s, k_{B} is Boltzmann's constant and T is the temperature. The expectation value of an observable X is

$$\langle X \rangle = \frac{1}{Z} \sum_{\{s\}} X_s e^{-\frac{E_s}{k_{\mathrm{B}}T}}, \tag{2.30}$$

where X_s is the value of the observable X when the system is in state s.

The summations are over all possible states of the system which is a space that is enormous for most physically relevant systems. However, most of the states occur with vanishingly small probabilities, computed by the formula $\frac{1}{Z}\exp\left(-\frac{E_s}{k_{\mathrm{B}}T}\right)$. Hence in order to avoid summing over all possible states, which is an intractable problem and a wasteful attempt, one usually uses *importance sampling*, in which the sampling is done over states that are chosen according to the probability distribution $\frac{1}{Z}\exp\left(-\frac{E_s}{k_{\mathrm{B}}T}\right)$ [26].

Given two states of the system and their energies, it is trivial to compute their relative probabilities according to their Boltzmann factors $\exp\left(-\frac{E_s}{k_{\mathrm{B}}T}\right)$. However, computing the absolute probability of a state requires computing Z, which we wish to avoid. The most commonly used way of computing expectation values without evaluating the partition function is by creating a Markov chain of states in which each state only depends on the state that immediately precedes it [27]. Starting from a configuration S_i with a Boltzmann factor p_i, a new trial configuration S_j with a

Boltzmann factor p_j is generated and accepted with probability π_{ij}. The probability of occupying the state S_j should be equal to the sum of the probabilities of arriving at state S_j from any given state S_i, i.e.

$$\sum_i p_i \pi_{ij} = p_j. \tag{2.31}$$

At equilibrium, this Markov process should obey *detailed balance*, i.e.

$$p_i \pi_{ij} = p_j \pi_{ji}. \tag{2.32}$$

In general, the transition probabilities π_{ij} are the product of two factors: a probability g_{ij} of proposing to move to state S_j from state S_i, and an acceptance ratio A_{ij} of accepting the proposed transition from S_i to S_j. Thus we can write

$$p_i g_{ij} A_{ij} = p_j g_{ji} A_{ji}, \tag{2.33}$$

or

$$\frac{g_{ij} A_{ij}}{g_{ji} A_{ji}} = \exp\left(-\frac{E_j - E_i}{k_{\mathrm{B}} T}\right). \tag{2.34}$$

For a given problem, g_{ij}, A_{ij} are specified by the algorithm such that equation (2.34) is satisfied and the sampling efficiency is maximized.

Finally, a valid Monte Carlo algorithm must be ergodic, i.e., any state must be reachable from any other state via a succession of moves.

2.6.2 Metropolis algorithms for discrete lattice models

The most common Monte Carlo algorithm for discrete lattice models such as the Ising model is the so-called Metropolis algorithm. Let us describe this algorithm in the context of our lattice model which we describe in more detail in Chapter 5.

The Hamiltonian of our two dimensional discrete lattice mode is

$$H = \sum_{i,j} E\left(\sigma\left(i,j\right)\right) + \sum_{i,j} J_x\left(\sigma\left(i,j\right), \sigma\left(i+1,j\right)\right) + \sum_{i,j} J_y\left(\sigma\left(i,j\right), \sigma\left(i,j+1\right)\right), \tag{2.35}$$

where (i,j) are the positions on the discrete lattice along the (x,y)-directions, $\sigma\left(i,j\right)$ is the state on lattice site (i,j), $E\left(\sigma\left(i,j\right)\right)$ is the site energy of the state $\sigma\left(i,j\right)$, $J_x\left(\sigma\left(i,j\right), \sigma\left(i+1,j\right)\right)$ is the nearest-neighbor interaction energy along the x-direction, and $J_y\left(\sigma\left(i,j\right), \sigma\left(i,j+1\right)\right)$ is the nearest-neighbor interaction energy along the y-direction. $J_x\left(\sigma_1, \sigma_2\right) = J_y\left(\sigma_1, \sigma_2\right) = 0$ if $\sigma_1 = \sigma_2$. In this model, there are N types of states, i.e. σ is a function that maps a lattice site onto one of $s_1, s_2, \ldots, s_N$. Note that the lower-case s are different from the upper-case S used above, which denoted the state of the whole system, which would be the collection of states s on all lattice points for this model.

The two dimensional Ising model is a special case of our model, where $N = 2$. The external magnetic field can be included by having $E\left(s_1\right) \neq E\left(s_2\right)$, and anisotropy can be included by having $J_x\left(s_1, s_2\right) \neq J_y\left(s_1, s_2\right)$.

The Metropolis algorithm would operate on our N-state model as follows:

1. Pick a lattice site at random. Let us call the state on the site s_i. Let us call the state of the initial system S_μ.

2. Propose to flip the state s_i to another state s_f, chosen among all non-s_i states with equal probability $\frac{1}{N-1}$. Let us call the state of the system if the proposed flip occurs S_ν. Thus $g_{\mu\nu} = \frac{1}{N-1}$ (see equation (2.34)). The probability of proposing the inverse move, i.e. going to S_μ from S_ν is clearly the same, hence $g_{\nu\mu} = g_{\mu\nu} = \frac{1}{N-1}$.

3. Compute the energy difference $E_\nu - E_\mu$ between S_ν and S_μ. This is simple, since the only difference is the state change of state s_i to s_j, and the energy difference

is localized to the site energy and the couplings with the nearest neighbors of that site.

4. The acceptance ratios are obtained by equation (2.34):

$$\frac{A_{\mu\nu}}{A_{\nu\mu}} = \exp\left(-\frac{E_\nu - E_\mu}{k_\mathrm{B}T}\right). \tag{2.36}$$

   A common way of achieving this equation is by setting:

$$A_{\mu\nu} = \begin{cases} \exp\left(-\frac{E_\nu - E_\mu}{k_\mathrm{B}T}\right) & \text{if } E_\nu > E_\mu \\ 1 & \text{if } E_\nu \leq E_\mu \end{cases} \tag{2.37}$$

To find the expectation value of an observable X, X is computed at each step of the simulation that comprises of a finite number of steps, and then simply averaged. This is the merit of *importance sampling*, which takes care of the relative probabilities of states through equation (2.34), therefore the observables can simply be averaged.

2.6.3 Wolff cluster algorithms

The success of a Monte Carlo algorithm is usually measured by how easy it can generate “independent” samples, i.e. how many attempts it takes to go from a state S_μ to another state S_ν such that S_μ and S_ν are “uncorrelated” (namely, decorrelation time). The “single-flip” Metropolis algorithm is conceptually simple and easy to implement. However, at each simulation step, the state only slightly changes, so the decorrelation time can be large. For models with a second order phase transition, such as the two dimensional Ising model, this algorithm suffers from “critical slowing down” where, close to the critical temperature of the model, the decorrelation time diverges [28].

This issue can be solved by algorithms that propose states that are sufficiently modified from the preceding state. A family of such algorithms is called cluster

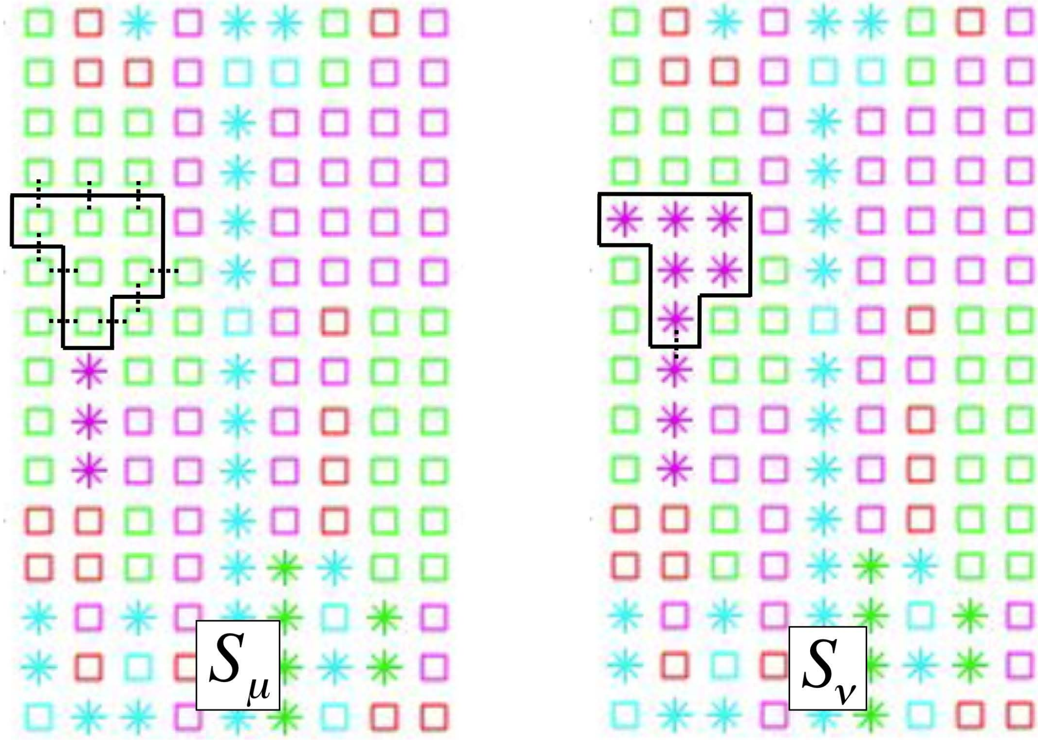


Figure 2.2: A sample instant of a Wolff cluster simulation of an N-state lattice model, prior to (S_μ) and after (S_ν) the switching of a cluster. The boundary of the cluster is shown by solid lines, and the bonds at the boundary of the cluster are shown by dotted lines. Each color-shape combination denotes a type of state in our 10-state lattice model, described in detail in Chapter 5.

algorithms, where rather than switching the state on a single site, the state on a groups of sites ("a cluster") is switched simultaneously [29]. Here we modify the Wolff cluster algorithm [30], originally developed for the Ising model, to simulate our N-state model:

1. Pick a lattice site i at random. Let us call the state on the site s_i. Let us call the state of the initial system S_μ.

2. Add each of the nearest neighbors j of the site i to the cluster, with the probability p_{add}, provided that the states on sites i and j are the same, and the "bond" between i and j has not yet been considered.

3. Once all the neighbors of site i have been considered, move to the next site in the cluster. Repeat step 2 for this site. If all the sites in the cluster have gone through step 2, the cluster has been built. Move to step 4.

4. Propose to flip the state s_i to another state s_f, chosen among all non-s_i states with equal probability $\frac{1}{N-1}$, for all the sites in the cluster. Let us call the state of the system if the proposed flip occurs S_ν.

5. Compute the number of bonds at the boundary of the cluster. The two neighboring states of the same kind are said to have a bond that is intact. When the cluster is "flipped" the bonds at the boundary will be broken. In Figure 2.2, we illustrate the formation of a cluster for a given state S_μ of the lattice, shown on the left. The number of bonds at the boundary (shown as dotted lines in the figure) is $n_\mu = 9$. The proposed state S_ν is shown on the right. The number of bonds at the boundary in the proposed state is $n_\nu = 1$.

6. Compute the energy difference $E_\nu - E_\mu$ between S_ν and S_μ. This requires accounting for all the nearest-neighbor interactions at the boundary of the cluster in both the initial and the final states.

Finding the correct acceptance ratio for this algorithm is somewhat involved. Let us assume that the cluster in S_μ in Figure 2.2 is built in the following order:

1. The site at the upper left corner of the cluster is randomly picked.

2. The site to the right is added with probability p_{add}, the other neighboring sites of the same kind (above and below) are rejected with probability $(1 - p_{\text{add}})^2$.

3. The site to the right is added with probability p_{add}, the other neighboring sites of the same kind (above and below) are rejected with probability $(1 - p_{\text{add}})^2$.

4. The site below is added with probability p_{add}, the site above is rejected with probability $(1 - p_{\text{add}})$.

5. The site to the left is added with probability p_{add}, the other neighboring sites of the same kind (to the right and below) are rejected with probability $(1 - p_{\text{add}})^2$.
6. The site below is added with probability p_{add}, the site to the left is rejected with probability $(1 - p_{\text{add}})$.
7. Both neighboring sites (to the left and to the right) are rejected with probability $(1 - p_{\text{add}})^2$.

The total probability of this process in this order is $p_{\text{add}}^5 (1 - p_{\text{add}})^{10}$. The same process can be repeated for the cluster in S_ν in Figure 2.2 built in the exact same order, which yields a probability of $p_{\text{add}}^5 (1 - p_{\text{add}})^2$.

The ratio of proposal probabilities of the forward and backward moves is then

$$\frac{p_{\text{add}}^5 (1 - p_{\text{add}})^{10}}{p_{\text{add}}^5 (1 - p_{\text{add}})^2} = (1 - p_{\text{add}})^8 , \tag{2.38}$$

where 8 is the difference in the number of bonds at the boundary for S_μ and S_ν, i.e. $n_\mu - n_\nu = 8$. It is evident that for any given order for building the same cluster, the ratio of proposal probabilities of the forward and backward moves will be $(1 - p_{\text{add}})^{n_\mu - n_\nu}$. Because $g_{\mu\nu}$ is the sum of the probabilities of all moves that propose S_ν from S_μ and $g_{\nu\mu}$ is the sum of the probabilities of all moves that propose S_μ from S_ν, we can write

$$\frac{g_{\mu\nu}}{g_{\nu\mu}} = (1 - p_{\text{add}})^{n_\mu - n_\nu} . \tag{2.39}$$

Therefore equation (2.34) yields

$$\begin{aligned} \frac{A_{\mu\nu}}{A_{\nu\mu}} &= (1 - p_{\text{add}})^{n_\nu - n_\mu} \exp\left(-\frac{E_\nu - E_\mu}{k_{\text{B}} T}\right) \\ &= \exp\left(-\frac{E_\nu - E_\mu - k_{\text{B}} T (n_\nu - n_\mu) \log(1 - p_{\text{add}})}{k_{\text{B}} T}\right) . \end{aligned} \tag{2.40}$$

If we define

$$\Delta_{\mu\nu} \equiv E_\nu - E_\mu - k_{\mathrm{B}}T\left(n_\nu - n_\mu\right)\log\left(1 - p_{\mathrm{add}}\right), \tag{2.41}$$

we can set the acceptance ratios (in analogy with equation (2.37)) to be

$$A_{\mu\nu} = \begin{cases} \exp\left(-\frac{\Delta_{\mu\nu}}{k_{\mathrm{B}}T}\right) & \text{if } \Delta_{\mu\nu} > 0 \\ 1 & \text{if } \Delta_{\mu\nu} \leq 0 \end{cases}. \tag{2.42}$$

In the original Wolff cluster method for the Ising model, p_{add} is defined as a function of temperature such that the acceptance ratios are always 1. This makes for a rejection-less algorithm which is able to switch clusters of different sizes at any temperature. However, in our model there is no simple relationship between $E_\nu - E_\mu$ and $n_\nu - n_\mu$ as in there is in the Ising model. Therefore p_{add} cannot be defined *a priori* to make $\Delta_{\mu\nu}$ vanish in equation (2.41), which in turn would guarantee $A_{\mu\nu} = 1$ in 2.42. After empirical tests on our simulations, we have set $p_{\mathrm{add}} = \frac{1}{2}$ for the results presented in Chapter 5. Improving the acceptance ratios through the choice of p_{add} is the subject of future research.

Chapter 3

Engineered unique elastic modes at a $BaTiO_3/(2 \times 1)$-Ge(001) interface

Note: This chapter has been adopted from an article we have published in *Physical Review Letters* [11].

3.1 Introduction

Strain coupling in perovskite-perovskite systems has led to the control and enhancement of a wide range of physical properties including metal-insulator transitions in the rare earth nickelates [31, 32, 33, 34], ferroelectricity and ferromagnetism in the titanates [35, 36, 37], and improper ferroelectricity in titanate multilayers [38]. Epitaxial constraints are particularly important in controlling the physical properties of most perovskite transition metal oxides whose properties are strongly coupled to local and long-range structural distortions of the oxygen octahedral sublattice [2, 39]. These properties can be continuously tuned by modifying the amplitudes of the lattice distortions. However, by coupling to more complex interfacial degrees of freedom that go beyond simple elastic strain, electronic and magnetic effects can potentially be coupled in order to give rise to multiferroic effects and complex orderings of electronic charge and spin [40, 41, 42]. Recent advances in the epitaxial growth of perovskite oxides on semiconductors such as Si [43, 44, 45, 46], Ge [47, 48, 49], and GaAs [50, 51] allow for structural coupling to even more complex symmetry degrees of freedom that are not accessible in standard perovskite-perovskite epitaxy [52]. Separately, in a more technological vein, the growth compatibility of complex oxides with semiconductors has allowed the monolithic integration of the functional properties of oxides with well-established complementary metal-oxide-semiconductor technologies [53].

In this chapter, we demonstrate the design of a non-bulk structural phase of the transition metal ferroelectric $BaTiO_3$ (BTO) using an interfacial structural motif that strongly couples to symmetry-lowering structural distortions in the BTO. Bulk BTO has a tetragonal P4mm structure at room temperature with an in-plane lattice constant of 3.99 Å and a ferroelectric polar distortion along its c axis [54]. We report on lattice distortions in thin BTO films grown on the reconstructed Ge(001) surface with a lower Pm planar symmetry arising from asymmetric Ge surface dimer rows [55, 56]. The net result is a strong interfacial coupling to in-plane symmetry-breaking distortions in the BTO that stabilize a non-bulk and distorted BTO interfacial phase. The non-bulk features of this structure include breathing mode distortions of the oxygen octahedra and cation rumpling. This phase leads to an in-plane component of the polarization in the BTO in addition to significant oxygen octahedral distortions: in other complex transition metal perovskite oxide systems, such distortions are associated with charge, spin, and orbital ordering [57, 58, 59]. While the in-plane polarization can be related to the room temperature stabilization of the bulk low-temperature orthorhombic BTO phase, the period-doubling octahedral distortions that we observe are unique to the BTO/Ge interface and are associated with stiff bulk modes that are stabilized by the forces exerted by the reconstructed Ge surface. Using a combination of synchrotron X-ray diffraction, transmission electron microscopy, and density functional theory calculations, we show that the observed distortions can be parametrized by considering the lattice distortion modes excited by the symmetry of the imposed epitaxial constraint.

3.2 Experimental findings

Thin films of BTO with thicknesses of 2.5 and 5.5 unit cell (u.c.) are grown on single crystalline Ge(001) substrates using reactive molecular beam epitaxy [46, 48, 3, 60]. A 0.5 monolayer of Ba is deposited at 440 °C on the Ge surface to passivate the surface and prevent the oxidation of the Ge surface. The wafer is then cooled to room temperature where an additional 0.5 monolayer of Ba is deposited, followed by co-deposition of two monolayers each of BaO and TiO_2. The low-temperature deposition leads to the formation of amorphous BTO as evidenced by a diffuse RHEED pattern. The wafer is then heated to 500 °C in UHV where the amorphous BaO and TiO_2 react to form crystalline $BaTiO_3$. To create the 5.5 u.c. sample, three additional monolayers of BaO and TiO_2 are co-deposited at a substrate temperature of 570 °C. We note that the higher temperature deposition of the three additional layers for the 5.5 u.c. sample may lead to changes in the stoichiometry and interfacial structure as discussed below. Finally, for the 2.5 u.c. sample, a 7-nm-thick capping layer of amorphous Ge is deposited at room temperature to protect the ultra thin $BaTiO_3$ film.

To characterize the structure of the BTO films, transmission electron microscopy measurements are carried out at Brookhaven National Laboratory. Figure 3.1(a) and Figure 3.1(b) show high angle annular dark field images obtained along Ge [110] and Ge [100] zone axes, respectively, for the 2.5 u.c. BTO/Ge sample. The images indicate an abrupt interface with coherent atomic registry between the BTO film and the Ge. As explained below, the reduced contrast for the interfacial Ge layer is attributed to a Ge reconstruction and not to structural disorder, in agreement with Fredrickson et

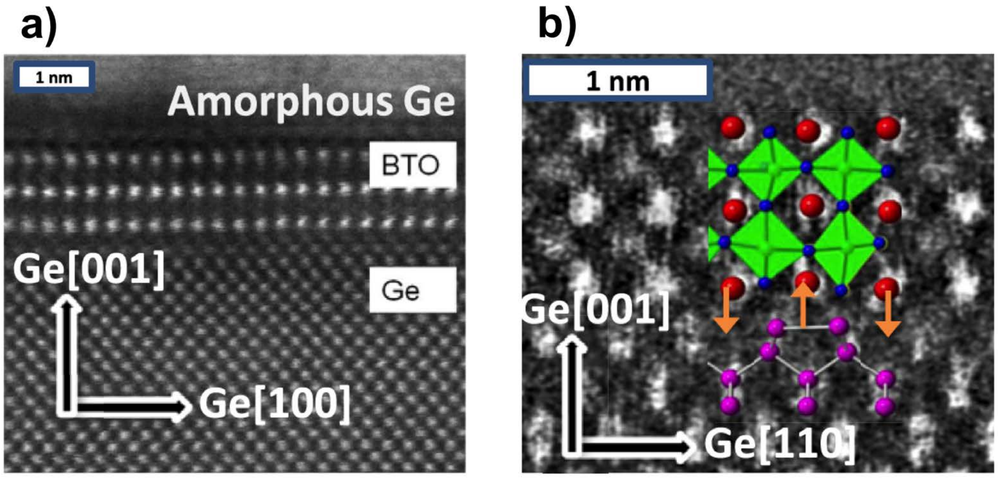


Figure 3.1: (a) High angle annular dark field TEM image of 2.5 u.c. BTO/Ge along the Ge [100] zone axis. (b) Superimposed on the TEM image taken along the Ge [110] zone axis is a model structure of the system determined from surface X-ray diffraction measurements. The orange arrows indicate the rumpling of the interfacial Ba along the [001] axis.

al. where the interface was prepared starting with a phase consisting of half a ML of Sr on Ge(001) [49].

To confirm the ordered nature of the interface and the effect of the coupling across the interface, we use synchrotron X-ray diffraction measurements to determine the atomic structure of the system: atomic-scale structure of the samples is determined from the analysis of crystal truncation rods (CTRs) [61]. In addition to integer-order rods observed for both the 2.5 and 5.5 u.c. samples, superstructure reflections consistent with a 2×1 symmetry with respect to the bulk Ge are observed for the 2.5 u.c. sample. From the periodicity of the superstructure reflections in the direction in reciprocal space perpendicular to the film-substrate interface, the 2×1 symmetry is determined to extend from the interface into the BTO film.

Atomic positions are derived from the measured CTRs and superstructure reflections by fitting using a chi-square minimization genetic fitting algorithm program, GenX [62]. For this analysis, we consider the coherent average of two 90° rotational

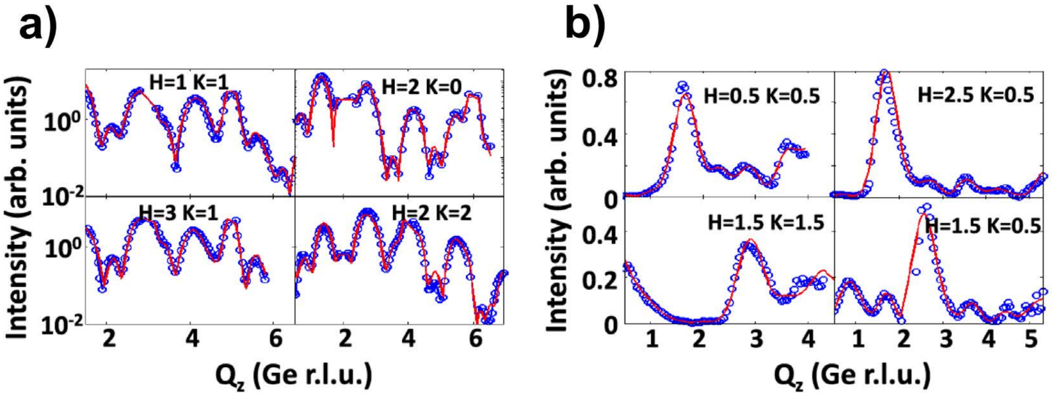


Figure 3.2: (a) Integer-order (logarithmic vertical scale) and (b) superstructure (linear vertical scale) crystal truncation rods measured (blue circles) for a 2.5 u.c. BTO film grown on Ge(001) using molecular beam epitaxy. The model structure of the BTO/Ge interface shown in Figure 3.1 (b) is derived from the fits (red lines) to the measured data.

domains that arise due to atomic layer steps on the Ge(001) surface. Figure 3.2 shows measured diffraction intensities along (a) integer and (b) half-order rods, and their associated fits for the converged 2.5 u.c. sample. The XRD-determined structure of the 2.5 u.c. sample is shown superimposed on the TEM image of the same sample in Figure 3.1 (b). The main features of the XRD structure are well reproduced in the TEM image: i.e., the rumpling of a full monolayer of Ba at the interface and the displacement of Ti along the BTO [100] direction. The O positions, which are not resolved in the TEM image, are determined from the X-ray CTR analysis.

In contrast to the 2.5 u.c. film, for the thicker 5.5 u.c. sample the CTR analysis and the absence of superstructure Bragg reflections show that the entire system of BTO and interfacial Ge have a 1×1 symmetry. The 5.5 u.c. BTO film has lattice constants $a = b = 3.99$Å and $c = 4.12$Å in agreement with those of bulk tetragonal BTO. Fits to the CTRs for the 5.5 u.c. film are shown in Figure 3.3. This points to a link explored below between the growth procedure for the thicker film and the manifestation or suppression of the interfacial distortions in the BTO.

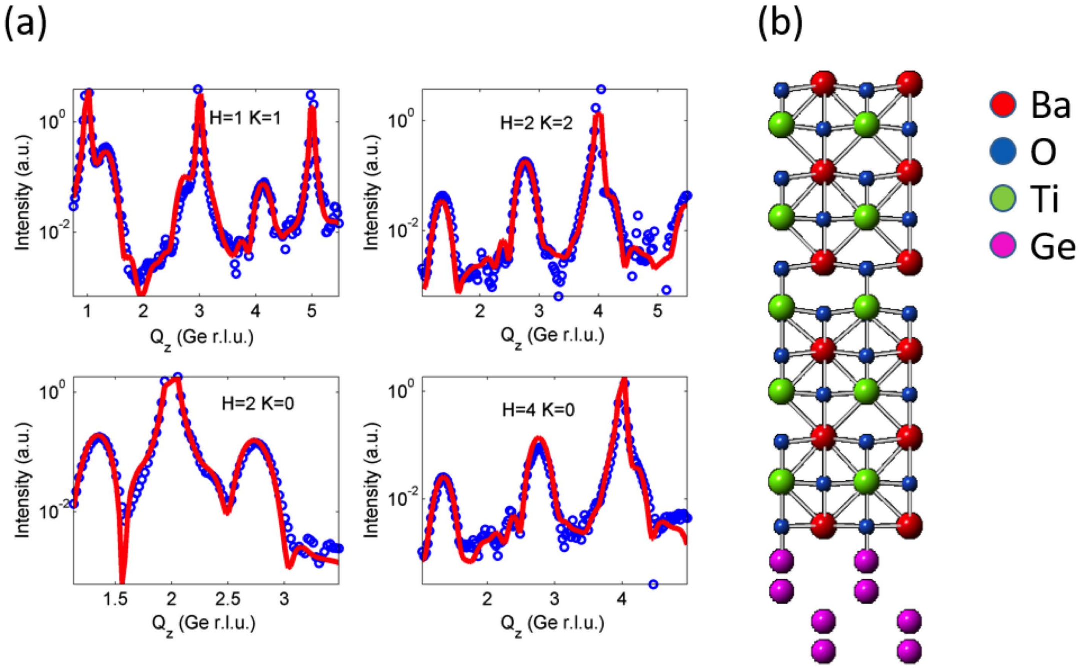


Figure 3.3: (a) Measured crystal truncation rods (blue circles) and associated fits (solid red lines) for a 5.5 u.c. $BaTiO_3$ film grown on Ge(001). (b) The atomic configuration of the BTO film derived from the CTR data.

Recent results show that a diverse range of interface structures can be stabilized for interfaces with varying cation and oxygen stoichiometry. BTO films grown with a higher alkaline earth metal content show that the Ge dimers are stable, even for thicker films, using a 0.5-ML Sr template layer on Ge [49]. In addition, prior theoretical calculations on the similar $SrTiO_3$/Si interface indicate a strong dependence of the dimerization of the interfacial semiconductor layer on both the cation and oxygen stoichiometry, where low oxygen concentrations at the $SrTiO_3$/Si interface destabilize the Si dimers resulting in a 1×1 interface [63]. We have performed systematic theoretical calculations (see section 3.4.2 for full details) to clarify the effects of oxygen stoichiometry and film thickness on the relative stabilities of 1×1 and 2×1 BTO/Ge interface structures. In brief, oxygen deficiency stabilizes the 1×1 interfacial structure: this is linked to electron doping of the interface (due to oxygen vacancy formation) and concomitant weakening of Ge-Ge dimers caused by electron occupation of anti-bonding states of the dimerized Ge layer. This observation means that experimentally controlling the interfacial oxygen stoichiometry at oxide-semiconductor interfaces can stabilize functional structural distortions.

3.3 Computational methods

We perform density functional theory (DFT) simulations, using the Perdew–Burke–Ernzerhof generalized gradient approximation (PBE GGA) to approximate the exchange-correlation functional [18] along with ultrasoft pseudopotentials [19], an in-plane 4×8 Monkhorst–Pack k-point sampling, a plane-wave energy cutoff of 35 Ry, and a Marzari–Vanderbilt smearing [64] with a width of 0.02 Ry. We employ the QUANTUM ESPRESSO software package [65]. We use a 2×1 supercell consisting of 8 layers of Ge and 2.5 u.c. of BTO on the top (100) surface of Ge, whereas the bottom Ge surface is passivated with hydrogen. Due to the asymmetric nature of the supercell, we use a dipole correction to prevent periodic copies of the supercell in the z-direction from having artificial electrostatic interactions. This self-consistent correction is achieved by introducing a fictitious dipole deep into the vacuum whose magnitude is automatically chosen to cancel out the potential difference caused by the film, leading to a flat potential in the vacuum [66]. The in-plane lattice constant is fixed to the computed one for bulk Ge at 4.07 Å. All atoms in the supercell, except the bottom two layers of Ge which are fixed to their bulk positions, are relaxed until atomic forces are smaller than $10^{-4}\mathrm{Ry}/a_0$ along all axial directions.

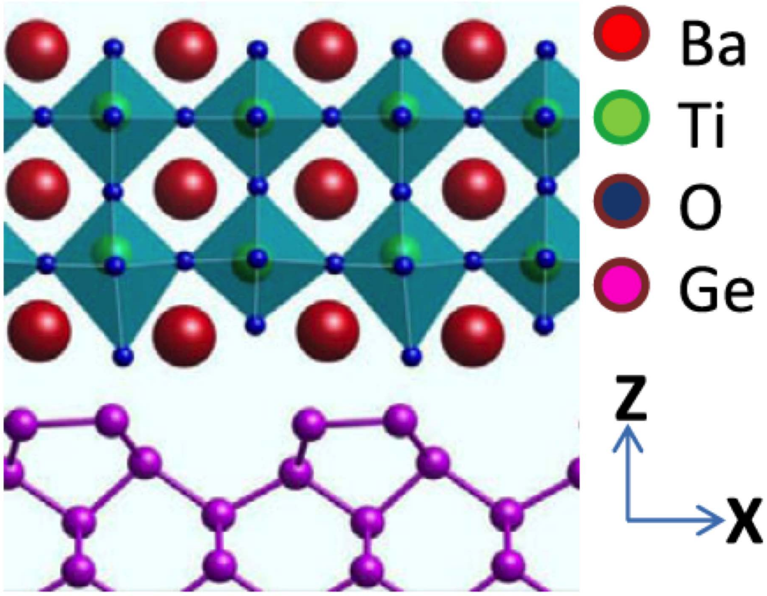


Figure 3.4: DFT-predicted ground state structure of the 2.5 u.c. thick BTO/Ge system with 2×1 periodicity. TiO_6 octahedra are shown as semi-transparent cages.

3.4 Results

3.4.1 Elastic modes at the BTO/Ge interface

We have first confirmed the stability of the observed structure for the 2.5 u.c. sample from first-principles calculations. The ground-state relaxed atomic coordinates of the BTO/Ge system are determined theoretically (see Figure 3.4) and are found to be in agreement with the experimentally measured structure (shown in Figure 3.1).

To understand the BTO structural response for the 2.5 u.c. thick films, we consider the lattice modes associated with bulk BTO [67]. The elastic response of atom l of the BTO lattice to a force $F_{l\alpha}$ is given by $F_{l\alpha} = \sum_{l'\beta} K_{l\alpha,l'\beta} u_{l'\beta}$, where $K_{l\alpha,l'\beta}$ is the interatomic force constant matrix connecting the force along Cartesian direction α on atom l to the displacement $u_{l'\beta}$ of atom l' in direction β. The interatomic force constants are calculated for a reference bulk phase of BTO using *ab initio* density functional perturbation theory calculations [20]. We then diagonalize the force constant matrix to obtain a set of static normal modes or eigenmodes. By projecting the computed BTO lattice distortions in the heterostructure (relative to the undistorted bulk BTO reference structure) onto these static eigenmodes, the distortions are described in the basis of eigenmodes of the bulk force constant matrix that immediately identifies unstable versus stiff modes. We have performed this analysis using the eigenmodes of the cubic high-symmetry phase of bulk BTO. We choose this reference phase for simplicity. We have explicitly verified that other choices of reference (e.g., tetragonal phase or strained bulk phases) yield similar results: the largest projections are always onto the soft modes at zero wave vector.

Figure 3.5 shows eigenmode dispersion curves along high-symmetry directions in reciprocal space for BTO in its high-symmetry cubic perovskite structure. An $8 \times 8 \times 8$ q-point mesh is used to sample the Brillouin zone. Intermediate q-points are determined by Fourier interpolation. Modes with imaginary eigenvalues are soft or unstable. The circles in Figure 3.5 indicate the magnitude of the projections of the atomic displacements in the 2.5 u.c. thick BTO slab onto the eigenmodes. Displacement patterns and projections are calculated for each 1 u.c. thick BTO that can be extracted from the BTO slab. Figure 3.5 (a) shows the projections for the two atomic layers (BaO and TiO_2) which form the first BTO u.c. adjacent to the interface. Each atom is distorted away from a reference configuration given by a pair of strained cubic cells (joined in the x-direction to make a $2 \times 1 \times 1$ cell) in which the atoms are at the high-symmetry locations of the perovskite structure. The geometric center of this reference cell is determined such that it coincides with the center of the atomic positions in the particular $2 \times 1 \times 1$ cell that is taken from the BTO film. This choice for the center makes the projections onto the acoustic modes at Γ vanish. (For more information regarding the choice of reference cell, see Section §3.5.) We decompose the displacements in a $2 \times 1 \times 1$ cell into the displacement modes obtained from a $1 \times 1 \times 1$ cell, which corresponds to the modes at the Γ and the X points of

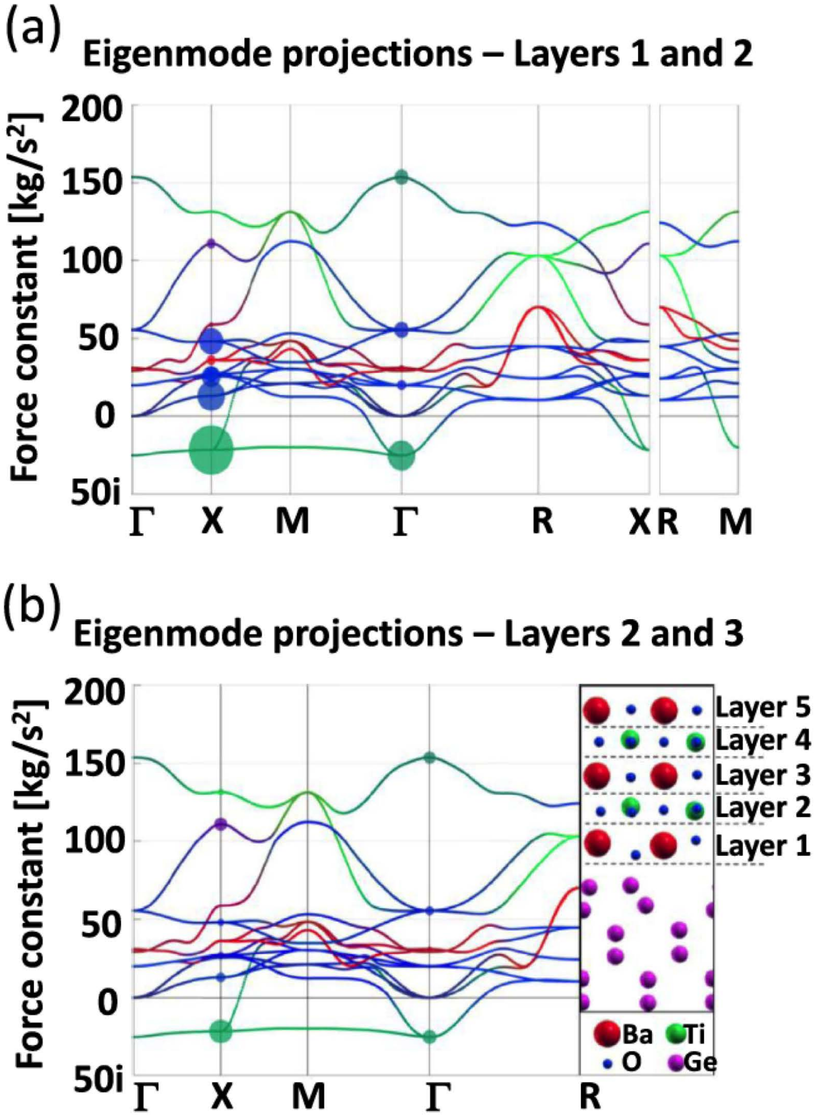


Figure 3.5: Force constant matrix eigenmode dispersion curves in reciprocal space for bulk cubic $BaTiO_3$. The colors correspond to the atomic content of each mode: red, green and blue correspond to amplitudes of motion of Ba, Ti and O atoms, respectively, following Ghosez et al. [67]. Solid circles superimposed on the dispersion curves express the decomposition of atomic displacements at the interface onto modes with Γ and X symmetry. The radii of the circles are proportional to the magnitude of the projections of the atomic displacements in the 2.5 u.c. thick BTO slab onto the eigenmodes. (a) The projections of the displacements of the first BaO and TiO_2 layers adjacent to the interface. (b) The projections for the first TiO_2 layer and the second BaO layer above it. The theoretical ground state of the BTO/Ge system is reproduced in the inset with the numbering of the atomic layers and the color coding of the atom types.

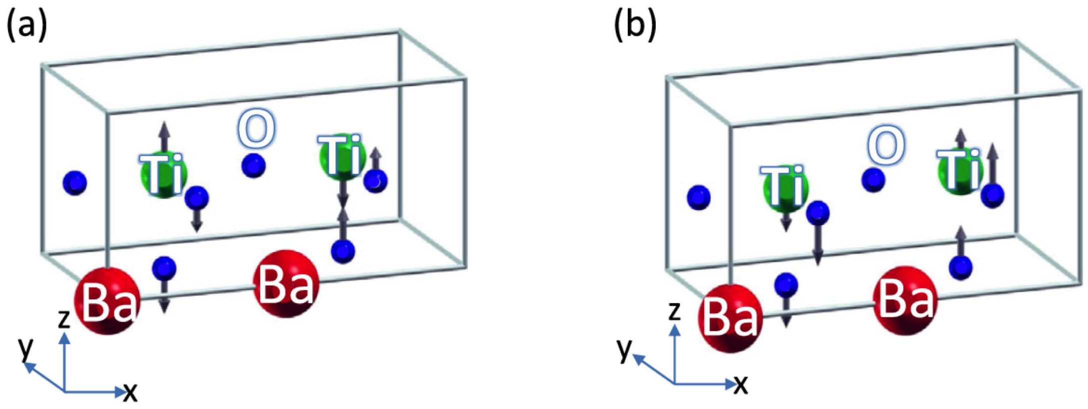


Figure 3.6: (a) Atomic motions for force constant matrix eigenmode contributing the most to the displacements of the 2.5 u.c. thick BTO slab. The mode is soft at the X point so that it involves anti-parallel displacements of planar O and Ti. (b) Lowest energy stiff acoustic mode at X, where the Ti and O atoms move in the same direction.

the $1 \times 1 \times 1$ cell (30 projections in total). Figure 3.5 (b) shows the projections for the next 1 u.c. thick subset of the slab. Away from the interface, the projections reduce in amplitude, but the dominant modes remain the same. As expected, the unstable modes with displacements normal to the interfacial plane account for most of the observed structural distortions.

In Figure 3.6, the two leading eigenmodes (two largest circles at the X point in Figure 3.5 (a)) for the interfacial BTO are shown. Figure 3.6 (a) shows the most dominant displacement mode that involves out-of-plane O and Ti motion, where the Ti atoms move in the opposite direction as the O atoms that are on the same (011) plane. Shown in Figure 3.6 (b) is a stiff acoustic mode, which is the next leading mode where the Ti atoms move in the same direction as the O atoms.

A significant part of the lattice distortion involves zone center distortions about the Γ point that result in nonzero polarization in the x- and z-directions, which corresponds to the bulk orthorhombic phase stable between 183 and 278 K (see Section §3.5 for a discussion of projections of BTO bulk orthorhombic displacements).

The ionic polarization $\vec{P}$ of the interfacial layer of the 2.5 u.c. film is determined from the summation $\vec{P} = ({}^{e}/_{\Omega}) \sum_i Z_i^* x_i$, where Z_i^* and x_i are the Born effective charges computed for a reference strained cubic bulk BTO, as discussed above, and the displacements from the reference position of the ith ion, respectively. Ω is the volume of the $2 \times 1 \times 1$ u.c. over which the summation is carried out. The spontaneous ionic polarization of the first interfacial BTO u.c. is estimated to be 54 $\mu\text{C/cm}^2$ along the BTO $\{101\}$ axis, which is comparable to the bulk polarization for orthorhombic BTO of 36 $\mu\text{C/cm}^2$ [68]. The large value may be a consequence of strain imposed by the Ge substrate. The lattice parameter of Ge (3.99 Å) is smaller than the a and b parameters of orthorhombic $BaTiO_3$ and larger than the c axis of orthorhombic $BaTiO_3$, increasing the orthorhombic distortion and the larger observed polarization. We note that since an equivalent BTO structure is obtained by applying a reflection operation to the system through the yz-plane, the in-plane component of the polarization does not have a preferred direction and, hence, in principle, may be switched. In practice, however, the computed energy barrier (using the nudged elastic bands method [25]) of 0.6 eV per 2×1 cell is on the large side.

3.4.2 Effects of film thickness

We have performed theoretical computations to understand the relative stability of the 1×1 and 2×1 reconstructions of the BTO/Ge interface as a function of BTO thickness and the oxygen content of the interface. One additional consideration is the strain state of the BTO. Experimentally, the lattice constant of tetragonal BTO is almost a perfect match to that of Ge ($a_{\text{ex}} = 3.99$Å), but theoretically our DFT optimized bulk lattice parameters are 3.99Å for BTO and $a_{\text{th}} = 4.07$Å for Ge. Thus, using the theoretical Ge lattice parameter puts tensile strain on the BTO that is not representative of the experiment (in fact no choice of theoretical lattice parameter reproduces the precise experimental situation). Therefore, we have performed a sep-

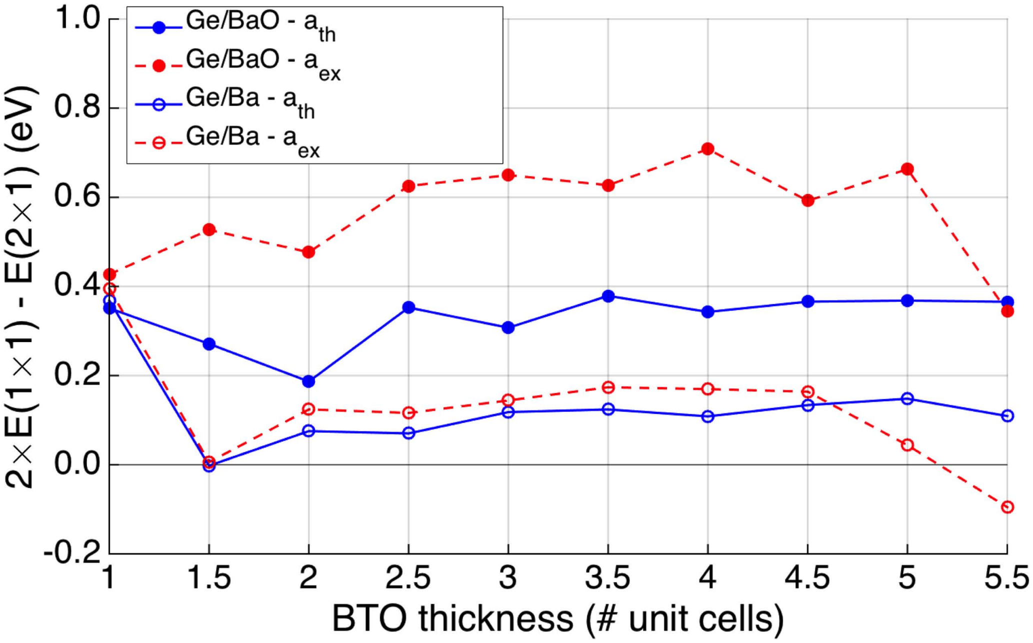


Figure 3.7: Total energy of the BTO/Ge system with 1 × 1 periodicity subtracted from the 2 × 1 periodicity with the same stoichiometry, as a function of BTO film thickness. Energies are in eV per 2 × 1 in-plane unit cell, and BTO thicknesses are in unit cells of BTO. For reference, the experimentally grown thicknesses are 2.5 u.c. and 5.5 u.c. The lines with solid markers correspond to interfaces with a fully oxygenated BaO layer adjacent to Ge, while hollow markers correspond to de-oxygenated Ba layers adjacent to Ge (i.e., 1 ML Ba on Ge(001)). The energies are reported for the theoretically optimized Ge substrate lattice constant of $a_{\text{th}} = 4.07\text{Å}$ (blue solid lines) and at the experimental lattice constant $a_{\text{ex}} = 3.99\text{Å}$ (red dashed lines).

arate set of computations using the experimental lattice parameters (a_{ex}) for Ge to investigate the strain dependence of the energies. The results of a total of 40 fully relaxed calculations are shown in Figure 3.7.

At 5.5 unit cells, we calculate that the 1 × 1 structure is more stable than the 2 × 1 structure, as observed experimentally, consistent with the main trend shown in Figure 3.7. That is, the introduction of oxygen vacancies at the interface strongly reduces the energy difference between 1 × 1 and 2 × 1 periodicities (i.e., they stabilize the 1 × 1 structure). We attribute this reduction partly to the fact that O vacancies donate electrons to the interface that enter into the higher energy anti-bonding states of the interfacial Ge, which weakens the Ge dimerization in the 2×1 case, whereas the 1×1 structure has more dangling orbitals available at lower energies to accommodate the added electrons. This view is corroborated by the fact that Ge dimers become 12% longer in the 2×1 oxygen deficient interfaces compared to the stoichiometric interfaces.

These trends strongly support the view that de-oxygenation of the interface stabilizes the 1×1 structure. However, a more detailed study dealing with the dependence of the energy difference on the exchange-correlation approximation as well as further details (e.g., why in certain cases in Figure 3.7 the energies become negative) is the subject future work.

3.5 Conclusion

In summary, we have shown how the symmetry of an epitaxial constraint can induce a phase of $BaTiO_3$ not found in the bulk. The structure of this phase is a seemingly complex distortion of the perovskite structure at an interface that is accurately calculated by applying first-principles theory to a many-atom model of the interface. We have reduced this complexity by showing that many of the structural details of the specific structure are encoded in the bulk force constant matrix of $BaTiO_3$ and expressed when coupled to the 2×1 symmetry of the Ge substrate. This observation suggests an approach to design materials properties by using a catalog of force constant matrices to identify materials with specific structural motifs. Substrates with symmetries that couple to soft modes of the material can then be chosen that enhance these motifs. For example, replacing the B-site Ti ions with other transition metal ions such as Mn and Ni, where breathing-mode distortions are associated with charge ordering, is likely to result in interfacial structural distortions similar to those measured here for BTO on Ge that will lead to charge “stripes,” metal-insulator transitions, or enhanced coupling with ferroelectric polarization.

Appendix: Theoretical details regarding the elastic modes

To understand the polar distortions observed in the interfacial $BaTiO_3$ layers, we project the displacements in the orthorhombic phase of bulk BTO from the high symmetry cubic configuration onto the bulk cubic BTO eigenmodes of the force constant matrix.

We consider the lattice modes associated with bulk BTO. The elastic response of atom l of the BTO lattice to a force $\vec{F}_l$ is given by $\vec{F}_l = -\sum_{l'} \overleftrightarrow{K}_{ll'} \vec{u}_{l'}$, where $\overleftrightarrow{K}_{ll'}$ is the interatomic force constant matrix, $\vec{u}_{l'}$ are the atomic displacements of atom l', and l' sums over all atoms. The interatomic force constants are calculated using first principles density functional perturbation theory. Due to the linear relationship, $\vec{u}_l$ can be determined from the inverse equation

$$\vec{u}_l = -\sum_{l'} \overleftrightarrow{K}^{-1}_{ll'} \vec{F}_{l'} \tag{3.1}$$

which means that observing displacements directly reflects imposed forces. To derive the eigenmode decomposition, we repartition the variables and sum to go over unit cells (instead of individual atoms):

$$\vec{u}_{\vec{R}} = -\sum_{\vec{R}'} \overleftrightarrow{K}^{-1}_{\vec{R}\vec{R}'} \vec{F}_{\vec{R}'} \tag{3.2}$$

where $\vec{u}_{\vec{R}}$ ($\vec{F}_{\vec{R}}$) is a 15×1 vector that describes the displacements of (forces on) all of the five atoms in the unit cell at $\vec{R}$ in Cartesian coordinates, and $\overleftrightarrow{K}^{-1}_{\vec{R}\vec{R}'}$ is the 15×15 interatomic force constant matrix (IFCM). (For a similar formulation, see Marder, *Condensed Matter Physics*, 2nd Ed., pp. 346-350 [69].) For a periodic lattice with translational symmetry, the IFCM between $\vec{R}$ and $\vec{R}'$ is diagonal in Bloch momentum representation ($\vec{q}$) so that the above equation be written as:

$$\vec{u}_{\vec{R}} = -\sum_{\vec{R}',\vec{q}} e^{i\vec{q}\cdot\left(\vec{R}-\vec{R}'\right)} \overleftrightarrow{K}_{\vec{q}}^{-1} \vec{F}_{\vec{R}'}. \tag{3.3}$$

Since the eigenmodes $\hat{\varepsilon}_{\vec{q}\alpha}$ of the IFCM $\overleftrightarrow{K}_{\vec{q}}$ (where α runs from 1 to 15 and labels the band) form a complete orthonormal basis for the 15×1 vectors, we can insert unity or $\sum_\alpha \hat{\varepsilon}_{\vec{q}\alpha}\hat{\varepsilon}^*_{\vec{q}\alpha}$ into the equation to get

$$\vec{u}_{\vec{R}} = -\sum_{\vec{R}',\vec{q},\alpha} e^{i\vec{q}\cdot\left(\vec{R}-\vec{R}'\right)} \overleftrightarrow{K}_{\vec{q}}^{-1} \hat{\varepsilon}_{\vec{q}\alpha} \left(\hat{\varepsilon}^*_{\vec{q}\alpha} \cdot \vec{F}_{\vec{R}'}\right). \tag{3.4}$$

which finally yields:

$$\vec{u}_{\vec{R}} = -\sum_{\vec{R}',\vec{q},\alpha} \frac{e^{i\vec{q}\cdot\left(\vec{R}-\vec{R}'\right)} \hat{\varepsilon}_{\vec{q}\alpha} \left(\hat{\varepsilon}^*_{\vec{q}\alpha} \cdot \vec{F}_{\vec{R}'}\right)}{\kappa_{\alpha\vec{q}}}, \tag{3.5}$$

where $\kappa_{\alpha\vec{q}}$ is the eigenvalue of $\overleftrightarrow{K}_{\vec{q}}$ with eigenvector $\hat{\varepsilon}_{\vec{q}\alpha}$. Thus an arbitrary displacement pattern of atoms in a unit cell can be written as a linear combination of the eigenmodes of the IFCM, and the coefficients are projections of the forces onto those modes in the manner shown above. At the BTO/Ge interface, the dimerized Ge surface imposes a force on the BTO atoms which have a 2×1 periodicity. Hence, the lattice displacements of the interfacial BTO atoms can be expressed as a linear combination of the eigenmodes of $\overleftrightarrow{K}_{\vec{q}}$ with wavevectors $\vec{q} = [0,0,0]$ (Γ) and $\vec{q} = [1/2, 0, 0]$ (X). Notice that the eigenmodes $\hat{\varepsilon}_{\vec{q}\alpha}$ of $\overleftrightarrow{K}_{\vec{q}}$ are related but not equivalent to the more conventional phonon modes $\hat{\eta}_{\vec{q}\alpha}$, which are the solutions to the dynamic equation $\overleftrightarrow{K}_{\vec{q}}\hat{\eta}_{\vec{q}\alpha} = \omega^2_{\vec{q}\alpha}\overleftrightarrow{M}\hat{\eta}_{\vec{q}\alpha}$, where $\overleftrightarrow{M}$ is the diagonal mass matrix and $\omega_{\vec{q}\alpha}$ is the frequency of the phonon. Eigenmodes of $\overleftrightarrow{K}_{\vec{q}}$ are independent of the masses and are appropriate for describing a static distortion pattern.

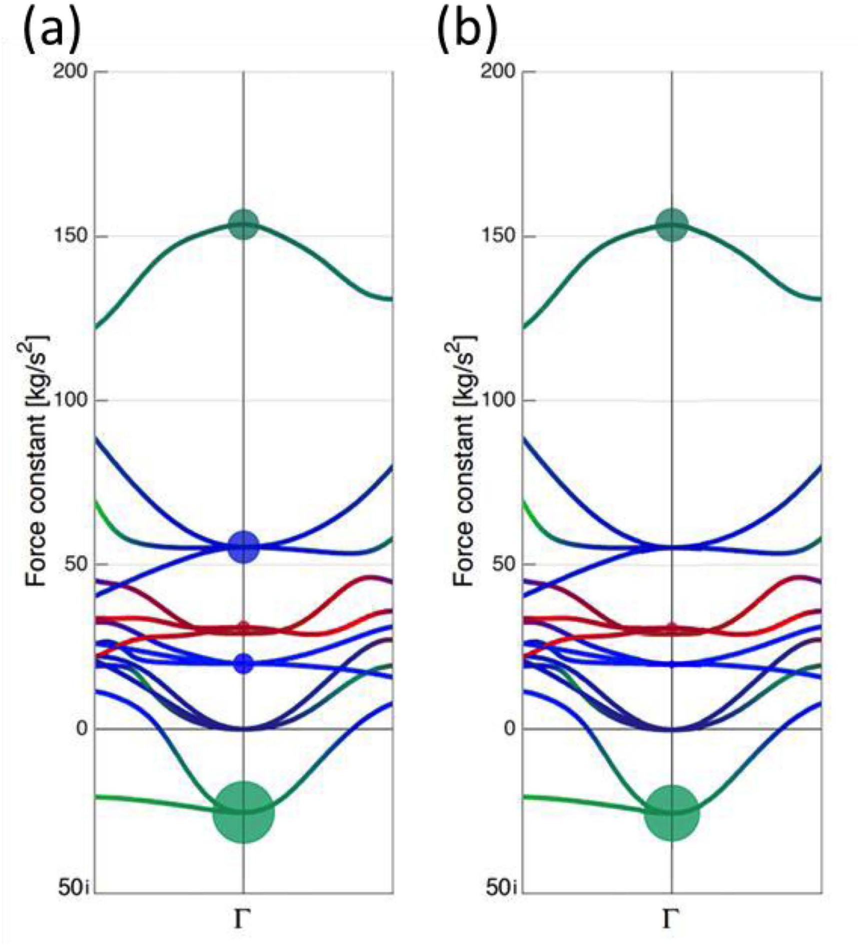


Figure 3.8: (a) Zone-center projections of the atomic distortions in the interfacial unit cell of $BaTiO_3$ grown on Ge onto the cubic eigenmodes are shown as large circles overlaid on the dispersion curves. (b) The projections of the orthorhombic bulk displacements of $BaTiO_3$ overlaid onto the cubic eigenmodes.

The amplitudes of the resulting projections are plotted as circles on the zone center dispersion curves for the interfacial BTO layer and the orthorhombic bulk BTO in Figure 3.8 (a) and Figure 3.8 (b), respectively. For this operation to be well defined, a linear transformation mapping the orthorhombic unit cell onto a cubic unit cell is applied on the atomic positions in the orthorhombic phase, which results in a configuration that is polarized equally in the x and z directions. We find that the orthorhombic displacements project almost exclusively onto the soft unstable mode and the stiff mode with the highest force constant. The displacement patterns of these modes are shown in Figure 3.9. The magnitudes and the displacement patterns of the projections of the orthorhombic displacements onto these modes are very similar to the projections of the interfacial BTO displacements, which indicates that the polar rearrangement of the interfacial BTO u.c. coincides with the bulk orthorhombic configuration. Lastly the polarizations of the bulk orthorhombic structure are calculated using the Born effective charges and found to be 38 μC/cm^2 along both (100) and (001) axes. This is in agreement with the fact that the zone center distortions in the interfacial BTO layers correspond almost entirely to orthorhombic distortions.

Finally, we comment on the choice of bulk reference cell. As explained above, we have calculated the displacements of the atoms in the BTO film with respect to the high symmetry cubic perovskite cell strained to the Ge(001) surface. This choice is due to its simplicity; however, since the film is free to relax in the (001) direction, it is not the only possible choice. In Figure 3.10 we compare the eigenmode projections for the interfacial unit cell obtained by (a) the high-symmetry cubic reference and (b) a high-symmetry tetragonal reference where the in-plane lattice constant is fixed to that of the Ge(001) surface but the out-of-plane lattice constant is set as the difference between the average z-coordinate of the co-planar atoms in the two consecutive BaO layers. This latter option reflects the possible choice of viewing the out-of-plane relaxation of the BTO layers as irrelevant for the eigenmode analysis. However, there

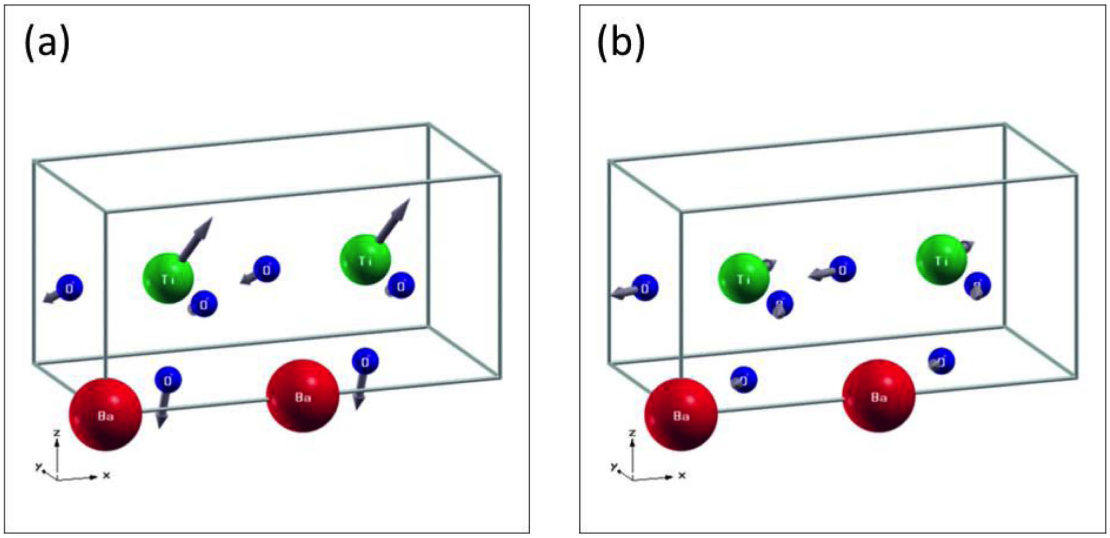


Figure 3.9: Zone-center force constant matrix eigenmodes for bulk cubic $BaTiO_3$ which contribute to the polar atomic displacements for bulk orthorhombic $BaTiO_3$. (a) Soft mode displacements and (b) stiff mode displacements.

is no qualitative difference between the two choices, and the quantitative differences are small. In Figure 3.10, the differences in size of the projections relative to their counterparts that are more than 10% are shown.

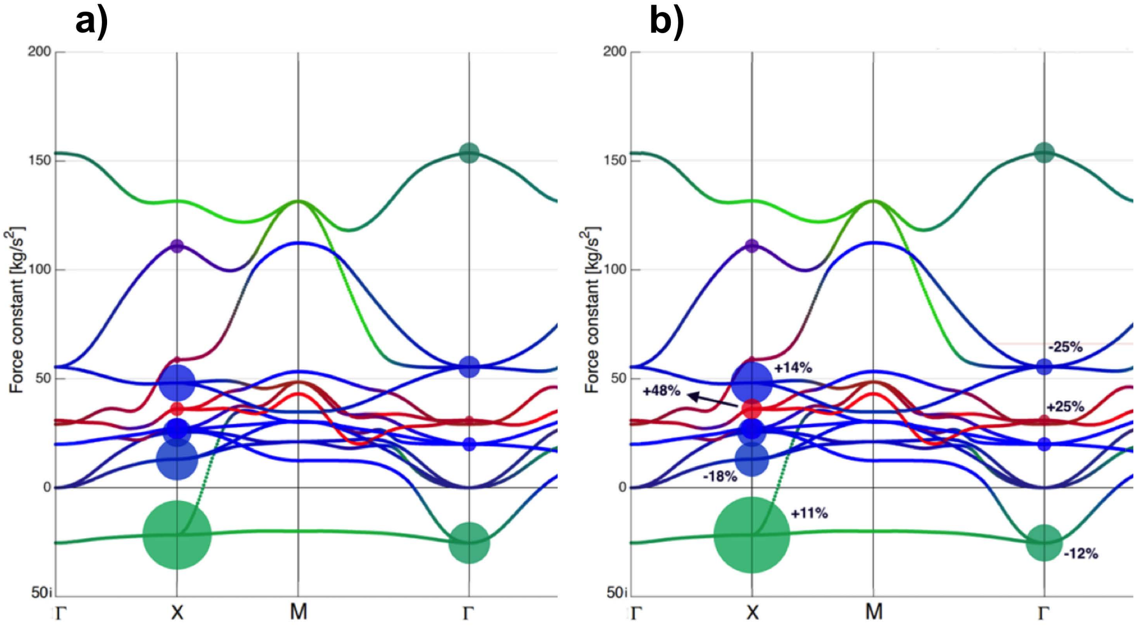


Figure 3.10: Eigenmode projections where the displacements are calculated from (a) the high-symmetry cubic reference and (b) a high-symmetry tetragonal reference, where the lattice constant in the z-direction is set as the vertical distance of the consecutive BaO layers (the position of a layer is taken as the average z-coordinate of the atoms in that layer). Additionally in (b) the differences in size of the projections relative to their counterparts in (a) are shown (only absolute values that are greater than 10% are included).

Chapter 4

Epitaxial $BaTiO_3$/Ge interface

4.1 Introduction

Perovskite oxides have long been of great interest because of a large variety of phenomena that can emerge in these materials, such as ferroelectricity, ferromagnetism and superconductivity [70]. The phenomena give rise to technological utilization of these materials ranging from memory to photovoltaics [71]. Epitaxially growing oxides on semiconductors with an abrupt interface opens up possibilities for device applications with enhanced functionalities [6, 72, 2, 73]. In the absence of interfacial amorphous layers (such as SiO_2), the mechanical and electrical properties of the semiconductor and the crystalline oxide can couple directly. Since the achievement of the first interface of this type between strontium titanate and silicon [46], $SrTiO_3$ has been used as a template for epitaxial growth of other oxides on Si [74, 75, 76, 77] and Ge [78, 79]. With a high dielectric constant, STO was considered a promising candidate for a gate dielectric to replace SiO_2 [80] until it became clear that band alignment between Si and STO does not favor an insulating state for the thin oxide [81, 82]. However, because of potential applications of oxide/semiconductor systems in which a non-zero current through the oxide is desired, such as tunnel junctions, STO/Si heterostructures are still an active area of research.

Another motivating factor for studying STO thin films has been the possibility of inducing ferroelectricity in these films through epitaxial strain [83]. A ferroelectric insulating oxide that has a crystalline interface with a semiconductor forms the basis of ferroelectric field effect transistors (FEFET). In such a device, the polarization of the oxide can be switched by the application of a gate voltage, which in turn modifies the transport properties of the semiconductor underneath. Therefore the state of the device is encoded in the polarization of the ferroelectric oxide and does not require the continuous application of a gate voltage, which makes a FEFET non-volatile. Achieving non-volatile transistors would dramatically decrease the energy consumption of a wide variety of devices, and has been an important research goal for decades [3]. Barium titanate is an excellent candidate for this application because of its multitude of ferroelectric phases in the bulk and its high dielectric constant [68].

$BaTiO_3$ has been successfully grown on Si [84, 85, 73]; however in these studies $SrTiO_3$ has been used as a buffer layer due to the high lattice mismatch between Si and BTO of $\sim 4\%$. On the other hand, germanium closely matches the lattice constant of BTO within 1%, which has prompted many studies of BTO/Ge heterostructures [6, 48, 49, 78, 79, 86]. Some of these studies have included direct epitaxial interfaces between Ge and BTO without buffer layers; but a report on the atomic structure of this interface has only recently been published [11]. In this recent study, we reported on the epitaxial growth of ultrathin BTO films (2.5 & 5.5 unit cells) on Ge(001) surface. We presented a detailed analysis of the interface-induced 2×1 structural distortions in the oxide film which are associated with charge, spin and orbital order effects [40, 59]. In this current work, we complement our initial report with a comprehensive first-principles study of the structure of thin BTO films on Ge(001). In section 4.3.2 we describe the atomic configurations of the various stable phases; in section 4.3.3 we analyze the chemistry of the semiconductor-oxide interface, and the resulting electronic structure and band alignments throughout the system; in section 4.3.4 we describe the ionic polarization in the oxide film; and in section 4.3.5 we discuss oxygen content and how the film is stabilized on Ge with a high quality non oxygenated interface.

4.2 Methods

We use density functional theory (DFT) with the Perdew–Burke–Ernzerhof generalized gradient approximation (PBE GGA) [18] and ultrasoft pseudopotentials [19]. We employ the QUANTUM ESPRESSO software package [65]. A plane wave energy cutoff of 35 Ry and an 8×8 Monkhorst–Pack k-point mesh (per 1×1 in-plane primitive cell) with a Marzari–Vanderbilt smearing of 0.02 Ry [64] are used. A sample 2×1 simulation cell is shown in Figure 4.1. A typical simulation cell consists of 16 atomic layers of Ge whose bottom layer is passivated with H on top of which 1.5 to 5.5 unit cells (or 3 to 11 monolayers) of BTO are placed. In some cases an electrode is added on top of the BTO. Each cell includes $\sim 12\,\text{Å}$ of vacuum to separate periodic copies of the system in the z-direction. The in plane lattice constant is fixed to the computed bulk Ge lattice constant of $4.06\,\text{Å}$. Because of the asymmetric nature of the simulation cell, the system has an overall dipole moment which can interact with the dipole moments of the periodic copies. In order to eliminate this artificial electrostatic interaction, we introduce a fictitious dipole in the vacuum region of the cell whose magnitude is such that it cancels out the electric field in vacuum [66]. All atoms (except the bottom 4 layers of Ge which are fixed to bulk coordinates) are relaxed until the forces on the atoms are less than $10^{-3}\text{Ry}/a_0$ in all axial directions (a_0 is the Bohr radius). In section 4.3.2 and section 4.3.5 we compute the transition energy barrier between different configurations using the nudged elastic bands (NEB) method with climbing images [25].

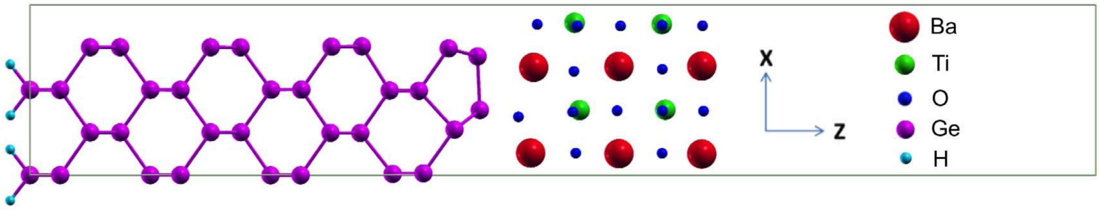


Figure 4.1: A sample simulation supercell for the BTO/Ge system without a top electrode. The displayed configuration corresponds to the 5 ML thick relaxed BTO, and is in agreement with experiment [11].

Bulk phase	Energy (eV per $BaTiO_3$)	Temp. (K) [1]
Rhombohedral	$\equiv 0.000$	≤ 180
Orthorhombic	0.002	$180 - 280$
Tetragonal	0.014	$280 - 400$
Cubic	0.068	≥ 400

Table 4.1: Total energies of the structural phases of $BaTiO_3$ at 0 K computed in this work, and temperatures at which each phase is most stable [1].

4.3 Results

4.3.1 Bulk $BaTiO_3$

$BaTiO_3$ is observed in four structural phases. At high temperature, the cubic perovskite phase is stable. As the temperature is lowered, three structural transitions to ferroelectric phases occur: first to the tetragonal, then to the orthorhombic and lastly to the rhombohedral phase. We show the computed energies of each BTO phase in Table 4.1, as well as the temperature range in which each phase is most stable.

In Table 4.2, we report the lattice parameters of each phase and compare them to previous theoretical [1] and experimental values [87]. All of the ferroelectric phases can be obtained by continuous breaking of symmetry in the cubic phase: Tetragonal, orthorhombic and rhombohedral phases are obtained by breaking the symmetry in the {001}, {101} and {111} directions, respectively, and allowing the atoms and the cells to relax fully. Breaking the symmetry in the {101} direction results in a monoclinic cell with $a = c$, however a twice as large unit cell with higher symmetry (orthorhombic) is used to label this phase.

4.3.2 Structure of the $BaTiO_3$/Ge(001) interface

4.3.2.1 Stoichiometry and periodicity

Researchers have reported studies of heterostructures involving $BaTiO_3$ and Ge [88, 49, 79] including $BaTiO_3$ thin films directly grown on Ge [86]. However, a study of the direct epitaxial BTO/Ge interface has only recently been published [11]. In this section we present a more detailed account of the *ab initio* investigation of the interfaces discussed in that work.

From a purely theoretical viewpoint, we have enormous freedom in terms of what type of epitaxial interfaces to explore. First, the atomic layer neighboring the Ge(001) surface, i.e. the interfacial layer, could be BaO or TiO_2. Second, the interfacial layer could be stoichiometric or not. Third, we do not know *a priori* what the in-plane periodicity will be once interfacial reconstructions occur. For the first point, we consult the experimental growth procedure, which starts by depositing 0.5 ML of Ba on the bare Ge surface at 440 °C and another 0.5 ML of Ba at room temperature. This strongly suggests that the interfacial layer is BaO with full barium stoichiometry. We also initially assume full oxygen stoichiometry to begin our analysis (we discuss variable oxygen content in section 4.3.5). Finally, we analyze the bare Ge(001) surface in order to decide on the lateral size of the simulation cell.

The Ge(001) surface has been widely studied [89, 90, 91] and it is known that the surface atoms strongly prefer to dimerize, which changes the periodicity from 1×1 to 2×1. We find that dimerization lowers the energy by 0.86 eV per dimer. Without

Bulk BTO phase	Parameters	This work (Å)	Theoretical ref. [1]	Experimental ref. [87]
Rhombohedral	a, α	4.05, 89.9°	4.00, 89.9°	4.00, 89.8°
Orthorhombic	a, b, c	3.97, 5.77, 5.82	3.98, 5.67, 5.69	3.99, 5.68, 5.69
Tetragonal	a, c	3.99, 4.18	3.99, 4.04	3.99, 4.04
Cubic	a	4.02	3.94	4.00

Table 4.2: Lattice parameters of the bulk phases of $BaTiO_3$ compared with theoretical and experimental references.

Ge surface	Energy (eV/dimer)	Ref. [89, 90]	Ref. [91]
flat $p(2 \times 1)$	$\equiv 0.00$	$\equiv 0.00$	$\equiv 0.00$
buckled $p(2 \times 1)$	-0.32	-0.24	-0.30
buckled $c(4 \times 2)$	-0.39	-0.31	-0.38
buckled $p(2 \times 2)$	-0.40	-0.31	-0.38

Table 4.3: Energies of the four lowest-energy Ge(001) surface reconstructions computed in this work, compared with other theoretical studies.

changing the periodicity, germanium dimers buckle and lower their energy further by 0.32 eV per dimer. Other reconstructions involve nearby dimers buckling in opposite ways. Some of these further reconstructions lower the energy, which we report in Table 4.3 (see the references in the table for detailed descriptions of the surface configurations). Since increasing the periodicity beyond 2 × 1 does not significantly reduce the energy, we choose to simulate 2×1 cells. We have also checked the validity of this choice *a posteriori* as discussed below.

4.3.2.2 Interface configurations

We begin by investigating the stoichiometric interface with 5 ML thick BTO (2.5 unit cells), which is one of the two experimentally studied thicknesses [11]. We perform our simulations at the computed Ge lattice constant of 4.06 Å which puts a 1% tensile strain on cubic BTO whose computed lattice constant is 4.02 Å (see Table 4.2). The two lowest-energy structures we have found are shown in Figure 4.2. We name them "asymmetric" and "symmetric" based on whether they are reflection symmetric with respect to a yz-plane (e.g., a plane cutting through the Ge atoms in the 3rd and 4th Ge layers directly under a dimer). The symmetric structure is found to be 0.39 eV per dimer higher in energy than the asymmetric structure.

The asymmetric interface is chiefly characterized by a large vertical rumpling of

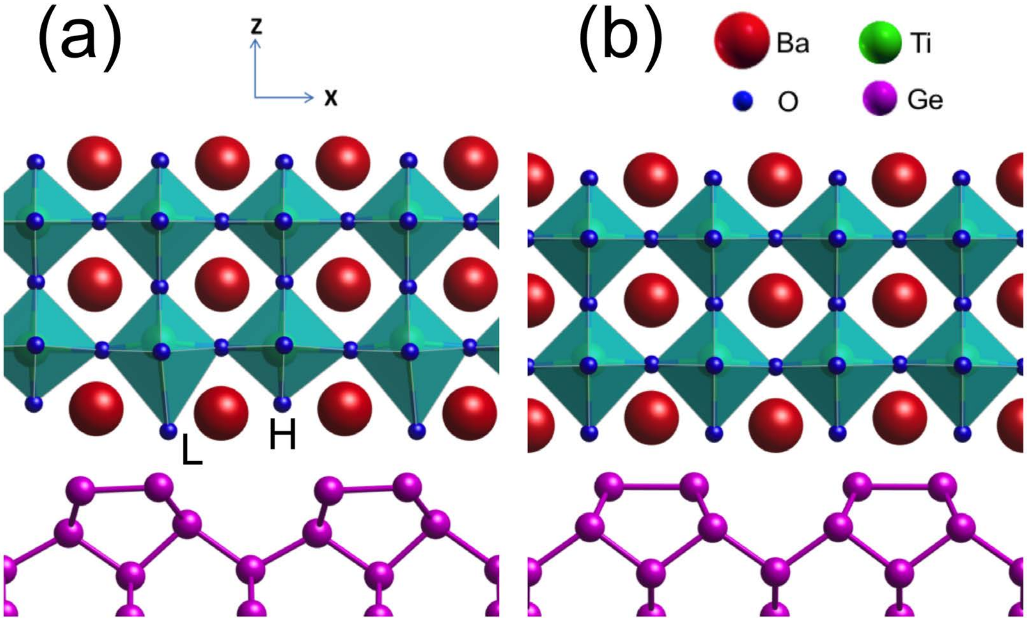


Figure 4.2: Computed configurations of (a) the ground state (asymmetric) and (b) the metastable state (symmetric) of the 5 ML BTO/Ge interface. The 2×1 unit cell, copied two times in the $x-$direction, is shown for both structures. TiO_6 octahedra are also displayed.

consecutive Ba atoms as well as consecutive O atoms in the x-direction. This causes the interfacial TiO_6 octahedra to have different volumes. The octahedra with the lower (higher) interfacial oxygen have a volume of 12.5 $Å^3$ (11.4 $Å^3$). The asymmetric interface also maintains the primary features of the 2×1 Ge(001) surface, i.e. buckled Ge dimers, though the buckling is significantly reduced. See Table 4.4 for a summary of dimer lengths and tiltings. The lower interfacial oxygen (labeled L in the figure) approaches the higher interfacial germanium and makes a bond of length 1.9 Å, whereas the higher interfacial oxygen (labeled H) approaches the titanium above and makes a bond of length 1.8 Å (compare to the 2.0 Å in the high-symmetry cubic bulk BTO). The interfacial oxygens lie in the same xz-planes as Ge dimers and the interfacial bariums lie in the xz-planes that are halfway between consecutive interfacial O-Ge planes (this also holds for the symmetric structure). We have also

Structure	Dimer length (Å)	Buckling angle
Ge(001) $p(2 \times 1)$	2.5	19.4°
BTO/Ge asymmetric	2.6	2.1°
BTO/Ge symmetric	2.7	0.0°
BTO/Ge no-oxygen	2.9	0.0°

Table 4.4: Ge dimer length and buckling angle for the bare Ge(001) surface and for the three interface structures studied.

computed the ground state configuration for the 2×1 slab with no interfacial oxygen, which is analogous to the symmetric oxygenated interface in that it has the same reflection symmetry. Because in this structure the interfacial Ge atoms are neighbors to Ba atoms alone, they accept electrons and therefore dimerize less prominently, as indicated by the elongated Ge-Ge bond length in Table 4.4.

For the symmetric structure, all the interfacial TiO_6 octahedra have the same volume of $11.6\,\text{Å}^3$. They approach the interfacial germaniums and make bonds of length 1.9 Å, whereas their distance from respective titaniums above increase to 2.2 Å. The energy reduction by the formation of the asymmetric (symmetric) interface from the Ge(001) surface and the 5 ML BTO slab in vacuum is 1.18 (0.79) eV per 2×1 cell.

After finding the ground state structure in the 2×1 cell, we have checked whether there are lower energy structures with 2×2 and $c\,(4 \times 2)$ periodicities in the following way. We have generated a "reflected" version of the asymmetric interface by reflecting the structure through a yz-plane that leaves the location of the dimer unchanged, but reverses its buckling. This plane leaves the bulk regions of Ge commensurate in the reflected and the unreflected configurations when they are joined in larger simulation cells. Then we have generated 2×2 and $c\,(4 \times 2)$ cells by appropriately joining the asymmetric and the reflected asymmetric structures. In both cases we have found

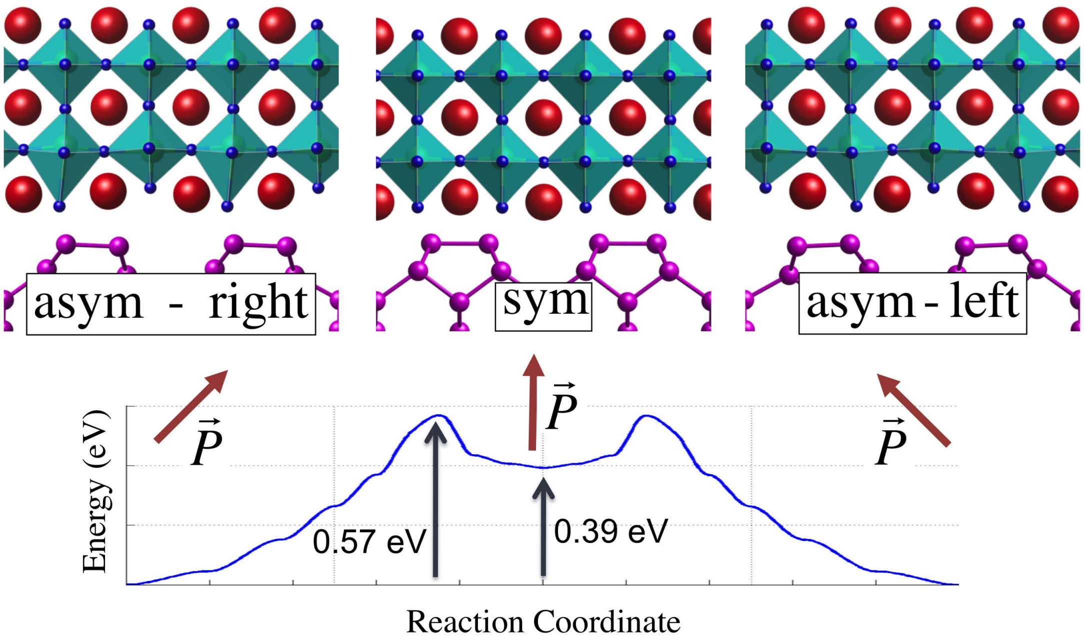


Figure 4.3: Minimum-energy transition path from the asymmetric structure (labelled "asym - left") to the physically equivalent reflected asymmetric structure (labelled "asym - right"), computed via the NEB method with climbing images. The transition passes through the symmetric structure, which is a local energy minimum. The energy barriers into and out of the symmetric local minimum are 0.57 eV and 0.22 eV, respectively. The polarization direction for each structure is also shown in the figure. The two symmetry related asymmetric structures have opposite in-plane polarizations, but the same out-of-plane polarization.

that the energy is raised by 0.39 eV per 2×1 cell.

Finally, we have computed the barrier for the uniform transition from the asymmetric to the symmetric structure as 0.57 eV per 2×1 cell, using the NEB method. We demonstrate the results of this computation in Figure 4.3. Because the transition from the asymmetric structure (labelled "asym - left") to the reflected asymmetric structure (labelled "asym - right") passes through the symmetric structure, the energy landscape has a triple well character: The middle well is higher in energy, the other two wells are symmetry related, and the energy barriers into and out of the middle well are equal to 0.57 eV and 0.22 eV, respectively.

Top BTO layer or electrode	$E(\text{sym}) - E(\text{asym})$ (eV)
Top layer: Ba	0.65
Top layer: $\text{BaO}_{\frac{1}{2}}$	0.54
Top layer: BaO	0.39
Na (2.28 eV)	0.44
Al (4.08 eV)	−0.09
Au (5.10 eV)	−0.34
Pt (6.35 eV)	−0.73

Table 4.5: Energy differences between the 5-ML-thick BTO films with the asymmetric and the symmetric interfaces for different top surfaces and electrodes. Oxygen vacancies in the top BaO layer act as electron donors, so an oxygen deficient surface layer acts as an electrode with a low work function. In the bottom four rows the top layer of the oxide is BaO, and the work functions (ϕ) of the electrodes are listed in parentheses [92].

4.3.2.3 Effects of capping electrode

We have found in addition that the energy difference between the two physically distinct configurations depends on the choice of the top electrode. We have computed this energy difference for four choices of electrodes: Na, Al, Au and Pt. The results are shown in Table 4.5. We find that as the work function of the electrode increases, the symmetric interface becomes increasingly more favorable compared to the asymmetric interface. We have tested if the trend holds for surface layers that have smaller electron affinity than BaO, such as $\text{BaO}_{\frac{1}{2}}$ and Ba. Because of the electron donating oxygen vacancies in these layers, they act as electrodes with low work functions. We show in Table 4.5 that the trend holds in both directions: the energy difference between these two configurations can be tuned by the choice of the surface boundary conditions. The underlying reason for this dependence, which is due to preferential electron transfer, is explained in section 4.3.3.

In particular, the two polarization states can be made degenerate by the correct choice of the top electrode. In such a situation, when the polarization state is switched by the application of a gate voltage (which is then turned off), there would be no energetic drive to switch back to the other state. Using the four different metal electrodes, we find that the energy difference between the two states depends essentially linearly on the experimental work function of the metal (taken from [92]) and we estimate that an electrode with $\phi = 3.82$ eV would cause the two states to be degenerate in energy.

4.3.3 Electronic structure

4.3.3.1 Interfacial chemistry

To understand the electronic structure of our interfaces, we first analyze the chemical bonds between the top Ge layer and the interfacial oxide layer. A simple description can be given as follows. For the fully oxygenated interfaces, the primary chemical bonding occurs between the sp^3-like dangling hybrid orbitals of the dimerized Ge atoms (named h_1 and h_2) and the $2p_z$ orbitals of the interfacial O atoms (named p_1 and p_2). For the asymmetric structure (see Figure 4.4(a), left panel), only h_2 and p_2 overlap to bond. Before the interface has formed, the oxygens are approximately in the O^{2-} state, so p_2 holds two electrons and h_2 holds one electron. After the $(h_2 p_2)$ bond has formed, one electron is released and gets accepted by h_1 (see Figure 4.4(a), middle panel for a level diagram). The $2p_z$ orbital belonging to the other oxygen (namely p_1) does not significantly participate in the process and hence is not shown in the figure. We shall also note that h_1 orbitals weakly mix with the neighboring Ba outer shell orbitals, but not enough to affect the simple description we have given.

For the symmetric interface, on the other hand (see Figure 4.4(b), left and middle panels), both pairs (h_1, p_1) and (h_2, p_2) bond and two electrons (per 2×1 cell) are released. Due to the lack of available interface states, these electrons are accepted by the Fermi level, and hence the interfacial region becomes doped.

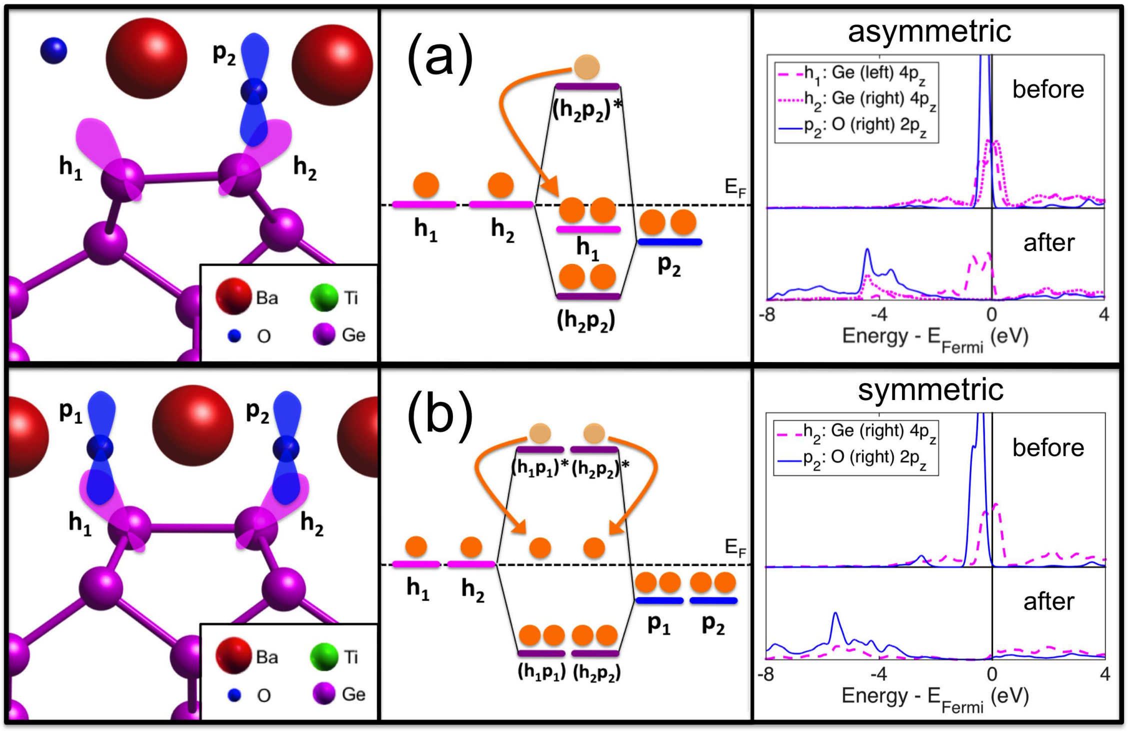


Figure 4.4: Chemical bonding at (a) the asymmetric and (b) the symmetric interfaces. For each interface configuration, atoms with the orbitals that contribute to the process are shown in the left panel (h_1, h_2 and p_2 for asymmetric; h_1, h_2, p_1 and p_2 for symmetric—see text for details); a level diagram that describes the electron occupations of atomic states both before and after the formation of the bonds is shown in the middle panel; and densities of states projected onto the participating orbitals (PDOS) both before and after the formation of the interface are shown in the right panel (Only one pair of orbitals, i.e. h_2 and p_2, is shown for the symmetric interface because of the equivalence of the other pair due to the symmetry of this structure). The zero of energy is taken as the Fermi level and is shown as a solid vertical line.

In order to further demonstrate the bonding, we show the relevant projected densities of states (PDOS) for both interfaces before and after interface formation in the right panels of Figure 4.4. We use Ge $4p_z$ orbitals to represent the dangling orbitals h_1 and h_2 because we expect them to closely align with the $z-$axis due to the dimerization of the interfacial Ge atoms. We have verified this by plotting all PDOS and observing that $4p_z$ orbitals are the only Ge orbitals that are relatively sharp and centered at the Fermi level. For the asymmetric interface we observe that h_2 (initially half-occupied) and p_2 (initially fully occupied) interact to create a bonding state at low energy, whereas the initially half-filled h_2 becomes mostly filled. For the symmetric interface we only display one pair of orbitals, i.e. h_2 and p_2, because the two pairs are equivalent by symmetry. The fully occupied p_2 and the half-occupied h_2 interact to create a bonding state at low energy which is filled, and the conduction band edge is pushed below the Fermi level to accommodate the donated electrons.

This picture of the interfacial chemistry is corroborated by the fact that electrodes with higher electron affinity lower the energy of the symmetric interface with respect to the asymmetric interface (see Table 4.5). We find that more electrons are transferred to the electrode in the symmetric case compared to the asymmetric, since there are more mobile electrons in the former. The difference in the transfer of charge is visible in Figure 4.5, where the spatial electronic density redistribution is plotted for (a) the asymmetric and (b) the symmetric interfaces. For each structure, we have computed the electron density of the full Au/BTO/Ge system, $n_{\mathrm{Au/BTO/Ge}}(x,y,z)$, and the isolated Au and BTO/Ge systems with the same atomic positions as the full system, $n_{\mathrm{Au}}(x,y,z)$ and $n_{\mathrm{BTO/Ge}}(x,y,z)$. We have then computed the difference in the density due to the addition of the electrode: $\Delta n(x,y,z) = n_{\mathrm{Au/BTO/Ge}}(x,y,z) - n_{\mathrm{Au}}(x,y,z) - n_{\mathrm{BTO/Ge}}(x,y,z)$. After averaging the values in the 1× direction, the resulting $\overline{\Delta n}(x,z)$ is shown in Figure 4.5. We have also computed the total charge transferred to the electrode to be 0.24 (0.29) e^{-}

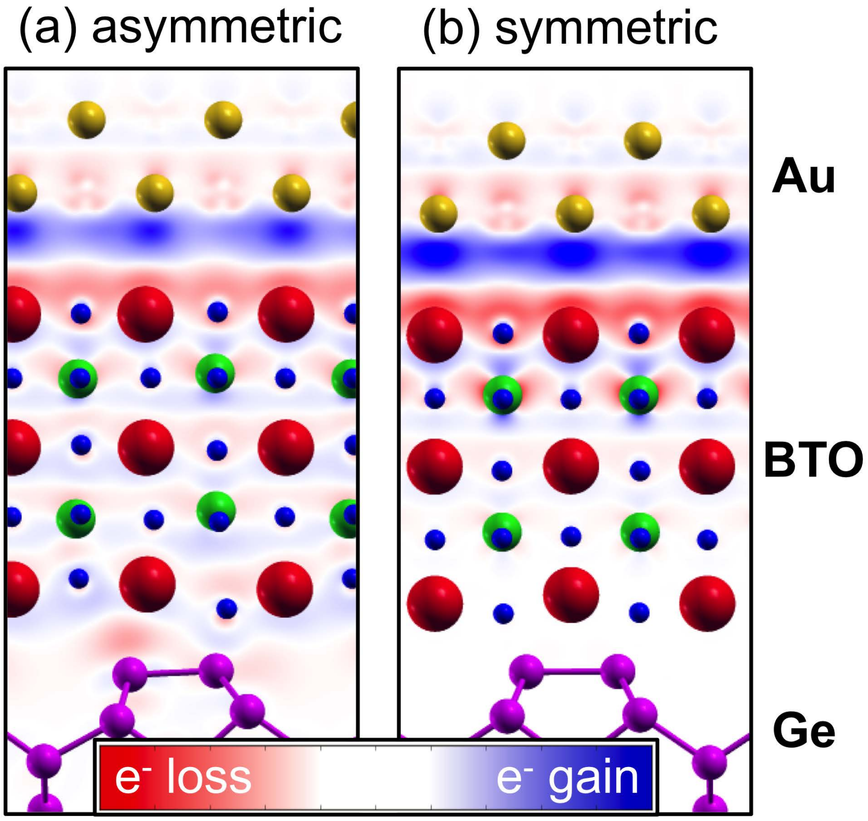


Figure 4.5: Electron density redistribution due to the addition of a Au capping electrode for (a) the asymmetric and (b) the symmetric structures. The plotting is done in the xz-plane by averaging the redistribution in the y-direction.

per 2×1 u.c. for the asymmetric (symmetric) structure. This allows the symmetric interface to have fewer electrons in the high energy conduction band states when an electrode is present. Using an electrode with a higher work function allows more of the mobile electrons to migrate, further lowering the energy of this structure. Conversely, oxygen vacancies in the BTO surface layer donates electrons into the system, further doping the conduction band in the symmetric case and increasing its energy.

For the oxygen-free interface, we have checked that prior to the interface formation the interfacial barium is approximately in the Ba^0 state using the PDOS data for Ba orbitals. The interaction between the interfacial Ge and Ba chiefly consists of ionic electron transfer from Ba to Ge, with a weak covalent mixing between the Ge dangling orbitals and Ba outer shell orbitals. Since there are four Ba valence electrons (per 2×1 cell) but room for only two electrons in the Ge dangling orbitals, two electrons are donated to the Fermi level, as occurs in the case of the symmetric interface. Because the same number of electrons are accommodated by the conduction band which consists of similar states in the two interfaces, the band alignments and the polarization profiles of these two interfaces are highly similar, as we shall present below.

4.3.3.2 Electron redistribution due to interface formation

To further illustrate the electronics of the interface, we have computed the spatial electronic density redistribution for all three interfaces as $\Delta n\left(x,y,z\right) = n_{\mathrm{BTO/Ge}}\left(x,y,z\right) - n_{\mathrm{BTO}}\left(x,y,z\right) - n_{\mathrm{Ge}}\left(x,y,z\right)$. After averaging the values in the 1× direction, we show the resulting $\overline{\Delta n}\left(x,z\right)$ is in Figure 4.6. A few comments can be made by inspecting the plots: (i) The most significant electronic redistribution occurs along Ge-O bonds; (ii) In the asymmetric structure, the Ge atom which does not neighbor an O accepts electrons; but the spatial region accepting region points to the neighboring Ba which is expected due to the positive formal charge state Ba^{2+}; and (iii) For both the sym-

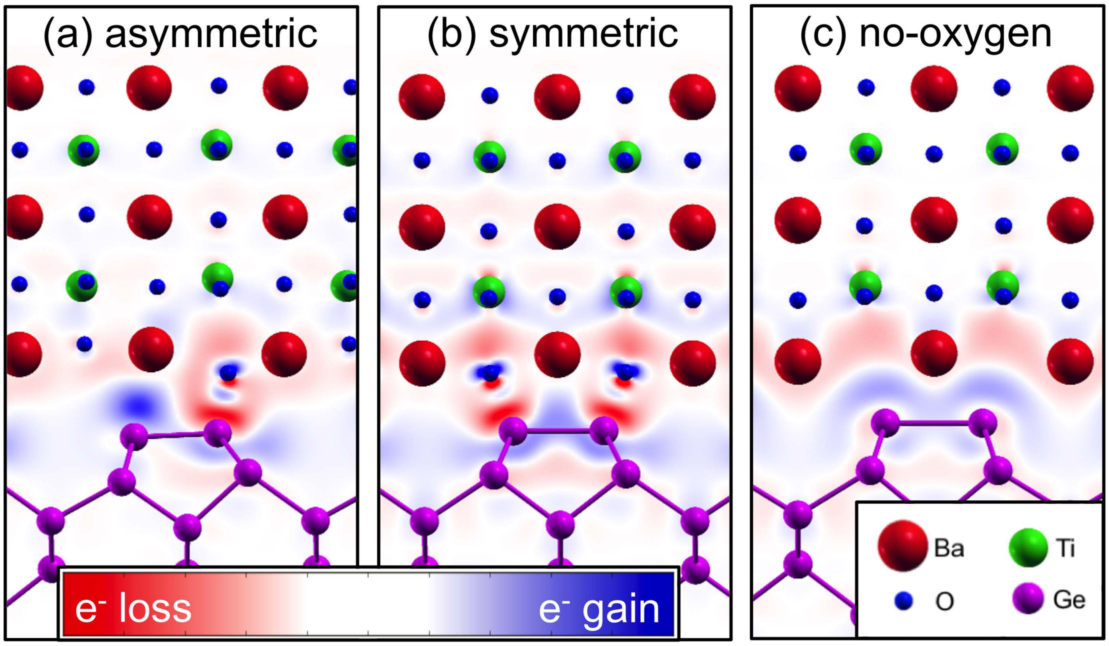


Figure 4.6: Electron density redistribution for (a) the asymmetric, (b) the symmetric and (c) the oxygen-free interfaces. The plotting is done in the xz-plane by averaging the redistribution in the y-direction.

metric and the no-oxygen cases, the interfacial Ba(O) layer loses some electrons to the neighboring Ge and TiO_2 layers.

4.3.3.3 Band alignments

We have found that the two different types of the stoichiometric BTO/Ge interface contains different chemical bonds and give rise to differences in electronic structure. In order to predict the electrical properties of this system for potential device applications, we now turn to a study of the alignments of energy levels. Our above analysis of the interfacial chemistry suggests that the position of the Fermi level is pinned by the interface in both cases. However, in order to infer the spatial behavior of the electronic energy levels throughout the system, we need to express the states in terms of localized functions in real space, such as projections onto atomic orbitals.

We determine the band alignments using a threshold method based on an examination of the layer-by-layer DOS of the semiconductor-oxide system. See Figure 4.7 for an example of the application of the threshold method. For each atomic layer, we sum the DOS projected onto all the atomic orbitals that belong to that layer and plot with a Gaussian broadening. We then find the energy values for each layer where the DOS of the layer is at a threshold to determine the layer by layer band edges throughout the system. We finally decide which layers are sufficiently bulk-like and report the band edges based on those layers. The threshold method as explained is sufficient for a qualitative description of the band alignment. However, the choice of the threshold requires careful analyses of the bulk band structure and DOS so that it reproduces the band edges in the bulk. In addition, due to the well known problem of underestimation of band gaps in density functional theory, we have modified the band edges in order to correct the gaps to match their experimental values in the two materials. For complete details of the determination of band alignments, see Section §4.4.

The main difference in the band alignments of the asymmetric and the symmetric interfaces is the nature of the interfacial states. For the asymmetric system, the interfacial region is dominated by the mostly filled dangling orbital of one of the germaniums (labelled h_1 above) around the Fermi level. These electronic states decay slowly into the Ge due to its small band gap. The weak covalent mixing between h_1 and the nearby Ba outer shell orbitals is manifest as a nonzero DOS for the interfacial BaO layer within the band gap of $BaTiO_3$. For these reasons, in the vicinity of the interface the projected densities of states are not bulk-like and the band edges are not well defined. We instead schematically display the approximate extent of the interface states as rectangles (see Figure 4.8 (a) and (c)). For the symmetric system,

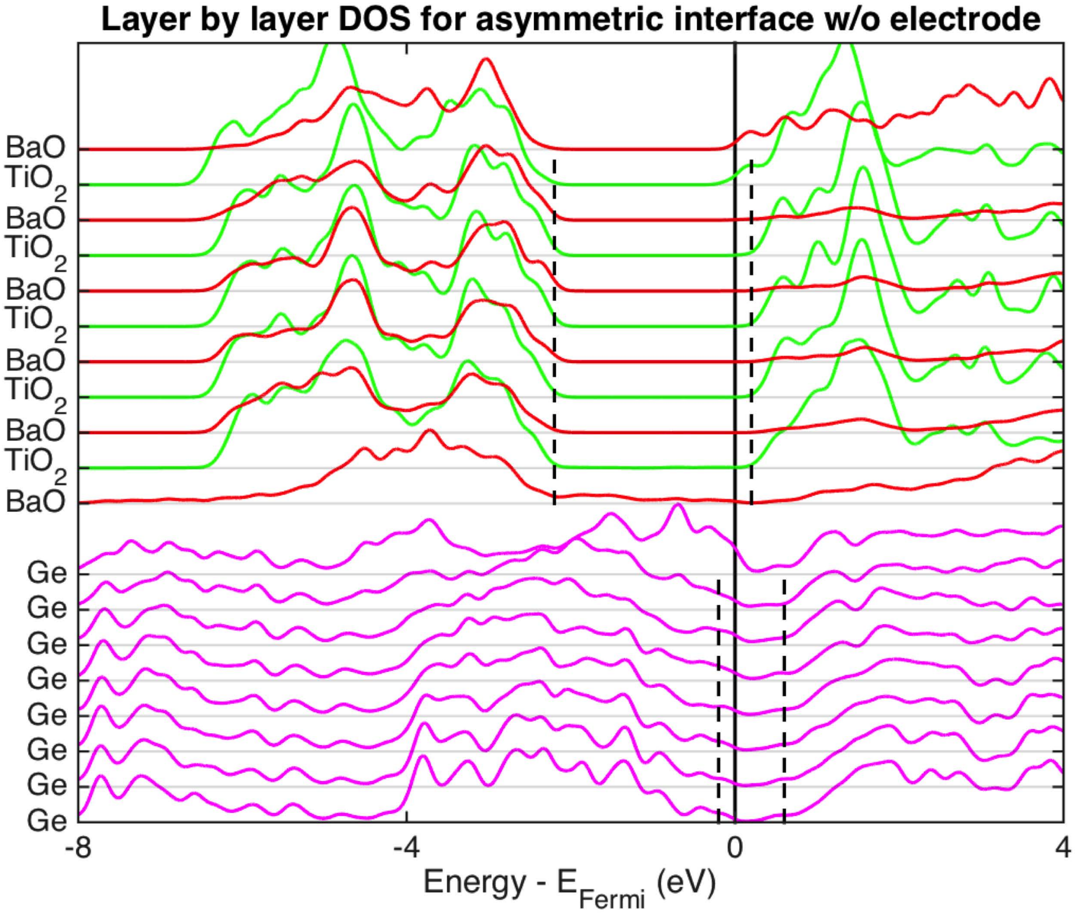


Figure 4.7: Layer-by-layer DOS for the asymmetric interface without a top electrode. The zero of energy is taken as the Fermi energy of the system. The valence and the conduction band edges in both sides of the interface are shown as vertical dashed lines. The band edges are determined by our threshold method (see text for a brief description and Section §4.4 for complete details). Only 8 layers of Ge are shown in the figure.

the PDOS curves in the vicinity of the interface show sufficient bulk-like character to allow us to trace the band edges all the way to the interface. The ejection of electrons from the interfacial region causes the structure to be electron doped in the vicinity of the interface (see Figure 4.8 (b) and (d)).

For both structures, the effect of a top electrode (in our case Au) is similar: It accepts electrons from the system and hence moves the bands upward with respect to the Fermi level. For the asymmetric structure, the effect is more striking because the valence band of the oxide is moved all the way up to the Fermi level. Because of the charge transfer to the electrode, an electron potential is created which increases from the interface to the electrode. This potential in turn enhances the out-of-plane ionic polarization of the film, which creates a potential that decreases from the interface to the electrode. For the asymmetric structure these two competing effects on the potential are in balance and thus the band edges in the oxide region are flat. For the symmetric structure, however, the electron transfer to the electrode is more significant than the enhancement of the ionic polarization; therefore the band edges in BTO have an upward slope going from Ge to Au. The experimental studies on the BTO/Ge heterostructures have reported the conduction band edges to be approximately aligned [49, 86], which suggests that the interfacial chemistry in these experimental systems is similar to the symmetric interface.

We observe that band edges in Ge are positioned differently in the two polarization states, suggesting carrier density modulation in the semiconductor. A similar field effect has been observed in BTO/STO/Ge heterostructures [79]. We further observe that the asymmetric and the symmetric structures have alignments that promote different charge carriers through the system: asymmetric favors holes, and symmetric favors electrons. Since these two configurations can be made approximately degenerate by the choice of the top electrode, it may be possible to switch between these phases with an electric field, as discussed in section 4.3.2. Because of the different

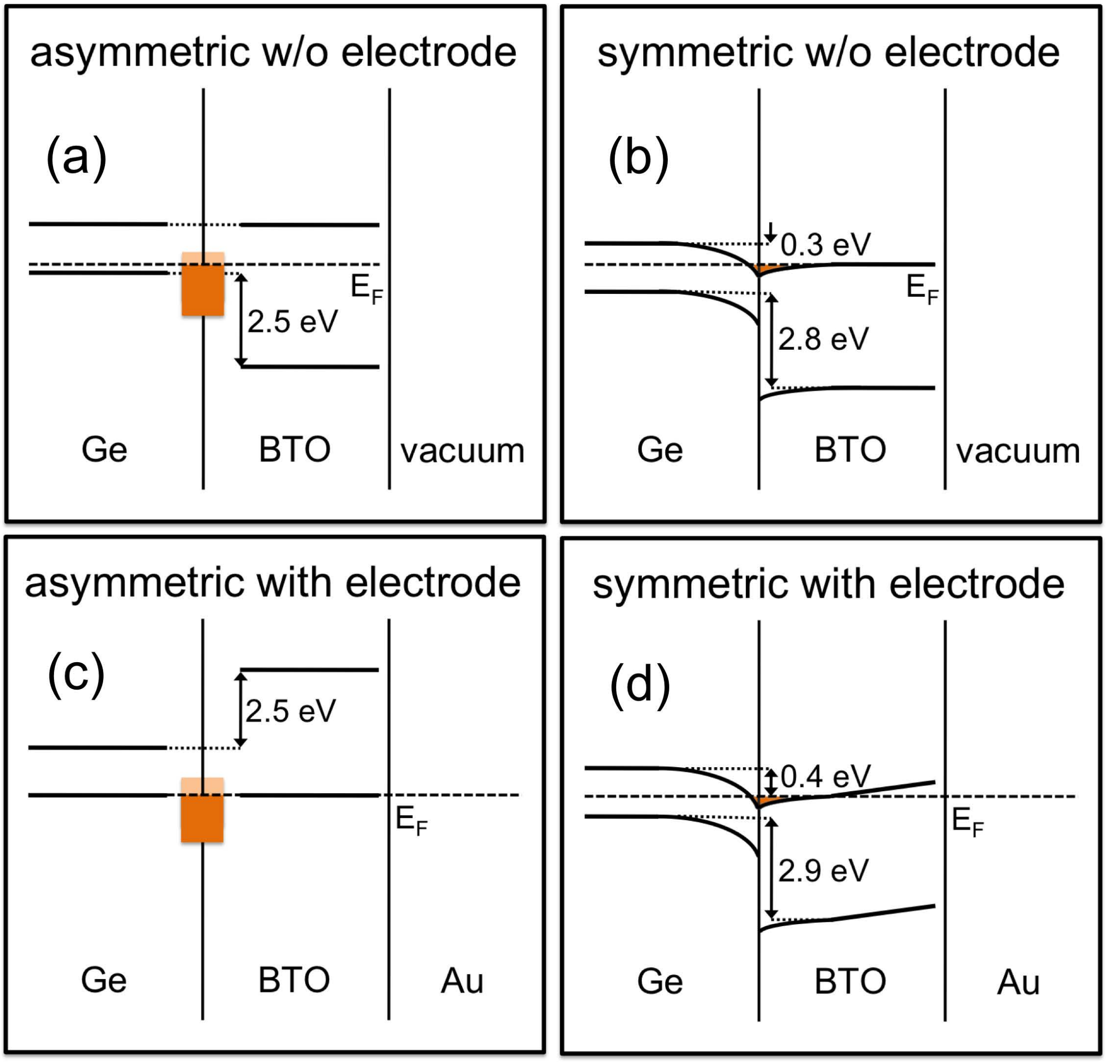


Figure 4.8: Band alignments for the asymmetric and the symmetric interfaces, with and without a gold electrode. The orange shaded rectangles in the alignments of the asymmetric structure represent interface states that extend into both materials and are partially filled. The orange shaded regions in the alignments of the symmetric structure represent the electron doping around the interface. The alignment for the no-oxygen interface is the same as for the symmetric interface.

charge carriers in the two phases, switching between them would change the polarization state as well as the dominant carriers for charge transport. To our knowledge this phenomenon of ferroelectricity combined with carrier type switching has not been observed before and may be useful in potential device applications.

4.3.4 Film polarization

4.3.4.1 Polarization profile

The chief difference between the asymmetric and the symmetric interfaces in terms of ionic polarization is that the asymmetric interface is polarized along $\pm x$- and z-directions, whereas the symmetric interface is polarized only along the z-direction. In order to characterize the out-of-plane polarization, we introduce the parameter δz as the average cation-anion displacement in a given layer, i.e. $\delta z = \overline{z\,(\text{cation})} - \overline{z\,(\text{oxygen})}$. We have calculated δz for each atomic layer of the BTO film for the asymmetric, symmetric and no-oxygen interfaces for various thicknesses. We have also introduced a capping electrode (2 monolayers of gold) in order to investigate the effect of boundary conditions.

We present our findings in Figure 4.9 for the 5 ML and 11 ML thick films, which are the two cases studied in the experiment [11]. When the film is not capped with an electrode (see Figure 4.9 (a)), we find that the top BTO layer pins the surface polarization to the same value for all interfaces, which causes a sudden drop in the polarization profile from the interface to the surface for the thinner films. For the thicker films, polarization reduces for the first two unit cells and remains approximately constant up to the surface. The symmetric configuration (labelled “sym”) has a significantly higher polarization than the asymmetric configuration (labelled “asym”) throughout the film. The structure with no interfacial oxygen (labelled “no-O”) has a similar profile to the oxygenated symmetric structure. The positive interfacial polarization, which is common in all interfaces, is best understood as follows: In all

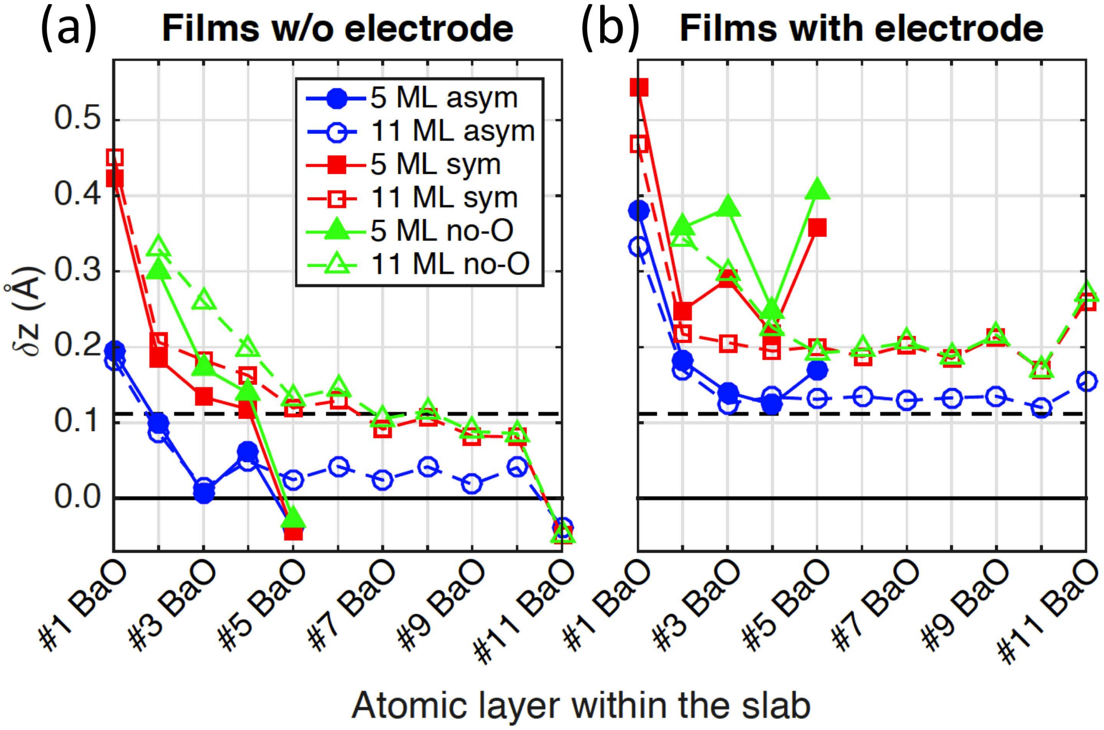


Figure 4.9: Layer-by-layer polarization profile as measured by cation-anion vertical displacement (δz) for 5 ML and 11 ML thick BTO films, plotted for the asymmetric, symmetric and oxygen-free interfaces. The profiles are displayed for the cases where the films (a) are not capped with an electrode and (b) are capped with 2 ML of Au. The #1 BaO layer is the interfacial oxide layer. The dashed horizontal lines show the computed value of δz for orthorhombic bulk BTO, strained in the xy-plane to Ge lattice parameters.

three interfaces there is a transfer of electrons from the interfacial region to the oxide. The resulting electronic dipole pushes the positive (negative) ions up (down), pinning the interfacial δz. In the absence of a well screening electrode, the depolarizing field causes the polarization to decay quickly. In “sym” and “no-O” because of the doping of the conduction band, there are mobile charges in the oxide which screen the depolarizing field so that the polarization in the interior is non-zero. These findings are quite similar to the findings of the previous work on the $SrTiO_3$/Si system, where presence of mobile charges determine the value of the layer polarization away from the interface [63].

When the film is capped with a gold electrode (see Figure 4.9 (b)), the top BaO layer bucklings are no longer pinned and have different values for different thicknesses and interfacial configurations. We observe also that with a gold electrode polarization increases throughout the film for all structures. This is because Au has a large work function and thus attracts electrons from the Ge/BTO system, which in turn enhances the ionic polarization toward the electrode, as we have discussed above. We also observe that “sym” and “no-O” interfaces have very similar profiles that are more polarized than “asym”. This is because of the combined effect of more mobile charges in the oxide and more electron transfer to the electrode in these structures compared to “asym”. We finally observe in Figure 4.9 (b) that δz is generally higher in thinner films than thicker films in a given oxide layer. This is because in thinner films there are fewer oxide states to accommodate the electrons donated from the interface, causing more of the electrons to go to the electrode, and thereby creating a larger electric field.

4.3.4.2 Effects of film thickness

In addition to the 5 ML and the 11 ML thick films studied experimentally [11], we examine the ionic polarizations of films ranging in thickness from 3 to 11 monolayers.

We summarize the results in Figure 4.10. We first observe that for all films with and without a capping electrode, the average polarization has a small modulation due to the surface termination of the oxide. For films without a capping electrode, the average δz has almost no thickness dependence. This is explained as follows: the higher values of δz are confined to the first few layers of the film, lowering the average polarization as the thickness increases. However, because the decay of the polarization becomes more gentle as the thickness increases, especially for "sym" and "no-O", this effect is mostly cancelled. For the films with a gold electrode, both because the profile is flat after the vicinity of the interface and because the near-interface δz is higher for thinner films, the decrease of the average δz with thickness is more significant. This finding is in agreement with the previously studied $SrTiO_3/Si$ system, where for well screened films the average polarization decreases with film thickness [63]. Finally we observe that the in-plane polarization that is present in the asymmetric film, i.e. δx is mostly unaffected by thickness and the presence of a capping electrode. The in-plane polarization, δx, is defined as the horizontal displacement of ions in an atomic $yz-$plane, averaged through all four inequivalent such planes in the 2×1 cells.

4.3.5 Oxygen content

4.3.5.1 Energetics of oxygen vacancies

In order to understand the oxygen content of the films, we have computed the formation energies of oxygen vacancies. First, we have relaxed both the asymmetric and the symmetric structures with one oxygen atom omitted per 2×1 cell. We have repeated this for each unique oxygen site. We have found that in both the asymmetric and symmetric structures, the smallest formation energy for an oxygen vacancy takes place when the vacancy is in the interfacial BaO layer. In Table 4.6 we summarize the formation energies of O vacancies vs layer in both structures. Formation energies are computed as the energy cost of removing an oxygen from its position in the film

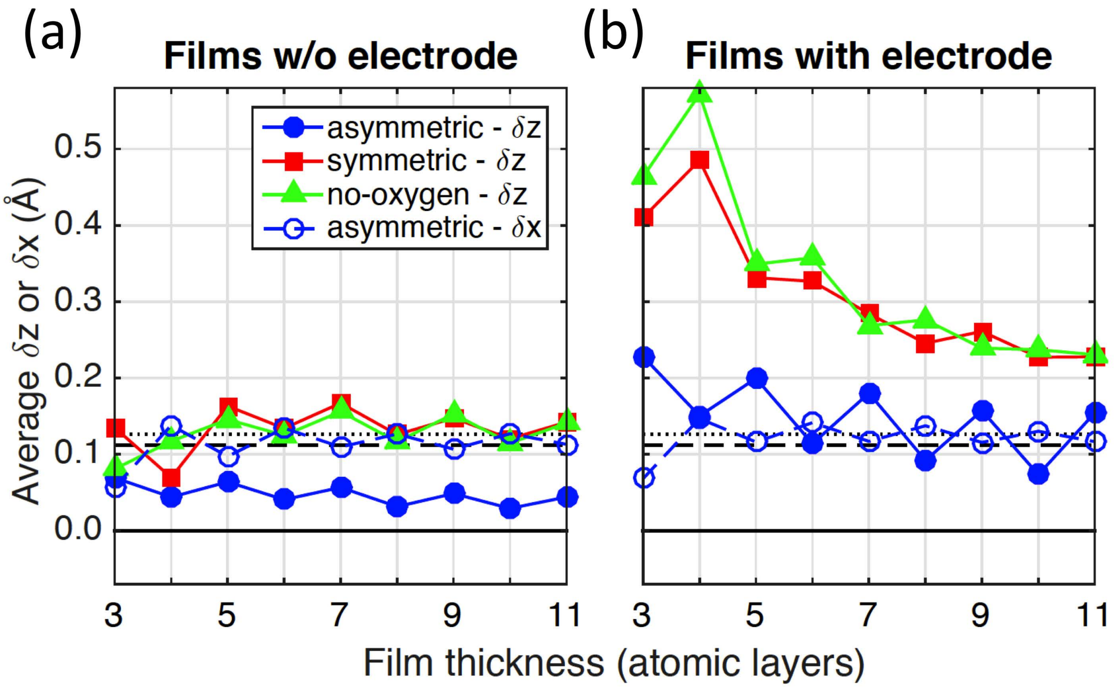


Figure 4.10: Average film polarization as measured by cation-anion vertical displacement (δz) for BTO films with different thicknesses, plotted for the asymmetric, symmetric and oxygen-free interfaces. The in-plane polarization of the asymmetric film is also displayed as measured by the parameter δx (described in text). The results are displayed for the cases where the films (a) are not capped with an electrode and (b) are capped with 2 ML of Au. The dashed (dotted) horizontal lines show the computed value of δz (δx) for orthorhombic bulk BTO, strained in the xy-plane to Ge lattice parameters.

Structure	#1 BaO	#2 TiO_2	#3 BaO	#4 TiO_2	#5 BaO
asym (a)	3.7	5.2	4.7	5.3	5.1
asym (b)	3.7	5.2	4.7	5.2	4.5
asym (c)	3.7	4.1	4.8	5.4	5.0
sym (a)	3.6	5.6	5.0	5.5	5.5

Table 4.6: Formation energies of O vacancies (eV per vacancy) for the asymmetric and symmetric structures, vs atomic layer. For each atomic layer, the lowest formation energy among possible vacancies is shown. #1 (#5) BaO layer is at the interface (surface). (a) 1 vacancy per 2 × 1 cell; (b) 2 vacancies per 2 × 2 cell; (c) 1 vacancy per 2 × 2 cell.

and placing it into an O_2 molecule in vacuum.

For the asymmetric structure, in addition to relaxing the structures with 1 vacancy per 2 × 1 cell, we have relaxed structures with 2 vacancies per 2 × 2 cell, where the vacancies are close enough to be considered neighbors. We list in Table 4.6 vacancy formation energies where the 2 vacancies in the 2 × 2 cell are in the same atomic layer (only the average energy of the 2 vacancies is shown for each layer). Formation energies of vacancies which reside in different atomic layers in these 2 × 2 cells are not significantly different, and are omitted to simplify the discussion. Finally, to compare these results with a different vacancy density, we list the formation energies of structures with 1 vacancy per 2 × 2 cell for the asymmetric structure. We find that the interface is the most energetically favored place for a vacancy for this case as well, indicating that this conclusion holds for the isolated vacancy limit.

Our finding that the interface is the least costly location for an oxygen vacancy irrespective of vacancy density and configuration is in agreement with the previous work on the $SrTiO_3$/Si system [93]. In that work it is argued that this is due to the chemical differences between the interfacial region and the interior of the oxide film. An oxygen vacancy can be thought to donate two electrons to the available states nearby. For a vacancy at the interface, there are low-energy Si dangling bond states that accept the donated electrons. However, for a vacancy in an interior layer of the oxide, it is found that high energy conduction band states with Ti $3d$ character get filled by the donated electrons [93]. Because Ge is isovalent with Si, we expect these findings for the $SrTiO_3$/Si system to apply to the $BaTiO_3$/Ge system as well.

4.3.5.2 Thermodynamics of oxygen content

Having established that oxygen vacancies favor the interfacial BaO layer, we investigate the thermodynamics of interfacial oxygen vacancies. We have computed the lowest energies for the interfaces with varying interfacial vacancy densities, using 2×1, 2×2 and 4×1 cells, and calculated the vacancy formation energies vs density. In order to build a phase diagram, we simply assert that a vacancy density is thermodynamically favored at temperature T and oxygen partial pressure p_{O_2} if $\mu_O(T, p_{O_2}) + E_{\text{formation}} < 0$. We then use the expression [93]

$$\mu_O(T, p_{O_2}) = \frac{1}{2} g_{O_2}(T) + \frac{1}{2} k_B T \ln\left(\frac{p_{O_2}}{1\text{atm}}\right), \tag{4.1}$$

where $g_{O_2}(T)$ is the Gibbs free energy of O_2 in gas phase at temperature T and $p_{O_2} = 1\text{atm}$. We use the experimental measurements for $g_{O_2}(T)$ [94] in order to create the phase diagram shown in Figure 4.11. We also show the growth conditions as a rectangular region where $300\text{K} < T < 800\text{K}$, $10^{-13}\text{atm} < p_{O_2} < 10^{-12}\text{atm}$. The details of the growth are presented in Ref. [11]. The phase diagram indicates that during the growth, vacancies are not thermodynamically stable. We have also computed the energies of the structures where the interface has 0.5 ML of excess oxygen. The region where excess interfacial oxygen is thermodynamically more stable is shown on the phase diagram, and it partially coincides with the growth conditions.

Lastly, for a comprehensive understanding of the oxygen content in the film, we

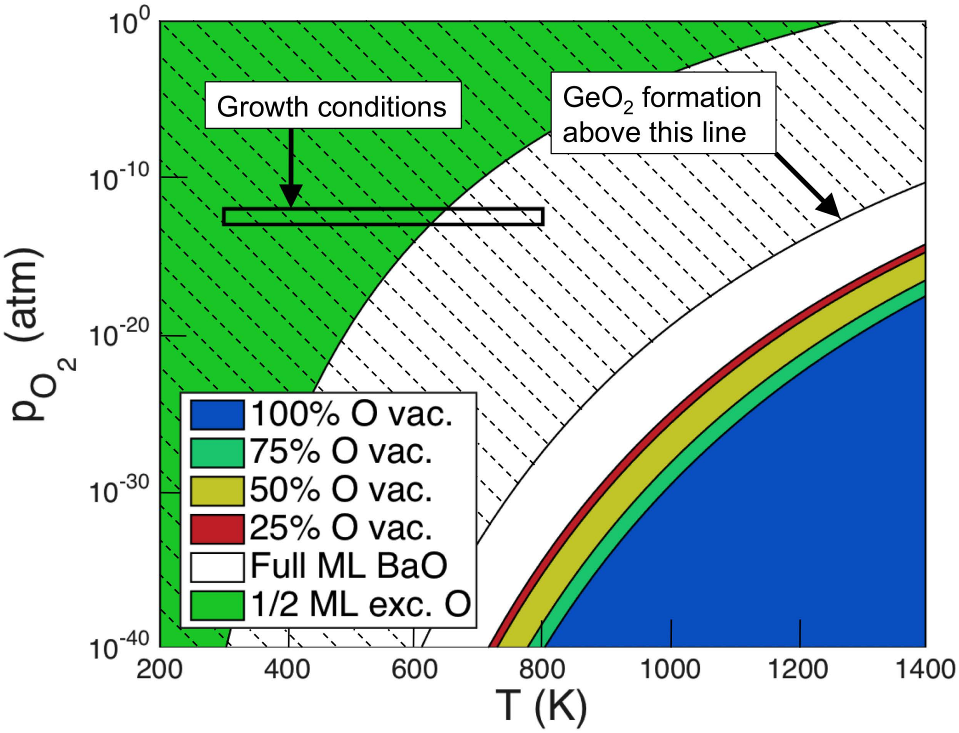


Figure 4.11: Computational phase diagram for oxygen stoichiometry in a BTO thin film grown on Ge(001). Larger and larger vacancy concentrations are thermodynamically favored as one moves to the lower right corner of the diagram. An interface with excess oxygen is preferred at the upper left corner. The vacancies or excess oxygens lie in the interfacial BaO_x layer, which is found to be the most favorable location for a given stoichiometry. The experimental growth conditions described in [11] are shown as a rectangular region, and the region unstable to GeO_2 formation is shaded.

have calculated the oxygen chemical potential at which bulk crystalline GeO_2 becomes as stable as bulk crystalline Ge and O_2 gas, as they exist separately. GeO_2 becomes more stable if $\mu_O\left(T, p_{O_2}\right) > \frac{1}{2}\left(E_{\mathrm{GeO}_2} - E_{\mathrm{Ge}} - E_{\mathrm{O}_2}\right)$, which holds for the shaded area in Figure 4.11, which covers the entire region corresponding to the growth. This well-known vulnerability of the Ge surface to oxygenation is experimentally overcome by kinetic trapping.

4.3.5.3 Kinetic trapping of oxygen

In order to gain theoretical insight about GeO_2 prevention as well as oxygen stoichiometry, we have performed NEB simulations for oxygen vacancy migration within the film, as shown in Figure 4.12. For the 5 ML asymmetric BTO, there are 14 distinct oxygen sites, each of which yields a unique relaxed structure when vacated. Figure 4.12 presents their vacancy formation energies as well as a set of representative minimum energy transition paths for inter-layer vacancy migration. We find that, in addition to the fact that a vacancy costs over an electron-volt more energy when it is not at the interfacial BaO layer, there are energy barriers of ~ 1 eV for vacancy migration between non-interfacial BTO layers. This indicates that once the initial few layers of BTO are deposited, the vacancies formed at the interface are kinetically trapped at the interface.

Let us consider the following process for increasing the oxygen content of the interface beyond stoichiometry: An O atom at the interfacial BaO layer breaks from the film and binds directly to the Ge surface. We have found the minimum of the energies of such configurations to be 1.9 eV higher than the stoichiometric asymmetric film, which causes this event to be unlikely. Once this occurs, a newly formed interfacial vacancy should migrate up into the film and reach the top BaO layer, which is both energetically and kinetically inhibited. Lastly the surface oxygen vacancy should be filled by atmospheric O_2, completing the process which in effect adds an extra oxygen

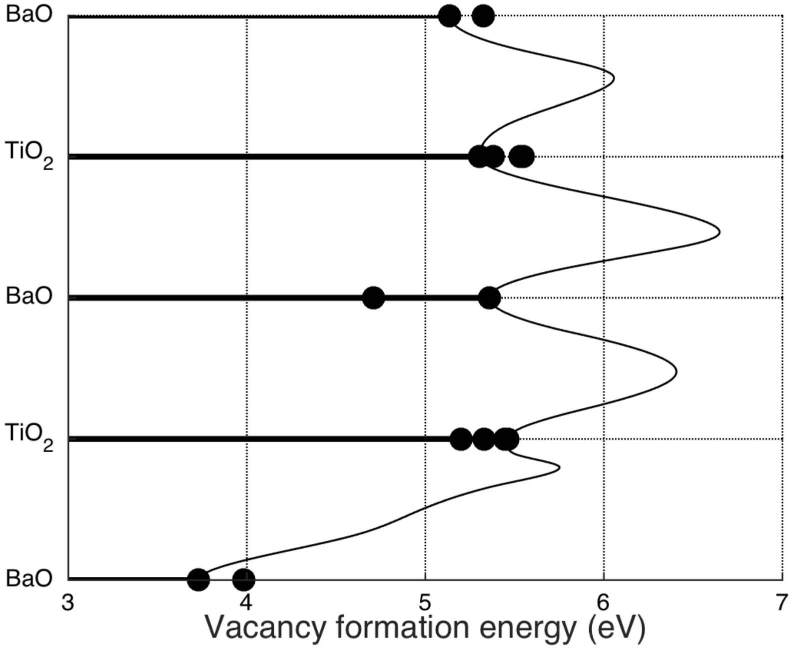


Figure 4.12: Vacancy formation energies for each unique site in the 5 atomic layer BTO film are shown layer by layer as solid circles. The bottom of the plot corresponds to the interface. A set of transition minimum energy paths that are computed via the NEB method are also shown as a solid curve going through a vacancy site in each layer.

to the interface. Therefore in spite of the fact that interfaces with excess oxygen, and eventually full oxygenation of the germanium surface, are thermodynamically favored for this film, once the initial few layers of BTO are deposited, the oxygen content remains stable. This mechanism for trapping the oxygen in the oxide film rather than oxygenating the substrate is similar to the proposed mechanism for $SrTiO_3$/Si [93].

4.4 Conclusion

We have conducted an *ab initio* study of the epitaxial $BaTiO_3$/Ge interface which has potential technological applications. We have found that ultrathin films of BTO grown on Ge(001) surface with full oxygen stoichiometry can occur in two 2×1 configurations with different ionic polarizations. This is in contrast to the similar $SrTiO_3$/Si system, where a 1×1 configuration is observed and the interface polarization is predicted to be fixed [83]. With a carefully chosen electrode, the two polarization states of BTO films on Ge can be made degenerate, enabling a potentially robust ferroelectric thin film oxide. We have analyzed the interface chemistries and the resulting electronic structures in the two polarization states. We have shown that the band alignments for these phases prove that there is significant carrier density modulation in the semiconductor, to the degree of changing the dominant carriers in the intrinsic semiconductor limit. Ferroelectric switching coupled with carrier type switching may have novel uses in device applications. We have examined the ionic polarization profile of these films for a range of thicknesses. We have found that the differences in interfacial chemistry lead to differences in polarization profile, as in the previously studied STO/Si case [63]. We have finally discussed the oxygen content of the ultrathin oxide by a detailed analysis of energetics of oxygen vacancies in the film. We have found that the thermodynamic tendency for the oxygenation of the Ge surface is kinetically inhibited, and once a few atomic layers of BTO are deposited, the oxygen stoichiometry becomes stable. This mechanism of kinetic trapping is similar to the mechanism that stabilizes the STO/Si interface [93].

Appendix: Band alignment determination

We determine the band edges in the two materials based on an analysis of layer-by-layer projected densities of states. We first plot all layer-by-layer DOS for a given configuration as in Figure 4.7. Then, for the layers whose DOS are bulk-like, we find where the DOS falls below a certain threshold around the Fermi level. We choose the threshold for each side of the interface based on the DOS of bulk Ge and bulk BTO, where we know the exact band structures and the band gaps. However, due to the well known DFT underestimation of band gaps, we obtain a bulk BTO with a gap of 1.7 eV compared to 3.2 eV measured experimentally [95]. Moreover, the bulk Ge we compute has no gap while the experimental gap is 0.7 eV [96].

A standard way to overcome this shortcoming of DFT is by applying a Hubbard U correction to atomic orbitals that comprise either the valence or the conduction band edge [97]. This method, commonly known as DFT+U, effectively decreases (increases) the energy of a DFT eigenstate, which overlaps with the chosen atomic orbital, if it is more (less) than half occupied. If the eigenstates that overlap with the chosen orbital comprise the VBE and/or CBE, this procedure increases the band gap. In addition to correcting the gaps, DFT+U simulations of the BTO/Ge system allow us to ascertain how the Fermi level is pinned and how the band edges in BTO and Ge relatively align as the gap in either material is varied. We have run simulations with Hubbard U applied to Ti $3d$ states and/or Ge $4p$ states. We present a representative set of results from these simulations in Table 4.7.

By close examination of the layer-by-layer DOS of each such simulation and application of the threshold method, we identify the band edges and how they depend

Hubbard U (eV)	(a) BE in asym w/o Au	(b) asym with Au	(c) sym w/o Au	(d) sym with Au
$U_{\mathrm{Ge}} = 0$, $U_{\mathrm{BTO}} = 0$	Ge: $-0.2,\ 0.6$ BTO: $-2.2,\ 0.2$	$-0.1,\ 0.7$ $0.0,\ 2.4$	$-0.5,\ 0.3$ $-2.4,\ 0.0$	$-0.4,\ 0.4$ $-2.4,\ 0.0$
$U_{\mathrm{Ge}} = 3$, $U_{\mathrm{BTO}} = 0$	Ge: $-0.3,\ 0.7$ BTO: $-2.1,\ 0.3$	$-0.2,\ 0.8$ $0.0,\ 2.4$	$-0.7,\ 0.3$ $-2.4,\ 0.0$	$-0.6,\ 0.4$ $-2.4,\ 0.0$
$U_{\mathrm{Ge}} = 3$, $U_{\mathrm{BTO}} = 3$	Ge: $-0.3,\ 0.7$ BTO: $-2.2,\ 0.4$	$-0.2,\ 0.8$ $0.0,\ 2.6$	$-0.7,\ 0.3$ $-2.6,\ 0.0$	$-0.6,\ 0.4$ $-2.6,\ 0.0$
Summary:	Midpoint of gap approx. fixed in Ge and BTO.	Midpoint of gap fixed in Ge, VBE at E_{Fermi} in BTO.	CBE at E_{Fermi} in Ge and BTO.	CBE at E_{Fermi} in Ge and BTO.
Final band edges:	Ge: $-0.1,\ 0.6$ BTO: $-2.6,\ 0.6$	$0.0,\ 0.7$ $0.0,\ 3.2$	$-0.4,\ 0.3$ $-3.2,\ 0.0$	$-0.3,\ 0.4$ $-3.2,\ 0.0$

Table 4.7: Band edges (BE) in the oxide and the semiconductor for the asymmetric and the symmetric interfaces, with and without the capping electrode, determined by the threshold method, for three sets of Hubbard U parameters. For a given configuration and a set of Hubbard U, the four numbers are VBE(Ge), CBE(Ge), VBE(BTO) and CBE(BTO), respectively. In each case the zero of energy is taken as the Fermi level and the units are electron-volts. See the text for further explanation of the determination of the final band edges. The no-oxygen interface has identical alignments as the symmetric interface.

on the applied U. For (a) the asymmetric interface without electrode, we observe that the Fermi level is pinned by the interfacial states, and the gap in either material changes with U by moving the VBE down and the CBE up by similar amounts. We conclude that for this case the midpoint of the gap in both materials is approximately fixed with respect to U and adjust the band edges accordingly. For (b) the asymmetric interface with electrode, we find that the Fermi level is pinned by the interfacial states, and that the VBE of BTO is fixed at the Fermi level. For the symmetric interface (c) without the electrode and (d) with the electrode, we find that the Fermi level is pinned at the CBE in both materials.

As a final step in determining the band edges, we have tested the accuracy of the threshold method in bulk BTO and Ge. We have run DFT+U simulations on the bulk materials where the Hubbard U is applied to Ti $3d$ states and Ge $4p$ states, with $U = 0, 1, 2, \ldots, 9$ eV. We summarize the results in Figure 4.13.

For (a) BTO the band gap is determined by the highest occupied state at the R point, and the lowest unoccupied state at the Γ point in the Brillouin zone. For very high values of U, we find that the order of lowest occupied bands switch and the gap starts to decline. Therefore DFT+U is unable to reproduce the experimental band gap of 3.2 eV. However, we find that it is possible to choose the threshold such that the band edges from the threshold method are indistinguishable from the band edges we obtain directly from the band structure. Therefore in the thin film simulations we take the band edges from the threshold method to be correct and rigidly shift them to set the gap to the experimental value.

For (b) Ge the band gap is determined by the highest occupied state at the Γ point, and the lowest unoccupied state at the L point in the Brillouin zone. However we find that for $U = 0, 1, 2$ eV the lowest unoccupied state at Γ is practically degenerate with the highest occupied state at Γ, and hence the gap vanishes. For higher values of U the gap monotonically increases as expected. We find that the threshold value that

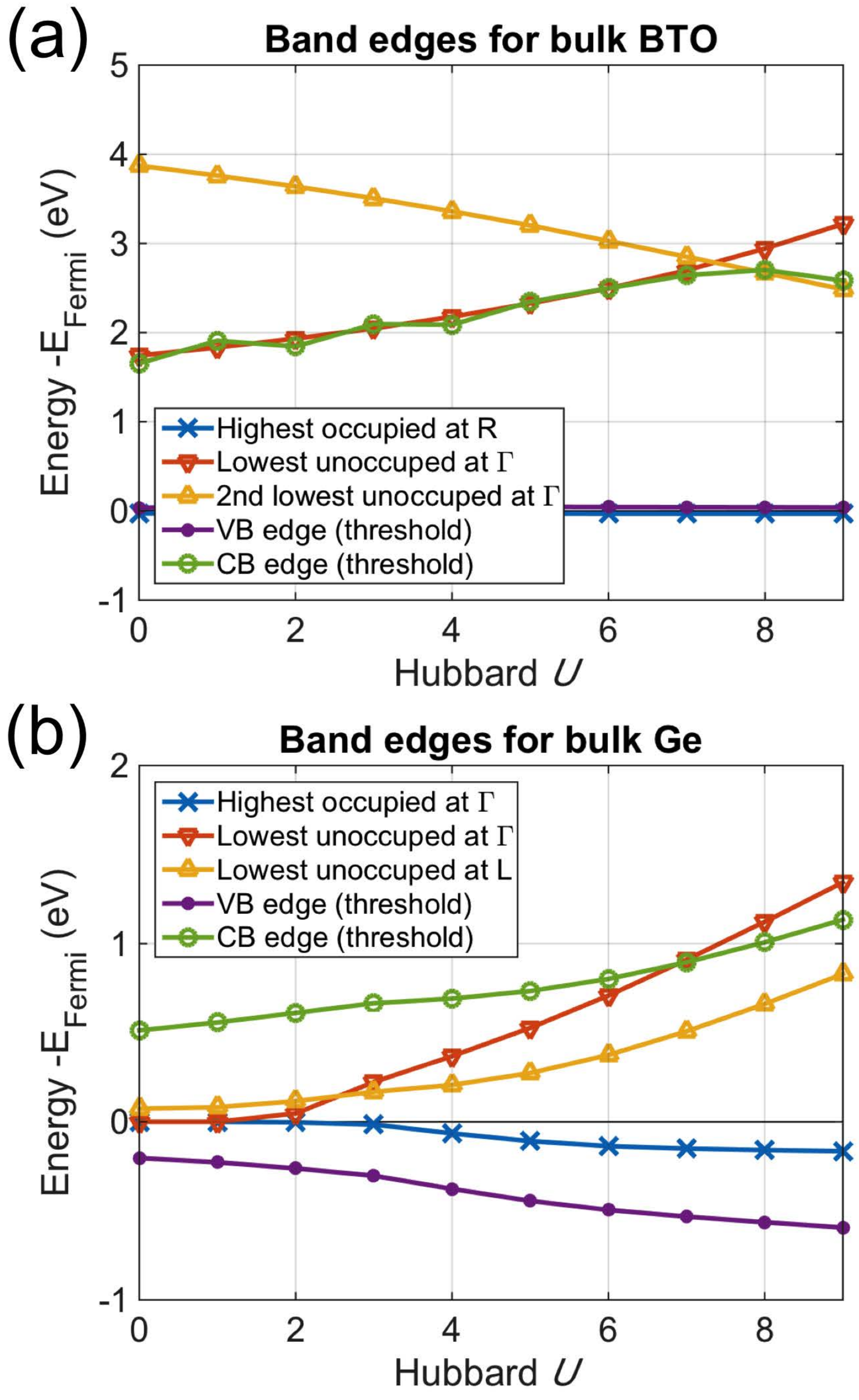


Figure 4.13: Bulk band edges of (a) BTO and (b) Ge vs a Hubbard U applied to Ti $3d$ states and Ge $4p$ states, respectively (see text for the description of the curves).

reproduces the band gap for higher U is too low to predict the band edges for lower U, because the DOS does not fall below the threshold around the Fermi level. Hence we choose a threshold which overestimates the band gap for a given U but reproduces the shapes of the VBE vs U and CBE vs U curves as closely as possible. We find that the experimental band gap is reproduced in the band structure for $U \simeq 7$, with VBE $= -0.1$ and CBE $= 0.6$. For $U = 0$ the threshold method yields VBE $= -0.2$ and CBE $= 0.6$. So in order to estimate the band edges in the film, after analyzing the Fermi level pinning and relative movement of the edges with respect to changing gaps, we determine the final positions of the Ge band edges for $U = 0$ case, and shift VBE up by 0.1 eV. Therefore we finally arrive at the band edges listed in Table 4.7 and displayed in Figure 4.8.

Chapter 5

ZrO_2 monolayers on Si(001)

5.1 Introduction

In Chapter 4, we described the scientific and technological importance of realizing ferroelectric thin films directly on semiconductors. For our test case of $BaTiO_3$ on Ge, we concluded that interfacial chemistry limits the directions in which the polarization can point, and the switching between two polarization states is feasible only if an electrode with the right work function is used. In order to overcome these limitations, in this Chapter we propose a novel approach where an oxide as thin as a monolayer is the candidate for a ferroelectric film on a semiconductor. Although a variety of monolayers may show ferroelectric behavior on a semiconductor, here we focus on monolayers of ZrO_2 on the (001) surface of Si as our candidate. In Section §5.2 we describe the growth procedure as applied by our experimental collaborators and show that an atomically abrupt interface is achieved with no significant formation of SiO_2. The ZrO_2 monolayers, which are amorphous in these experiments, are then incorporated into a gate stack with amorphous Al_2O_3. Ferroelectric behavior is then observed by C-V measurements with this gate stack. In Section §5.3, we describe our computational methods. In Section §5.4 we report on our computational study of these monolayers on the Si(001) surface. We find that these films have multiple (meta)stable structures with no significant chemical differences between them. Hence epitaxial monocrystalline growth may be challenging. However, locally switching between the (meta)stable configurations with similar atomic positions and chemical bonding may be responsible for the observed ferroelectricity in the amorphous films. This allows for a simpler growth process which is more conducive to manufacturing than an epitaxial system.

5.2 Experimental Findings

To grow ZrO_2 on silicon with an atomically abrupt interface, our collaborators first clean the silicon by removing the native oxide using 5% HF:H_2O and expose the resulting H-terminated Si surface to UV-ozone to regrow SiO_2 back on the surface to a thickness of approximately 0.5 nm. They then deposit 1 monolayer of Zr metal on top of the SiO_2 using an electron-beam (e-beam) source in a background O_2 pressure of $\sim 5 \times 10^{-9}$ Torr. To remove the SiO_2, the system is heated in vacuum until the SiO_2 desorbs, typically for < 2 min at 870 °C. Desorption of the SiO_2 is monitored using reflection high-energy electron diffraction (RHEED) and is complete when the background signal begins to decrease (Figure 5.1(a)). The gate stack is completed with a layer of amorphous Al_2O_3 deposited by thermally evaporating Al at $\sim 8 \times 10^{-7}$ Torr O_2, with the substrate kept at room temperature. To verify that the process results in an atomically abrupt interface of ZrO_2 on silicon (without SiO_2 at the interface), they use X-ray photoelectron spectroscopy (XPS). A sample with a thin (3 nm) Al_2O_3 layer on top of ZrO_2 is moved in-vacuo between the deposition and XPS analysis chambers. The effectiveness of the procedure to form an abrupt ZrO_2-Si interface is demonstrated by the dominance of the non-oxidized Si^0 component in the Si 2p spectrum, as shown in Figure 5.1(b). The relative area of the Si^{1+}

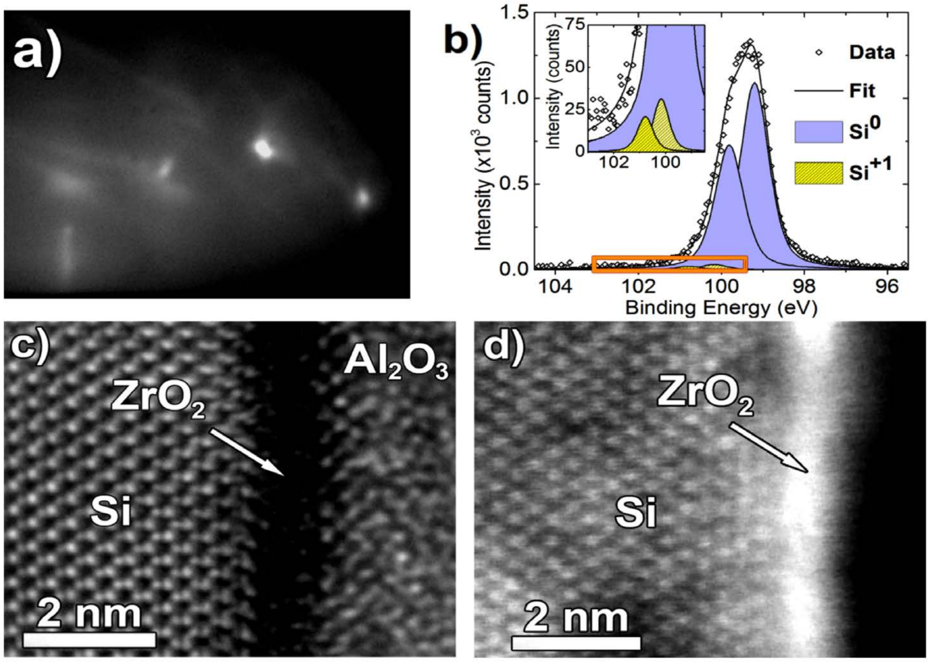


Figure 5.1: Physical properties of the Si-oxide interface. (a) RHEED image of the ZrO_2 monolayer prior to Al_2O_3 deposition. (b) Si 2p XPS spectrum of a ZrO_2/Si interface covered with a thin (3 nm) Al_2O_3 layer. The inset shows a magnification of the Si^{1+} region taken from the orange rectangle. (c) TEM and (d) HAADF STEM micrographs of the cross section of an Al_2O_3/ZrO_2/Si stack.

component is $\sim 2.5\%$ with respect to the Si^0 peak. From the XPS data, the amount of the suboxide is found to be less than 0.2 monolayers of Si^{1+} [98]. A cross-section transmission electron microscope (TEM) micrograph (Figure 5.1(c)) and a scanning TEM (STEM) high-angle annular dark field (HAADF) micrograph (Figure 5.1(d)) of an Al_2O_3 (22 nm)/ZrO_2 (1 nm)/Si sample show that the ZrO_2 layer is uniform and continuous, with a sharp interface with Si. In addition, XPS data further indicate that Zr atoms are overwhelmingly in Zr^{4+} and Zr^{3+} states, suggesting that the O:Zr ratio is near 2.

Capacitor devices are fabricated from the as-grown gate stack of Pt (50 nm)/Al_2O_3 (22 nm)/ZrO_2 (1 nm)/Si. The capacitance-voltage (C-V) behavior shown in Figure 5.2 demonstrates ferroelectric switching of the ZrO_2 interface. As a control, our experimental collaborators also deposit a gate stack with 12 nm-thick amorphous Al_2O_3 using the same procedure described above, but without ZrO_2 at the interface. The control sample (Figure 5.2, inset) shows C-V behavior typical of a metal oxide semiconductor (MOS) capacitor. At negative voltages, the Fermi level of the silicon at the oxide-silicon interface is near the Si valence band, and holes accumulate at the interface so that the measured AC capacitance is only that of the oxide gate stack. As the voltage is increased, the Fermi level at the interface moves into the silicon band gap, depleting charge at the interface, and the capacitance decreases due to the added capacitance from the silicon depletion layer.

The striking feature of the C-V characteristics for the gate stack with ZrO_2 at the interface (Figure 5.2) is the hysteretic behavior. This behavior is notably absent in the control sample that does not include the ZrO2 layer (Figure 3 inset). The sense of the hysteresis, indicated by the arrows in the figure, is consistent with ferroelectric switching of the ZrO_2 and a concomitant change in the Si Fermi level [99], as also predicted by theory (Figure 5.10). The magnitude of the hysteresis, ~ 0.4 V, is slightly smaller than the ~ 0.6 eV predicted by DFT (Figure 5.10), which may be because only a fraction of the ZrO_2 in the amorphous structure is switching. The hysteresis in the C-V behavior shows a switchable dipole at the interface: the field and polarization of the ZrO_2 point downwards for negative voltages. As the field across the ZrO_2 increases in the positive direction during the voltage sweep, the polarization of the ZrO_2 flips from down to up. This change in polarization induces a positive switchable field at the interface and shifts the C-V curve toward negative voltages, as observed in Figure 5.2.

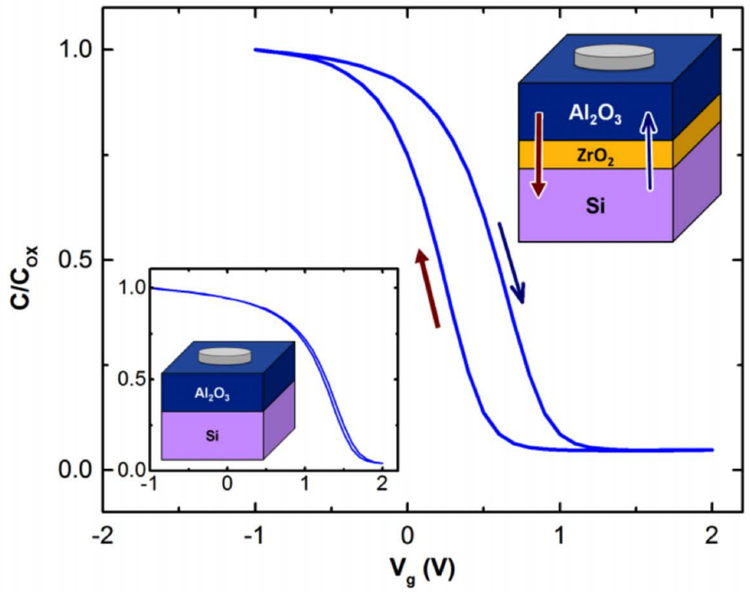


Figure 5.2: The ferroelectric behavior of ZrO_2 on Si. A capacitance-voltage (C-V) curve of a Pt/Al_2O_3 (22 nm)/ZrO_1 (1 nm)/Si MOS capacitor, with arrows indicating the scan direction. The top inset shows a schematic device structure with the polarization corresponding to the arrows marking the sweep direction. The bottom inset shows a C-V curve of a control sample with 12 nm Al_2O_3 and no ZrO_2 layer.

5.3 Computational methods

We theoretically model this system using density functional theory (DFT) with the Perdew–Burke–Ernzerhof generalized gradient approximation (PBE GGA) [18] and ultrasoft pseudopotentials [19]. We use the QUANTUM ESPRESSO software package [65]. A 35 Ry plane wave energy cutoff is used to describe the pseudo Kohn–Sham wavefunctions. We sample the Brillouin zone with an $8 \times 8 \times 1$ Monkhorst-Pack k-point mesh (per 1×1 in-plane primitive cell) and a 0.02 Ry Marzari–Vanderbilt smearing [64]. A typical simulation cell consists of 8 atomic layers of Si whose bottom layer is passivated with H and a monolayer of ZrO_2 (see Figure 5.3). Periodic copies of the slab are separated by ~ 12Å of vacuum in the z-direction. The in-plane lattice constant is fixed to the computed bulk Si lattice constant of 3.87Å. In general, the slab has an overall dipole moment that might artificially interact with its periodic copies. In order to prevent this unphysical effect, we introduce a fictitious dipole in the vacuum region of the cell which cancels out the electric field in vacuum and removes such interactions [66]. All atomic coordinates are relaxed until the forces on all the atoms are less than $10^{-3} Ryd/a_0$ in all axial directions, where a_0 is the Bohr radius (except the bottom 4 layers of Si which are fixed to their bulk positions to simulate a thick Si substrate). We use the nudged elastic bands (NEB) method with climbing images [25] to compute the transition energy barrier between different metastable configurations.

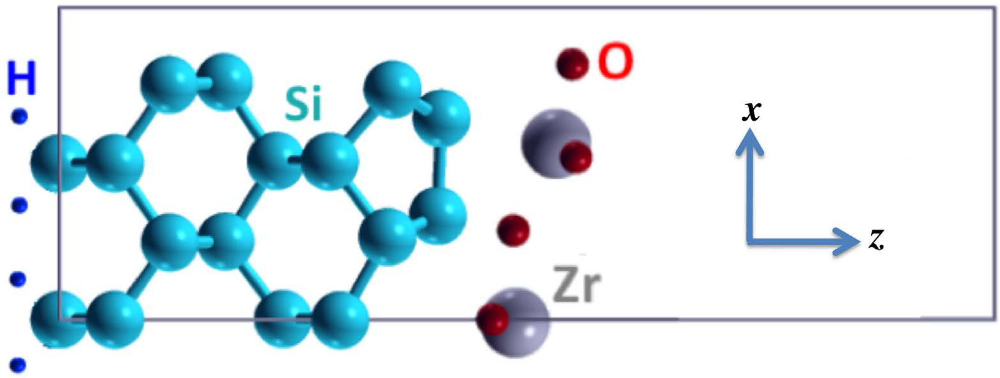


Figure 5.3: A typical simulation supercell with 2×1 in-plane periodicity. The bottom 2 layers of Si are passivated by H and fixed to bulk coordinates. There is ~ 12Å of vacuum along the $z-$direction to separate periodic copies.

5.4 Results

5.4.1 Free standing ZrO_2 monolayers

5.4.1.1 Background: bulk zirconia

Bulk ZrO_2 is observed in three structural phases. The high symmetry cubic phase (space group: $Fm\bar{3}m$) is shown in Figure 5.4. The lower symmetry tetragonal ($P4_2/nmc$) and monoclinic ($P2_1/c$) phases are obtained by continuously breaking the symmetries of the cubic phase. All three configurations are centrosymmetric and hence not ferroelectric. However, this binary oxide has a *layered structure* (along low-index directions) in which the cations and anions lie in different planes, which, in thin film stoichiometric form, would cause ultrathin ZrO_2 films to be polar. For instance, in Figure 5.4 a horizontal monolayer of ZrO_2 could be formed by the zirconium atoms in Layer 3, with (a) the oxygen atoms in Layer 2, or with (b) the oxygen atoms in Layer 4, or with (c) half of the oxygen atoms in each of Layer 2 and Layer 4. Before relaxing the atoms in these hypothetical monolayers, in case (a) the resulting polarization would be upward, in case (b) it would be downward, and

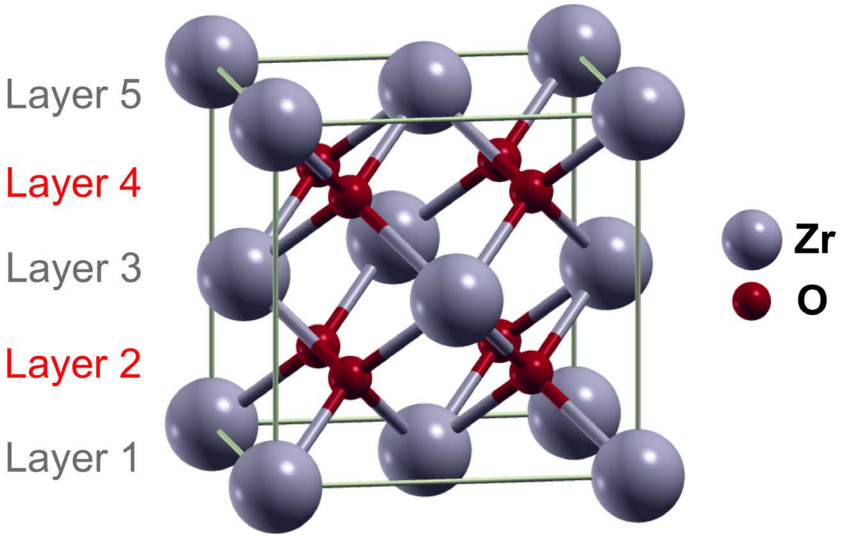


Figure 5.4: The high symmetry cubic phase (Fm$\overline{3}$m) of bulk ZrO_2. Atomic layers are labelled 1 through 5, where the odd (even) layers correspond to cation (anion) planes.

in case (c) it would be zero. This intrinsic layered structure, which is also preserved in the tetragonal and the monoclinic phases of zirconia, is a fundamental reason why ZrO_2 is an excellent candidate to promote ferroelectricity as a buffer between Si and a perovskite oxide such as $SrTiO_3$.

5.4.1.2 Structure of free standing monolayers

In order to check if this richness of structure due to the layered nature of the bulk is retained in the ultrathin film, we have simulated free standing ZrO_2 monolayers. A monolayer formed by a (001) plane of cubic ZrO_2 would have a square lattice with size 3.61 Å (from our DFT computations). To lattice match the Si substrate, we simulate the monolayers at the lattice constant of the Si(001) surface, which we find to be 3.87 Å. We have searched for minimum energy configurations for 1×1, 2×1, 2×2 and $c(4 \times 2)$ unit cells of monolayer ZrO_2 which are the periodicities of the low energy reconstructions of the bare Si(001) surface, as we shall discuss in section 5.4.2.

We find that the lowest and the second lowest energy configurations of the ZrO_2

monolayer are 2×1 and 1×1, respectively, as shown in Figure 5.5. The chief difference between the two configurations is that the lowest energy structure, named A, has a buckling of zirconiums in the $2\times$ in-plane direction, while for the second lowest energy structure, named B, the Zr are coplanar. We find that $E(B) - E(A) = 0.07$ eV per ZrO_2. Both of these configurations are inversion symmetric and hence non-polar. However, because neither A or B is symmetric with respect to the mirror plane reflection $z \to -z$, there are two more geometrically distinct minima, named $\overline{A}$ and $\overline{B}$, which are shown in Figure 5.5. $\overline{A}$ and $\overline{B}$ are obtained from A and B, respectively, by the mirror reflection. Notice that $\overline{A}$ can be obtained from A also by translating in the $2\times$ direction by half a 2×1 cell. However, since the underlying substrate will have at least 2×1 periodicity, this translation would not leave the entire system invariant.

5.4.1.3 Energy landscape of free standing monolayers

In order to analyze these configurations further, we parametrize the energy landscape of free standing ZrO_2 monolayers by using two coordinates: $z_1 \equiv z(\mathrm{Zr}_2) - z(\mathrm{Zr}_1)$ and $z_2 \equiv z(\mathrm{O}_1) - z(\mathrm{Zr}_1)$, where the atoms Zr_1, Zr_2 and O_1 are labelled in Figure 5.5 and are vertically aligned for all the displayed structures. Note that the structures B and $\overline{B}$ are treated as 2×1 structures in this analysis. To explore the energy landscape, we have made a 9×9 grid of (z_1, z_2) values and computed energies for structures whose z_1 and z_2 are fixed but all other coordinates are relaxed. In Figure 5.6, we plot the energy landscape using darker (lighter) colors to represent lower (higher) energies. The coloring is done by MatLab's interpolation scheme based on the DFT results on an equally spaced 9×9 grid. We also label the four metastable configurations on the landscape. The energies are reported for 2×1 cells such that $E(A) = 0$.

In Figure 5.6 we also present the minimum energy transition paths between these energy minima. We have studied these transitions using the NEB method with climbing images [25]. There are 6 pairs of metastable configurations and hence 6 transition

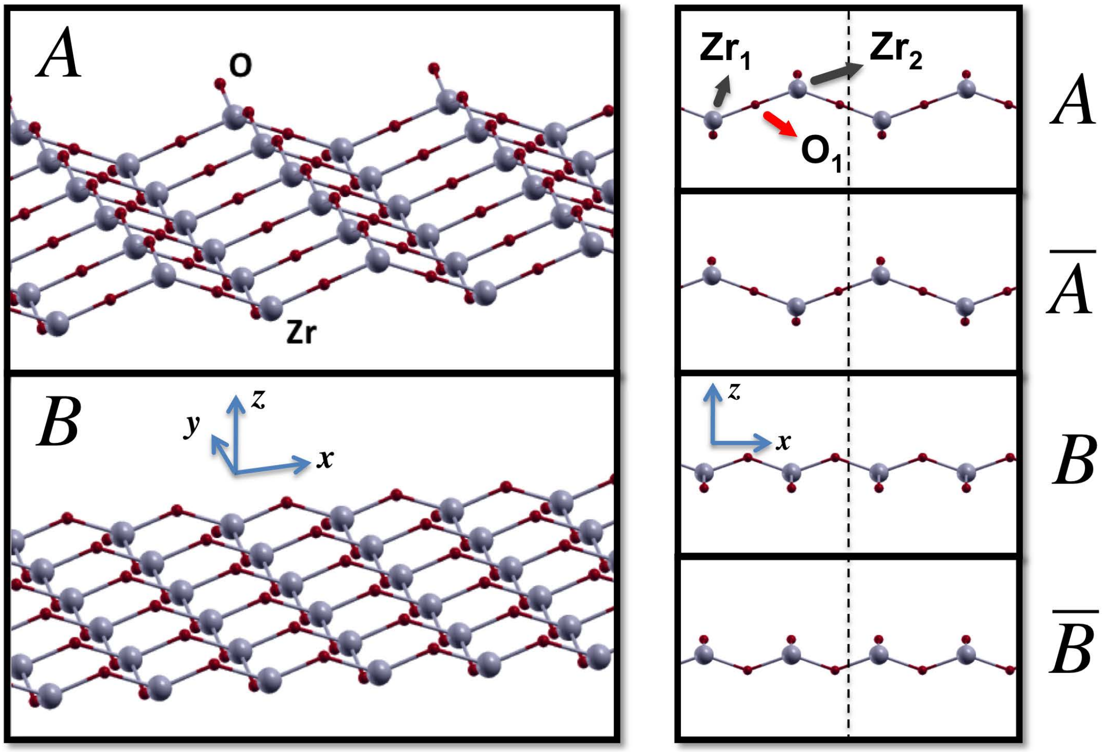


Figure 5.5: The lowest energy configurations of the free standing ZrO_2 monolayer. Structure B lies at 0.07 eV per ZrO_2 with respect to structure A. On the right, all four geometrically distinct metastable configurations are shown. $\overline{A}$ and $\overline{B}$ are obtained from A and B, respectively, by reflection in the $z = 0$ plane. For each structure, two copies of 2×1 cells are displayed and a vertical dashed line is drawn to separate the copies.

paths: $A \leftrightarrow \overline{A}$, $A \leftrightarrow B$, $A \leftrightarrow \overline{B}$, $\overline{A} \leftrightarrow B$, $\overline{A} \leftrightarrow \overline{B}$ and $B \leftrightarrow \overline{B}$. However, as seen from the figure, the transition paths of $A \leftrightarrow \overline{A}$ and $B \leftrightarrow \overline{B}$ go through another energy minimum and hence can be expressed in terms of the remaining 4 transitions. We have found that all of the four transitions have an energy barrier of hight 1.04 eV per 2×1 cell. These four saddle points, shown as diamonds in Figure 5.6, are related by reflection and/or translation operations, and hence are physically equivalent.

To sum up, we have found that as a free standing monolayer in vacuum, ZrO_2 is not polar but has two physically distinct stable configurations. In the presence of a surface that breaks the $z \to -z$ symmetry, the two configurations may turn into two new configurations that are differently polarized.

5.4.2 ZrO_2 monolayers on Si(001)

5.4.2.1 Bare Si(001) surface

To study the behavior of zirconia on Si(001), first we review the structure of the bare Si(001) surface. It is well known that on the Si(001) surface, neighboring Si atoms pair up to form dimers [100, 101], and we find that dimerization lowers the energy by 1.45 eV per dimer. The dimers can buckle (i.e., the two Si forming the dimer will not have the same out-of-plane z coordinate) which lowers their energy. If nearby dimers buckle in opposite ways, higher order reconstructions occur. We summarize the energies of these reconstructions in Table 5.1 (we refer the reader to the cited works for detailed descriptions of the surface configurations). In short, there is a strong drive for the surface Si atoms to dimerize (transition from a 1×1 to a 2×1 unit cell) and a weaker energetic drive to organize the dimers into structures with periodicities larger than 2×1. Because the metastable configurations of the ZrO_2 monolayers we have found are 2×1 or smaller, we have limited our search for Si/ZrO_2 interfaces to 2×1 simulation cells.

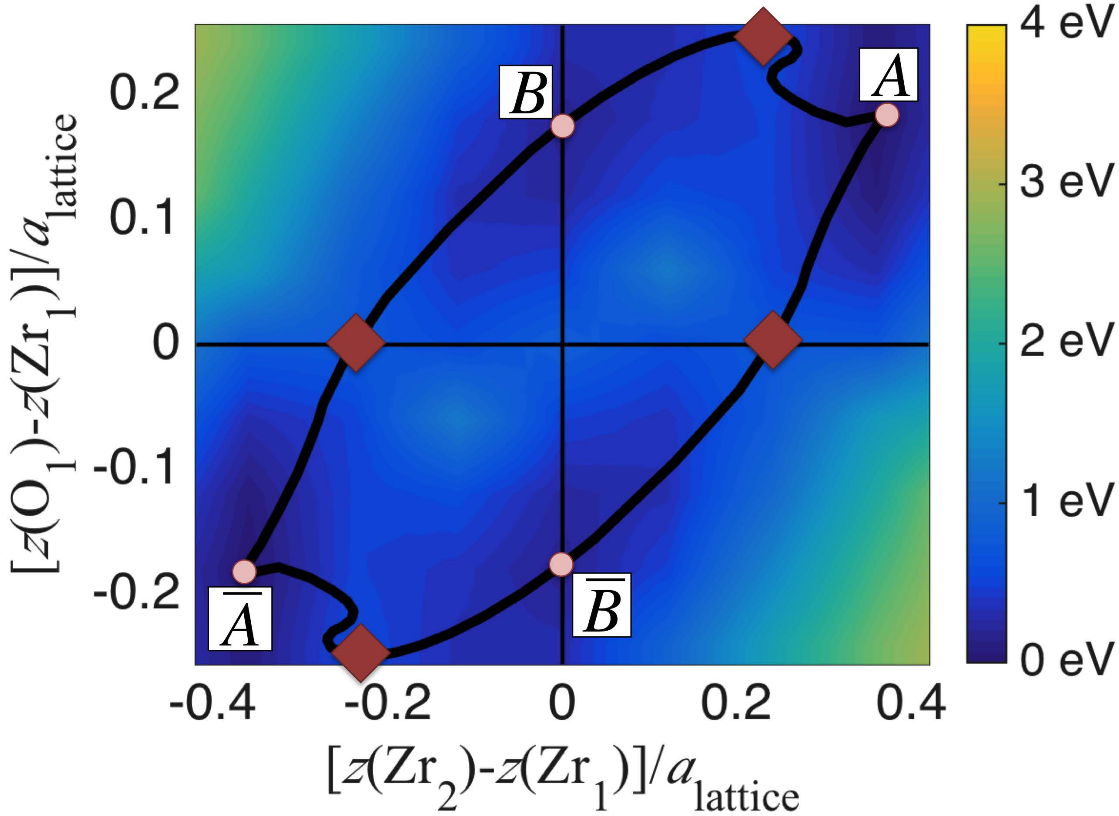


Figure 5.6: The energy landscape of the free standing ZrO_2 monolayer, as parametrized by a pair of coordinates $z_1 \equiv z\left(\mathrm{Zr}_2\right) - z\left(\mathrm{Zr}_1\right)$ and $z_2 \equiv z\left(\mathrm{O}_1\right) - z\left(\mathrm{Zr}_1\right)$ (See Figure 5.5 for labelings of the atoms). a_{lattice} is the computed lattice constant of silicon and is equal to 3.87 Å. All four local energy minima as well as the minimum energy transition paths between them are shown. The saddle points on the landscape (i.e., the transition states) are shown as diamonds. The zero of energy is taken to be the energy of structure A. All transition states lie at the same energy because they are related by reflection/translation operations. The energy landscape is computed by DFT on a 9×9 grid and then interpolated by MatLab to produce the smooth colored plot.

Si surface	Energy (eV/dimer)	Ref. [100]	Ref. [101]
flat $p(2 \times 1)$	$\equiv 0.00$	$\equiv 0.00$	$\equiv 0.00$
buckled $p(2 \times 1)$	-0.20	-0.12	-0.13
buckled $p(2 \times 2)$	-0.28	-0.17	-0.23
buckled $c(4 \times 2)$	-0.07	-0.17	-0.24

Table 5.1: Energies of the lowest energy Si(001) surface reconstructions per dimer. Two theoretical references are presented alongside our computed results. See the cited works for details of the listed reconstructions.

5.4.2.2 Structure of the monolayers on silicon

We have searched the configuration space for ZrO_2 on Si(001) as follows: First, we have created a $3 \times 3 \times 2$ grid of points inside the 2×1 in-plane unit cell on top of the bare Si surface where a Zr atom is placed (the 3×3 grid corresponds to points in the $xy-$plane and the $\times 2$ corresponds to the vertical distance from the substrate). A flat and high symmetry 1×1 zirconia monolayer is generated such that it includes this Zr atom. For each such structure, the atoms in the Si surface layer and the ZrO_2 monolayer are randomly and slightly displaced to generate 5 initial positions. This procedure, which yields $3 \times 3 \times 2 \times 5 = 90$ configurations, is done for dimerized and non-dimerized Si surfaces, so that there are 180 initial configurations in total. We have then relaxed all the atoms in ZrO_2 and the top 4 layers of silicon substrate to reach local energy minima.

We present the five lowest energy structures we have obtained in Figure 5.7. The horizontal axis is a quantity that describes the ionic polarization of the ZrO_2 monolayer and is defined as the mean vertical Zr-O separation $\delta z \equiv \overline{z\,(\mathrm{Zr})} - \overline{z\,(\mathrm{O})}$, where over-bars mean averaging of the coordinate over the structure. The vertical axis is the energy in eV per 2×1 cell measured with respect to the lowest energy structure, labelled $S1$. The energies of $S1$ through $S5$ are listed in Table 5.2.

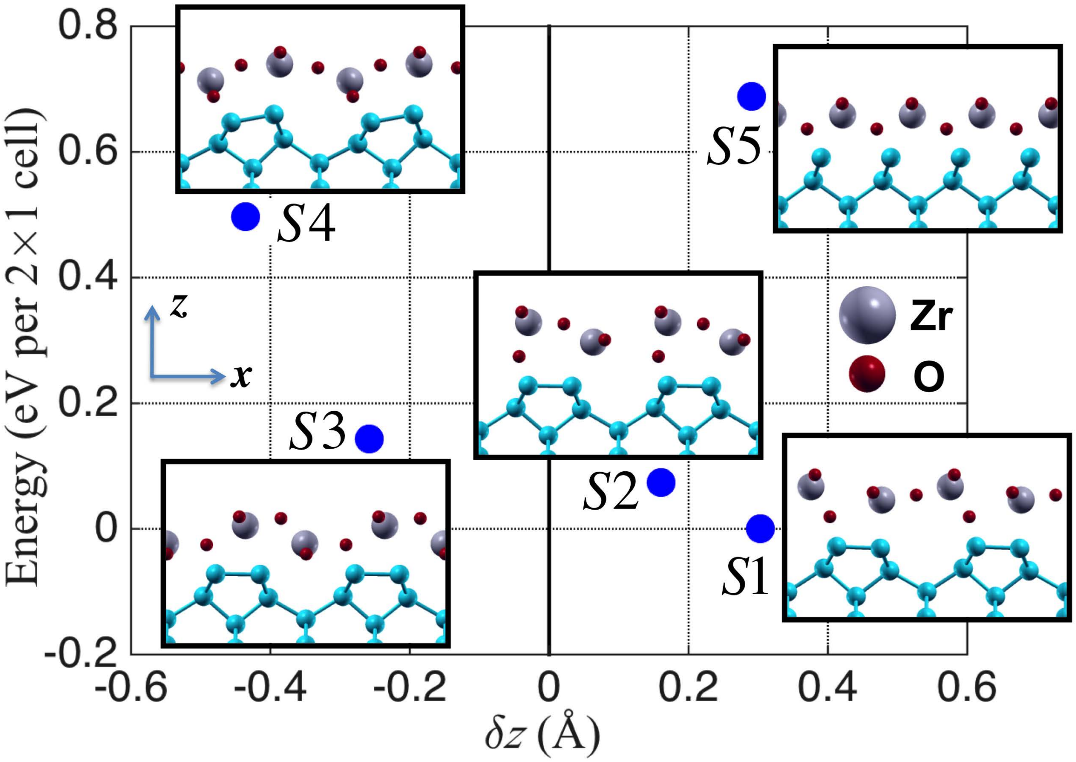


Figure 5.7: Five lowest energy configurations of ZrO_2 monolayers on Si. $\delta z \equiv \overline{z\,(\mathrm{Zr})} - \overline{z\,(\mathrm{O})}$ is a measure of ionic out-of-plane polarization for the monolayers. Energies are listed in eV per 2×1 in-plane cell measured with respect to the lowest energy structure $S1$.

	$S1$	$S2$	$S3$	$S4$	$S5$
Energy (eV per 2×1 cell)	$\equiv 0.00$	0.07	0.14	0.50	0.69

Table 5.2: Energies of the five lowest energy configurations of ZrO_2 monolayers on Si as they are labelled in Figure 5.7.

First, we observe that the metastable configurations lie on both sides of the $\delta z = 0$ line, which means that there is no polarization direction that is strongly preferred. Second, we find that the four lowest energy structures have a 2×1 periodicity with intact Si dimers. (In addition to $S5$, we have found three more 1×1 structures at energies higher than 1 eV that are not shown.) The energy difference of 0.69 eV per dimer between the lowest energy 1×1 and the lowest energy 2×1 structures (i.e. $S5$ and $S1$) is half of the energy of dimerization on the bare Si surface. Moreover, the length of the dimer in $S1$ is 2.42 Å which is longer than the 2.31 Å on the bare surface. Therefore, in general, the Si dimers are weakened but not broken by the ZrO_2 monolayer.

Third, we notice that for each configuration shown in Figure 5.7, a physically equivalent configuration is obtained by a mirror reflection by the $yz-$plane, which doubles the number of metastable structures in the configuration space. For our analysis of transitions between these configurations, we make the reasonable assumption that silicon dimers remain intact during the transition between two dimerized configurations. Hence, we reflect the atomic positions through a $yz-$plane which keeps the dimers in place in order to obtain the geometrically equivalent (but physically identical) set of structures $\overline{S1}$, $\overline{S2}$ etc.

5.4.2.3 Transitions between low energy states

We have computed the minimum energy transition paths between the three lowest energy configurations and their symmetry related counterparts ($S1, \overline{S1}, S2, \overline{S2}, S3, \overline{S3}$). When applying the NEB method to find transition states, each atom in the initial configuration is mapped to an atom in the final configuration. In principle, all possible matching choices should be attempted in order to find all inequivalent transition paths and barriers. However, this is impractical (and not physically necessary). For the case of free standing ZrO_2, in all the minimum energy configurations, all atomic

(x, y) coordinates line on a square grid, and by making the reasonable assumption that atoms do not swap sites during the transition, we can dramatically reduce the number of possible transition paths under consideration. Hence, we matched each atom in the initial configuration with the atom that sits at the same site in the final configuration in order to perform the NEB calculations. Even though no fixed square grid exists for the ZrO_2 monolayer on Si cases that applies to all the configurations, similar considerations are possible: (1) For the six configurations of interest, both Zr atoms and two of the O atoms in a unit cell align along the $y-$direction with the Si dimers ($y = 0$), and the other two O atoms lie half way between consecutive dimers ($y = 0.5a_{\mathrm{lat}}$). Both along the $x-$ and the $y-$directions, atomic chains of ...-Zr-O-Zr-O-... exist in all cases. So for each configuration, we can make a square grid in the $xy-$plane such that one Zr per cell sits at a lattice site and the other atoms are very close to the other lattice sites. For each transition process, the grid is assumed only to shift in the $x-$direction. (2) Because of the high energy cost of breaking Si dimers on the bare Si(001) surface, we assume that the dimers remain intact during a transition. (3) We assume that ...-Zr-O-Zr-O-... chains along the $y-$direction remain intact during a transition, so no movement in the $y-$direction is considered.

By using these constraints, we can reduce the number of possible matchings to four for each transition. We demonstrate these choices for the transition $S1 \to S2$ in Figure 5.8. The final state $S2$ is displayed upside down in order allow for a clear demonstration of ZrO_2 matchings. In the left panel, ...-Zr-O-Zr-O-... chains along the $y-$direction are circled by blue dashed rings. There are two possible ways in which the chains in $S1$ can be matched to the chains in $S2$ that do not cause large scale rearrangements. One of these matchings is shown as solid arrows, and the other is shown as dotted arrows. In the right panel, the same exercise is repeated for the remaining oxygens (circled by red dashed rings). Therefore there are $2 \times 2 = 4$ matchings in total. Note that the reverse processes correspond to the set of matchings

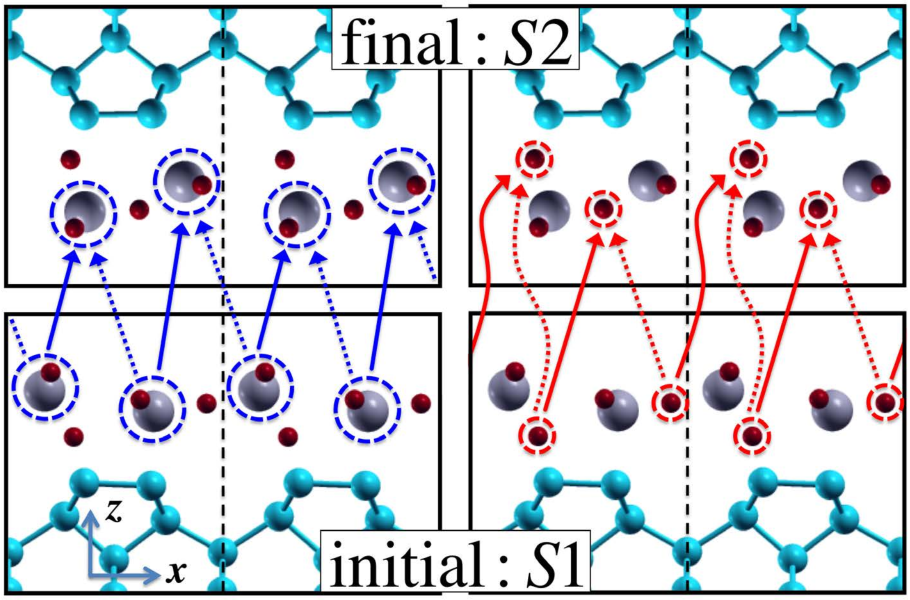


Figure 5.8: The possible matchings for the $S1 \rightarrow S2$ transition for the NEB simulation. The $S2$ structure is displayed upside down to allow for a clear demonstration. On the left panel, two possible choices for the two Zr-O pairs (or chains) in the $S1$ unit cell to be matched to the Zr-O pairs (or chains) in the structure $S2$ are shown. The set of solid arrows corresponds to one choice, and the set of dotted lines corresponds to another. Similarly, two choices for the remaining oxygens are displayed. See text for further details of the described matchings. Two copies of 2×1 cells are are shown and a dashed line is drawn to separate the copies.

that obey our rules for the transition $S2 \rightarrow S1$.

The resulting smallest energy barriers are listed in Table 5.3. Notice that the nine listed transitions cover all the possible transitions because, e.g., $S1 \leftrightarrow \overline{S2}$ is symmetry related to $\overline{S1} \leftrightarrow S2$. We observe that the transitions between the unbarred states are about 1 eV smaller than the transitions between unbarred and barred states. This is understood as follows: For all six structures, there is one oxygen per cell which binds to a silicon atom. The transitions that leave that oxygen in place (such as the dotted arrows in the right panel of Figure 5.8) have lower energy barriers. A transition between an unbarred state and a barred state necessarily involves displacing that

Transition	$E_{\text{barrier}}\,(\rightarrow)$ (eV)	$E_{\text{barrier}}\,(\leftarrow)$ (eV)
$S1 \leftrightarrow \overline{S1}$	1.63	1.63
$S1 \leftrightarrow S2$	0.79	0.71
$S1 \leftrightarrow \overline{S2}$	1.60	1.52
$S1 \leftrightarrow S3$	0.79	0.65
$S1 \leftrightarrow \overline{S3}$	1.60	1.46
$S2 \leftrightarrow \overline{S2}$	2.48	2.48
$S2 \leftrightarrow S3$	0.23	0.17
$S2 \leftrightarrow \overline{S3}$	1.57	1.51
$S3 \leftrightarrow \overline{S3}$	1.77	1.77

Table 5.3: Transition barriers, calculated via the NEB method, between pairs of low energy configurations of ZrO_2 monolayers on Si(001). Energy barriers are reported in eV per 2×1 cell.

oxygen and binding it to the other Si atom on the surface. Therefore a low energy path is not possible in such a case.

Focusing on the three low energy transitions, i.e. $S1 \leftrightarrow S2$, $S1 \leftrightarrow S3$ and $S2 \leftrightarrow S3$, we plot energy vs δz curves in Figure 5.9. The transition state of $S2 \leftrightarrow S3$ (dotted curve) and the shared transition state of $S1 \leftrightarrow S2$ and $S1 \leftrightarrow S3$ (solid curves) are marked by diamonds on the plot and their configurations are displayed. During these transitions, the oxygen atom that has bonded to a silicon (circled by red dashed rings in the figure) remains in place, while the remaining 5 atoms in the ZrO_2 layer (inside the blue dashed rounded rectangles) move in concert. Because this movement does not significantly alter the chemistry of the interface, the energy barriers are relatively low.

Because of the rich landscape of stable configurations at low energy with similar chemical bonding and small structural differences, we predict that growing single crystalline epitaxial films of ZrO_2 on Si(001) should be challenging. However, epitaxy

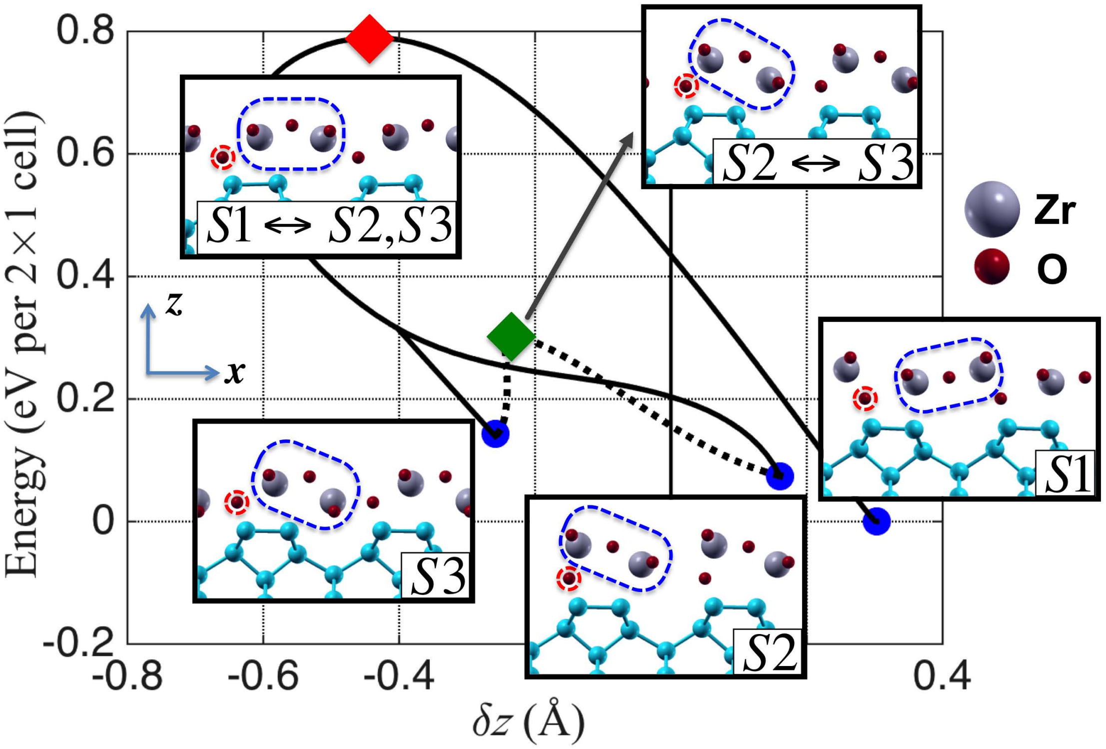


Figure 5.9: Three lowest energy configurations of ZrO_2 monolayers on Si and the transition paths between them calculated via the NEB method. The solid curve corresponds to the transitions $S1 \leftrightarrow S2, S3$ that share a transition state denoted by a red diamond. The dotted curve corresponds to the transition $S2 \leftrightarrow S3$ which has a transition state denoted by a green diamond. The circled oxygen atoms remain in place during the transitions, and the circled groups of five atoms move as a block with small internal displacements.

may not be a necessary condition for ferroelectricity in this system. A close examination of the structures shown in Figure 5.9 indicates that the symmetry of the silicon surface, as well as the inherently rumpled structure of ZrO_2, give rise to the switchable polarization. The switching of the dipole occurs by a continuous displacement of a group of atoms in the unit cell, while one oxygen remains in place. No significant chemical change occurs during these transitions. We note that open channels in the dimerized (001) face of silicon allow for the motion of the oxide atoms lacking silicon nearest neighbors, which stabilizes the three low energy polar ZrO_2 structures.

5.4.2.4 Coupling of polarization to electronic states in Si

In addition to the theoretical prediction that the three lowest energy structures may coexist in monolayer form, in section 5.4.3 we will explain why, at temperatures of interest to us, structures $S2$ and $S3$ may dominate the film. Because of the large difference in polarization together with a low energy barrier between these structures, we believe that the polarization switching described in Section §3.2 should correspond to switching between $S2$ and $S3$.

A first corroboration involves showing that the change in the silicon Fermi level observed in the experiment is comparable with our theoretical prediction. In Figure 5.10, we plot the density of states (DOS) of the ZrO_2/Si system projected onto an interior layer of the Si substrate for the cases of interface structures $S2$ and $S3$. We set the energy of the Si valence band edge (VBE) of $S2$ to zero and align the vacuum energy level in $S3$ to the vacuum energy energy in $S2$. We find a 0.6 eV VBE shift in Si, which is larger than, but quite comparable to, the experimental value of 0.4 eV. We believe that this is due to the fact that the experimental monolayers are not epitaxial but amorphous with multiple structural motifs present, so that application of the electric field is not as effective at polarization switching as is assumed in the clean, epitaxial and ordered theoretical simulations.

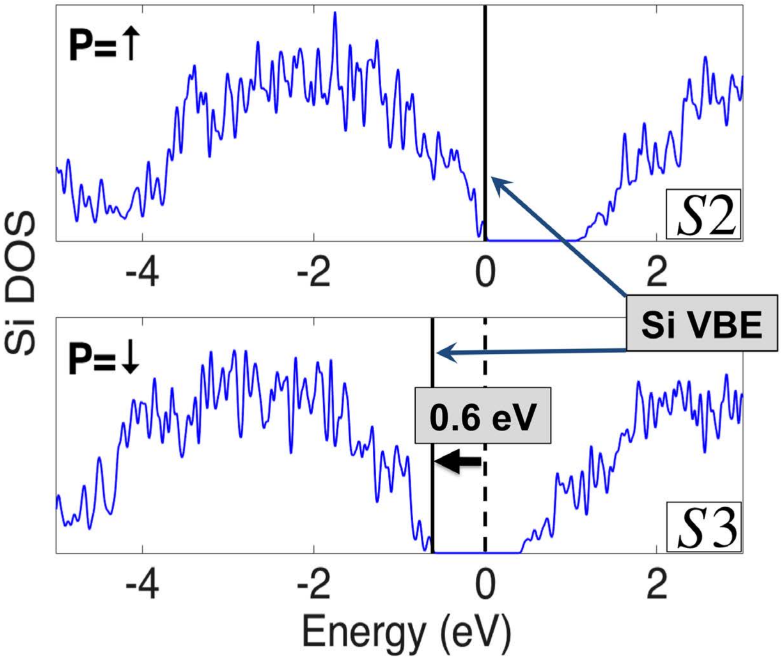


Figure 5.10: Density of states in an interior Si layer with the ZrO_2 film in its upwardly polarized ($S2$) and downwardly polarized ($S3$) forms. There exists a valence band edge (VBE) shift between the "up" state (top) and the "down" state (bottom).

5.4.3 Domain energetics

Up to this point, our theoretical study of the ZrO_2 monolayers on the Si(001) surface has shown that (meta)stable configurations with varying polarizations are present. We have also demonstrated that transitions between some of the lowest energy configurations do not require complicated rearrangements of atoms and have low energy barriers. Because of these findings as well as the fact that the experimental monolayer is amorphous, we expect there to be a multi-domain character to these monolayers at or near room temperature ($k_{\mathrm{B}}T = 0.026$). However, directly calculating the energy of a multi-domain region of the system an area larger than a few unit cells is not feasible. In this Section, we describe an approximate method to compute the energies of arbitrary regions of multiple domains, and use Monte Carlo simulations to find thermodynamic ground states at finite temperatures.

5.4.3.1 Domain wall energies

In order to investigate the behavior of finite domains, we have developed a lattice model where every 2×1 in-plane cell is treated as a site in a two dimensional lattice which couples to its neighbors via an interaction energy. Such models have been proposed for other two dimensional systems [102]. Such a model is reasonable if the interface (domain wall) between domains of different states is sharp, i.e., the atomic positions a few unit cells away from a domain boundary are indistinguishable from the atomic positions in the center of the domain. To find the degree of locality, we have computed domain wall energies as a function of domain size.

Sample simulation arrangements are shown in Figure 5.11. In (a) and (b), domain walls along the $y-$ and $x-$directions are formed, respectively, between the configurations $S1$ and $S2$. Three unit cells of $S1$ and $S2$ each are generated and attached together to build larger cells to simulate the domain walls: 12×1 and 2×6 cells to simulate the boundaries along the $y-$ and $x-$directions, respectively. In each of the 3 unit wide domains, the center unit is fixed to the atomic configuration of the uniform system. In Figure 5.11, for the $S1$ domain, the atoms in the unit labelled $S1$ are fixed, and the atoms in the units $S1l$ and $S1r$ are relaxed. The same is true for $S2$, but for clarity, fixed units of $S2$ are displayed on both sides. We then compute the domain wall energy between $S1$ and $S2$ by subtracting $3E\left(S1\right)+3E\left(S2\right)$ from the total energy of this supercell and dividing by two. We have checked for a few test cases that increasing the domain width from 3 to 5 changes the domain wall energies by small amounts on the order of 1-10 meV while typical domain wall energies are larger than 100 meV (see Table 5.4). This, together with visualization of the resulting structures, convinces us that the domains are sufficiently local for us to treat the domain walls as being sharp. Note that there are two inequivalent boundaries between $S1$ and $S2$ along a given direction. In Figure 5.11, these boundaries are shown as red and blue dashed lines. Due to the periodicity of simulation cells, it is not possible to separate the energies of these two boundaries, so we are forced to assume that their energies are equal.

The final step in determining the domain boundary energies is to survey the configuration space at a given boundary. For that purpose, for each domain boundary we have generated a number of initial configurations depending on the direction of the boundary:

- For a boundary along the $y-$direction such as in Figure 5.11(a), we have generated five initial configurations as follows. For each domain state (e.g., $S1$ or $S2$), we have labelled the Zr-O pairs along the $y-$direction and the remaining oxygens and numbered them in an increasing order in the $x-$direction. In the figure, the labelling for states $S1$ and $S2$ is shown. Note that for each cell, the sequence starts with a Zr-O pair and ends with an O atom. Hence in some cases the oxygen labelled 4 lies beyond the unit cell to which it belongs, such as in $S2$. To build a domain boundary such as the $S1r$-$S2l$ (shown as a blue dashed

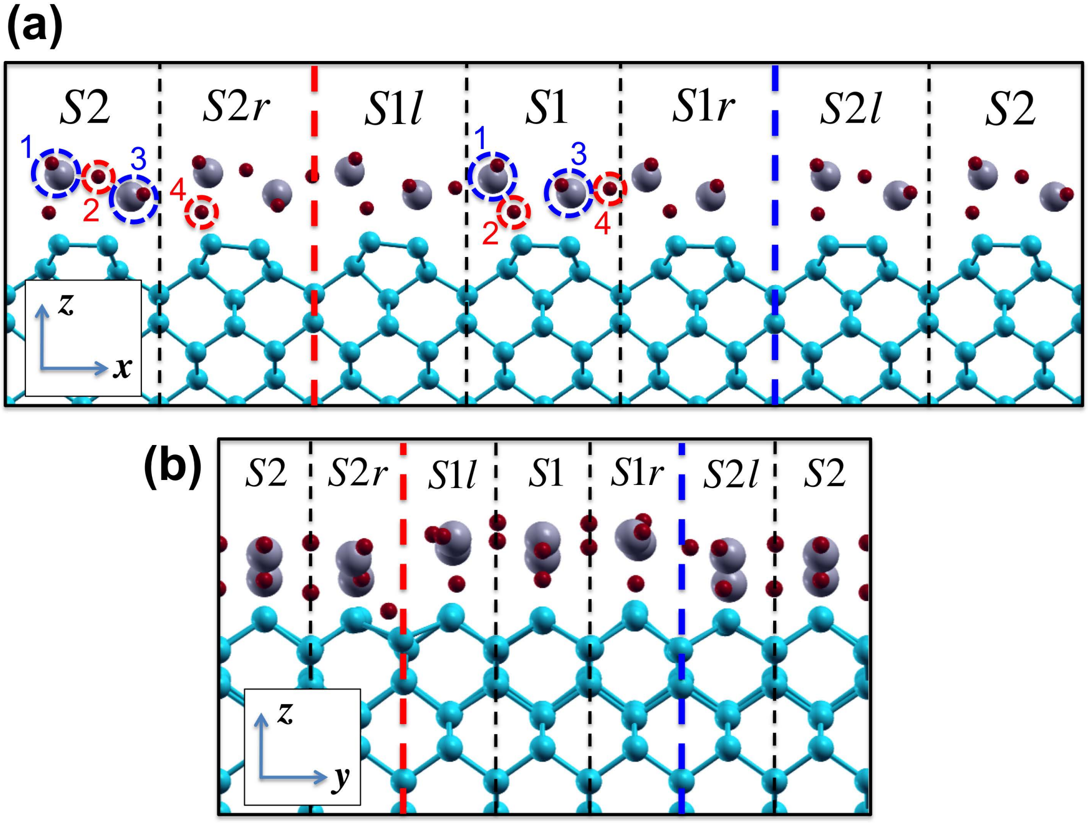


Figure 5.11: Simulation arrangements to compute the domain boundary energies between $S1$ and $S2$. (a) 3 cells each of $S1$ and $S2$ are stacked along the $x-$direction to form straight domain boundaries in the $y-$direction. The numberings of atomic groups within the unit cells are displayed using dashed circles. The boundary on the right (blue) is initially built by the atomic groups 1, 2 and 3 from $S1$ and 4 from $S2$ in the unit cell to the left of the boundary (labelled $S1r$), and the atomic groups 1, 2, 3 and 4 from $S2$ in the unit cell to the right of the boundary (labelled $S2l$). The boundary on the left (red) is constructed to preserve the number of atomic groups from each cell. (b) 3 cells each of $S1$ and $S2$ are stacked along the $y-$direction to form straight domain boundaries in the $x-$direction. Fully relaxed boundary configurations are shown.

line), we first place the atomic groups numbered $1-4$ from $S1$ to the left hand side of the boundary, and the atomic groups numbered $1-4$ from $S2$ to the right hand side of the boundary. This constitutes our first initial configuration. The second configuration is obtained by replacing atom 4 from $S1$ on the left hand side by atom 4 from $S2$. The third is obtained by replacing both group 3 and atom 4 from $S1$ by 3 and 4 from $S2$. The fourth choice is replacing atomic group 1 from $S2$ on the right hand side by atom 1 from $S1$; and lastly the fifth choice is replacing 1 and 2 from $S2$ by 1 and 2 from $S1$. The opposite operation is performed at the other boundary such as $S2r$-$S1l$ (shown as a red dashed line). We then take the smallest of the five domain energies computed as the final energy. Note that the relaxed structure shown in the Figure 5.11(a) for the $S1$-$S2$ domain boundaries is obtained via choice 2 for the $S1r$-$S2l$ boundary.

- For a boundary along the $x-$direction such as in Figure 5.11(b), we have generated a few initial configurations by slightly and randomly displacing the two oxygen atoms at the boundary along the y direction in order to break the $y \to -y$ symmetry inherent to these structures.

5.4.3.2 Construction of a lattice model

With the library of domain boundary energies for every pair of states along the x- and y-directions computed as described above, we approximate the energy of the system with an arbitrary configuration of domains by a two-dimensional anisotropic lattice Hamiltonian on a square lattice:

$$H = \sum_{i,j} E\left(\sigma\left(i,j\right)\right) + \sum_{i,j} J_x\left(\sigma\left(i,j\right), \sigma\left(i+1,j\right)\right) + \sum_{i,j} J_y\left(\sigma\left(i,j\right), \sigma\left(i,j+1\right)\right), \quad (5.1)$$

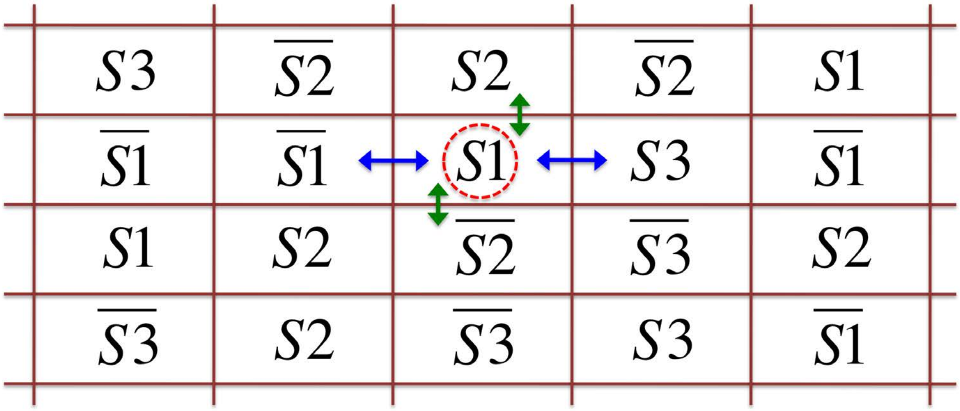


Figure 5.12: An example configuration of the two dimensional lattice that approximates the ZrO_2 monolayer on Si as a multi-domain system. Nearest neighbor sites couple through the coefficients J_x (blue arrows) and J_y (green arrows).

where $\sigma(i)$ donates the state at a given site i, $E(\sigma(i))$ is the energy (per 2×1 unit cell) of state $\sigma(i)$, and $J_\alpha(\sigma(i), \sigma(j))$ is the energy of interaction (i.e., domain wall energy) between the neighboring states i, j in the axial direction α. In our model, only nearest neighbor interactions are included. Because of the anisotropic nature of the film (the x and y directions are fundamentally different due to Si dimerization), the interaction term must distinguish between directions x and y so that J_x and J_y differ. The domain boundary energies calculated via DFT simulations are employed as nearest neighbor interaction energies in this model. In Figure 5.12, we illustrate an arbitrary configuration of such a lattice. As an example, the state $S1$ in the middle column couples to $\overline{S1}$ and $S3$ via $J_x\left(S1, \overline{S1}\right)$ and $J_x(S1, S3)$, respectively, and to $S2$ and $\overline{S2}$ via $J_y(S1, S2)$ and $J_y\left(S1, \overline{S2}\right)$, respectively.

For a model with N kinds of states, our interaction matrices J_α $(\alpha = x, y)$ have the following properties:

- The interaction energy between the sites of the same kind is zero by definition, $J_\alpha(\sigma_i, \sigma_i) = 0$. Hence the number of non-zero entries is $N^2 - N$.

- We have assumed that the domain wall energy between states σ_i and σ_j remains the same if we swap the states. Therefore the matrices are symmetric $J_\alpha(\sigma_i, \sigma_j) = J_\alpha(\sigma_j, \sigma_i)$, reducing the number of unique non-zero entries to $\frac{1}{2}(N^2 - N)$.

- In our particular system, every state has a counterpart which is obtained by the reflection $x \to -x$. Hence, e.g., the domain wall between $\overline{S1}$ and $\overline{S2}$ can be obtained from the domain wall between $S1$ and $S2$ by applying a single symmetry operation. Therefore many of the entires of $J_\alpha(\sigma_i, \sigma_j)$ are paired up in this way which further reduces the number of unique entries further to $\frac{1}{4}N^2$.

In Table 5.4 we list the unique entries of $J_\alpha(\sigma_i, \sigma_j)$ for states σ ranging over the the six lowest energy states. Note that since $N = 6$ for this table, there are $\frac{1}{4}6^2 = 9$ entries in the table. Because the unit cell is 2×1, the couplings J_x are expected to be smaller than the couplings J_y, which is generally correct. We have computed the domain wall energies for more possibilities of states including $S4$, $\overline{S4}$, $S5$ and $\overline{S5}$, and the resulting domain wall energies (listed in Table 5.6) are included in our treatment of the lattice model below.

We notice that some of the values in Table 5.4, namely $J_x\left(S2, \overline{S3}\right)$ and $J_y\left(S3, \overline{S3}\right)$, are very small, which we expect to be significant in finite temperature behavior of our model. We demonstrate the domain wall that corresponds to $J_y\left(S3, \overline{S3}\right)$ in Figure 5.13 with a top view. Because one of the …-Zr-O-Zr-O-… chains along the $y-$direction in the $S3$ unit cell is approximately aligned with the valley between consecutive Si dimers along the $x-$direction, it is approximately unchanged under the $S3 \to \overline{S3}$ transformation. Therefore when $S3$ and $\overline{S3}$ cells are attached in the $y-$direction, continuous and linear …-Zr-O-Zr-O-… chains are obtained (the top and the bottom black horizontal straight lines in Figure 5.13). The remaining …-Zr-O-Zr-O-… chain in the unit cell matches imperfectly, but the distortion is small (the winding black horizontal curve in the middle in Figure 5.13) such that the only atom

Domain boundary	J_x (eV)	J_y (eV)
$S1, \overline{S1}$	0.26	1.35
$S1, S2$	0.76	1.13
$S1, \overline{S2}$	0.96	0.99
$S1, S3$	0.61	4.81
$S1, \overline{S3}$	0.44	1.75
$S2, \overline{S2}$	0.38	1.64
$S2, S3$	0.17	0.98
$S2, \overline{S3}$	0.01	0.91
$S3, \overline{S3}$	0.73	0.002

Table 5.4: Domain boundary energies computed from first principles. These energies, along with the couplings that include the states $S4$, $\overline{S4}$, $S5$ and $\overline{S5}$ reported in Table 5.6, serve as the couplings of nearest neighbors in our lattice model.

with a modified environment is one of the oxygen atoms at the domain boundary (encircled with a red dashed ring in the figure). This near-perfect meshing of the ...-Zr-O-Zr-O-... chains after stacking $S3$ and $\overline{S3}$ structures along the $y-$direction may be said to cause the very small energy cost of making the domain boundary.

The model we have built is a general discrete lattice model that resembles two dimensional Ising models and, more generally, Potts models [103]. However, due to the lack of any simple pattern in site energies and couplings, it does not belong to any analytically solvable category of models.

5.4.3.3 Mean-field approach

To understand the thermodynamic behavior of this model at finite temperature, we begin with the standard mean-field approach which is based on the assumption that every site interacts in an averaged manner with its neighboring sites. For a model with N states $\sigma_1, \sigma_2, \ldots \sigma_N$, every site has a probability $p(\sigma_i)$ of being occupied by

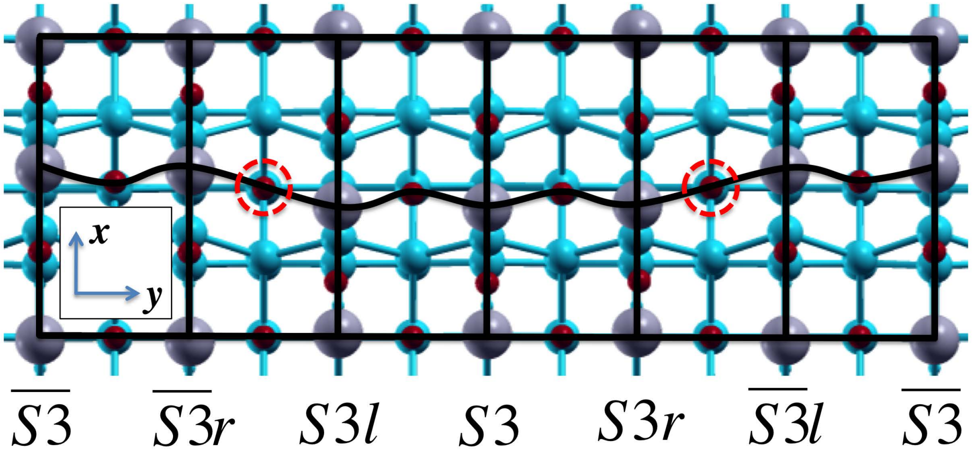


Figure 5.13: Top view of the domain boundaries along the x-direction between $S3$ and $\overline{S3}$, computed by stacking 3 unit cells of each structure along the $y-$direction. The domain energy, computed to be $J_y\left(S3,\overline{S3}\right) = 0.002$ eV per unit length, is very small due to the near-perfect meshing of the …-Zr-O-Zr-O-… chains in this configuration.

state σ_i. In mean field theory, the energy of such a site including its interactions with its nearest neighbors is

$$U\left(\sigma_i\right) = E\left(\sigma_i\right) + 2\sum_{j=1}^{N} p\left(\sigma_j\right) J_x\left(\sigma_i, \sigma_j\right) + 2\sum_{j=1}^{N} p\left(\sigma_j\right) J_y\left(\sigma_i, \sigma_j\right). \tag{5.2}$$

The probability $p(\sigma_i)$ is given by the the Boltzmann factor so that

$$p\left(\sigma_i\right) = \frac{\exp\left(-\frac{U(\sigma_i)}{k_\mathrm{B}T}\right)}{Z}, \tag{5.3}$$

where

$$Z = \sum_{j=1}^{N} \exp\left(-\frac{U\left(\sigma_j\right)}{k_\mathrm{B}T}\right) \tag{5.4}$$

is the mean-field partition function.

These equations form a self-consistent system of N equations for $p\left(\sigma_i\right)$ given T and the list of energies $E(\sigma_i)$ and couplings J_x, J_y. We present the solutions of the system

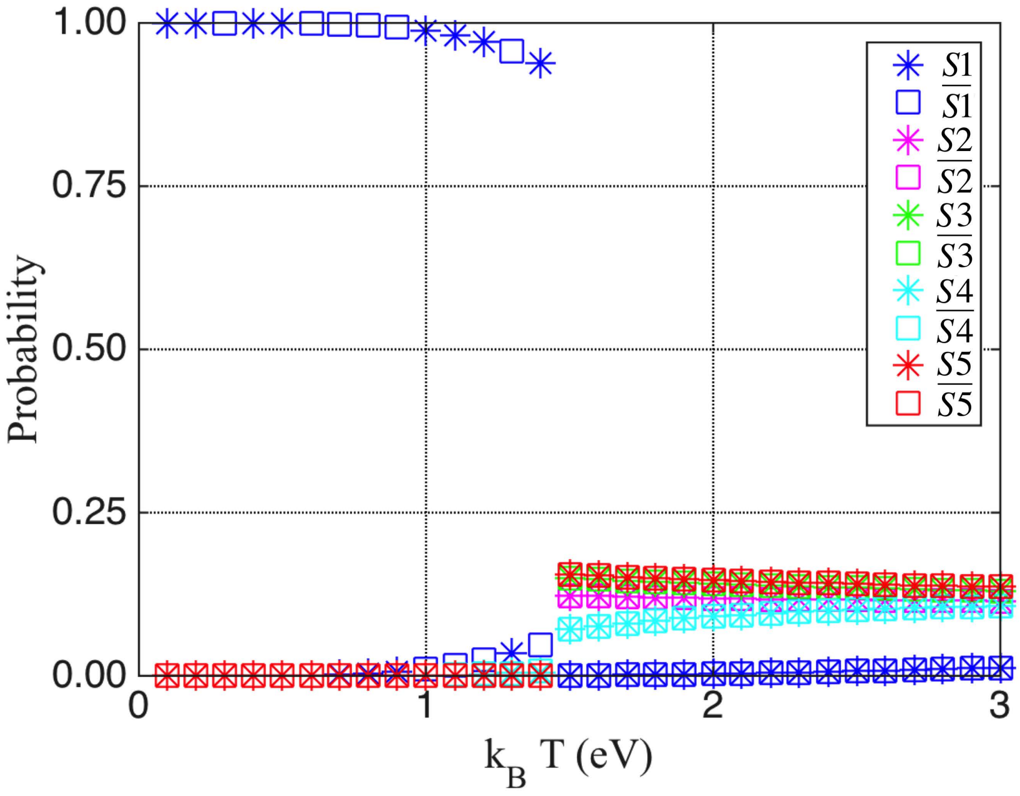


Figure 5.14: Probabilities of finding a type of state at an arbitrary site vs temperature, as computed by the mean-field equations for our lattice model.

of equations for temperatures ranging from 0.1 through 3.0 eV/k_B in Figure 5.14. We find that there is a first-order phase transition at $k_{\mathrm{B}}T = 1.4$ eV (i.e., about 16,000 K). Below this temperature, either one of the ground states $S1$, $\overline{S1}$ occupies close to all of the sites. In this regime, there is a spontaneous symmetry breaking which causes one of the ground states to fill the lattice. Above the transition temperature, the ground states are suppressed and the lattice gets filled by the rest of the states with an approximately equal distribution. At very high temperature (not shown in the figure), all states have equal probability, as expected.

It is known that in simpler two dimensional lattice problems, the mean-field approximation predicts the phase transition correctly but overestimates the critical temperature [104]. Overall, the mean-field approach assumes that each site interacts with its neighbors in an uncorrelated fashion equally, neglecting the fact that correlation lengths are finite. Moreover, as seen from equation (5.2), the mean-field equations are isotropic, which is an important shortcoming due to the major role anisotropy is expected to play in our system (see 5.4). In summary, we expect these mean field theory predictions to be informative but not quantitatively accurate.

5.4.3.4 Monte Carlo simulations

For a better understanding of our model at temperatures that are of interest to us, we have employed classical Monte Carlo simulations with a version of the Wolff cluster algorithm that we have developed. For further details of the method, please refer to Chapter 2. We have run simulations in a 50×150 lattice with free boundary conditions and a randomized initial condition, for $k_\mathrm{B}T = 0.016$, 0.032, 0.064, 0.128, 0.256 and 0.512 eV. We have used non-square cells because of the larger couplings in the y-direction compared to the x-direction, which lead to larger correlation lengths in the y-direction, as described below. In Figure 5.15, a sample configuration of a simulation with $k_\mathrm{B}T = 0.016$ eV ($T = 186$ K) is shown.

In Figure 5.16, the time autocorrelation function $C_\mathrm{auto}(t)$, and the horizontal and vertical spatial correlation functions $C_x(\Delta i)$ and $C_y(\Delta j)$ are plotted for a particular Monte Carlo run. These correlation functions are defined as:

$$C_\mathrm{auto}(\Delta t) = \underset{i,j,t}{\mathrm{mean}}\left[\langle\sigma(i,j,t)\,\sigma(i,j,t+\Delta t)\rangle - \langle\sigma(i,j,t)\rangle\langle\sigma(i,j,t+\Delta t)\rangle\right], \quad (5.5)$$

$$C_x(\Delta i) = \underset{i,j,t}{\mathrm{mean}}\left[\langle\sigma(i,j,t)\,\sigma(i+\Delta i,j,t)\rangle - \langle\sigma(i,j,t)\rangle\langle\sigma(i+\Delta i,j,t)\rangle\right], \quad (5.6)$$

$$C_y(\Delta j) = \underset{i,j,t}{\mathrm{mean}}\left[\langle\sigma(i,j,t)\,\sigma(i,j+\Delta j,t)\rangle - \langle\sigma(i,j,t)\rangle\langle\sigma(i,j+\Delta j,t)\rangle\right], \quad (5.7)$$

where $\sigma(i,j,t)$ denotes the state at the lattice site (i,j) at simulation time step t. In Figure 5.15, correlation functions for every type of state ($S1$, $\overline{S1}$ etc.) are computed

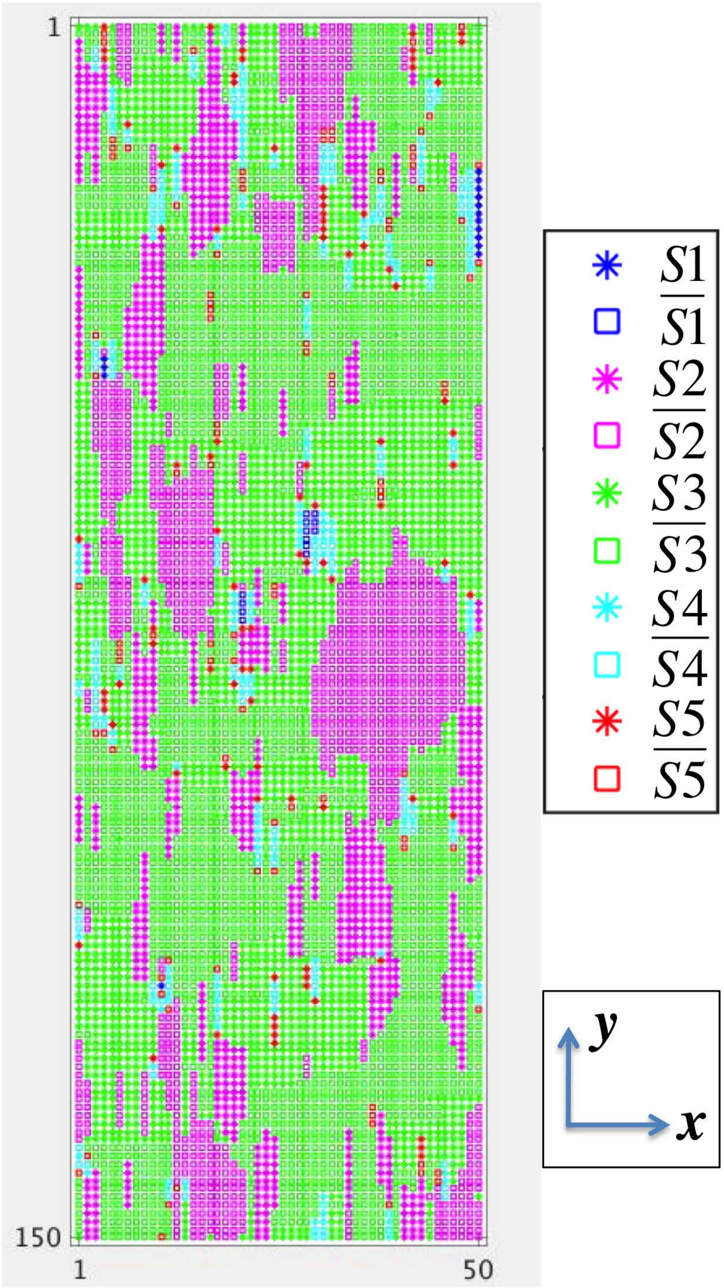


Figure 5.15: A sample of the Monte Carlo simulation of the lattice model at $k_{\mathrm{B}}T = 0.016$ eV ($T = 186$ K).

separately and overlaid. This is done by defining 10 functions $\sigma(i, j, t)$ (one for each state) such that for a given type, $\sigma(i, j, t)$ is 1 if the site i, j is occupied by that type at time t, and 0 otherwise.

We observe that for the run exemplified by Figure 5.15 and analyzed in Figure 5.15, (1) a 1000 step Monte Carlo simulation leads to decorrelation of states $S1$, $\overline{S1}$, $S4$, $\overline{S4}$, $S5$ and $\overline{S5}$ (i.e., equilibration), but not for $S2$, $\overline{S2}$, $S3$ and $\overline{S3}$. (2) The simulation cell of size 50×150 is successful in containing the domains that form at this temperature since the spatial autocorrelations become quite small by the half-way point in each direction: sites that are sufficiently far from each other are not correlated in the simulation. We have repeated these simulations 10 times for each temperature and have found that the correlation functions behave similarly when the initial state of the simulation cell is randomly chosen. For temperatures higher than 0.128 eV, all temporal correlations decay below 0.1 in the duration of the simulation.

The reason behind the slow temporal decay of the $S2$, $\overline{S2}$, $S3$ and $\overline{S3}$ autorcorrelations at low temperatures is that sufficiently large domains of these states form and the Monte Carlo algorithm becomes inefficient in flipping them to other configurations. To see what other effects are present in these simulations, we monitor two other physical quantities throughout the course of the simulations as displayed in Figure 5.17. The first quantity is the probability that any lattice site has a particular state (the ratio of the number of sites occupied by a state to the total number of sites in the simulation cell). The second quantity is the average domain size for each state at a given time (computed by determining all the domains of that state, including domains with only one site, and dividing the total number of sites occupied by the state to the number of domains). A large jump in the second quantity during the simulation usually indicates merging of two domains. The fact that these quantities change quickly at the beginning of the simulation and more slowly toward the end of the simulation in Figure 5.17 is indicative that the characteristics seen in Figure 5.15

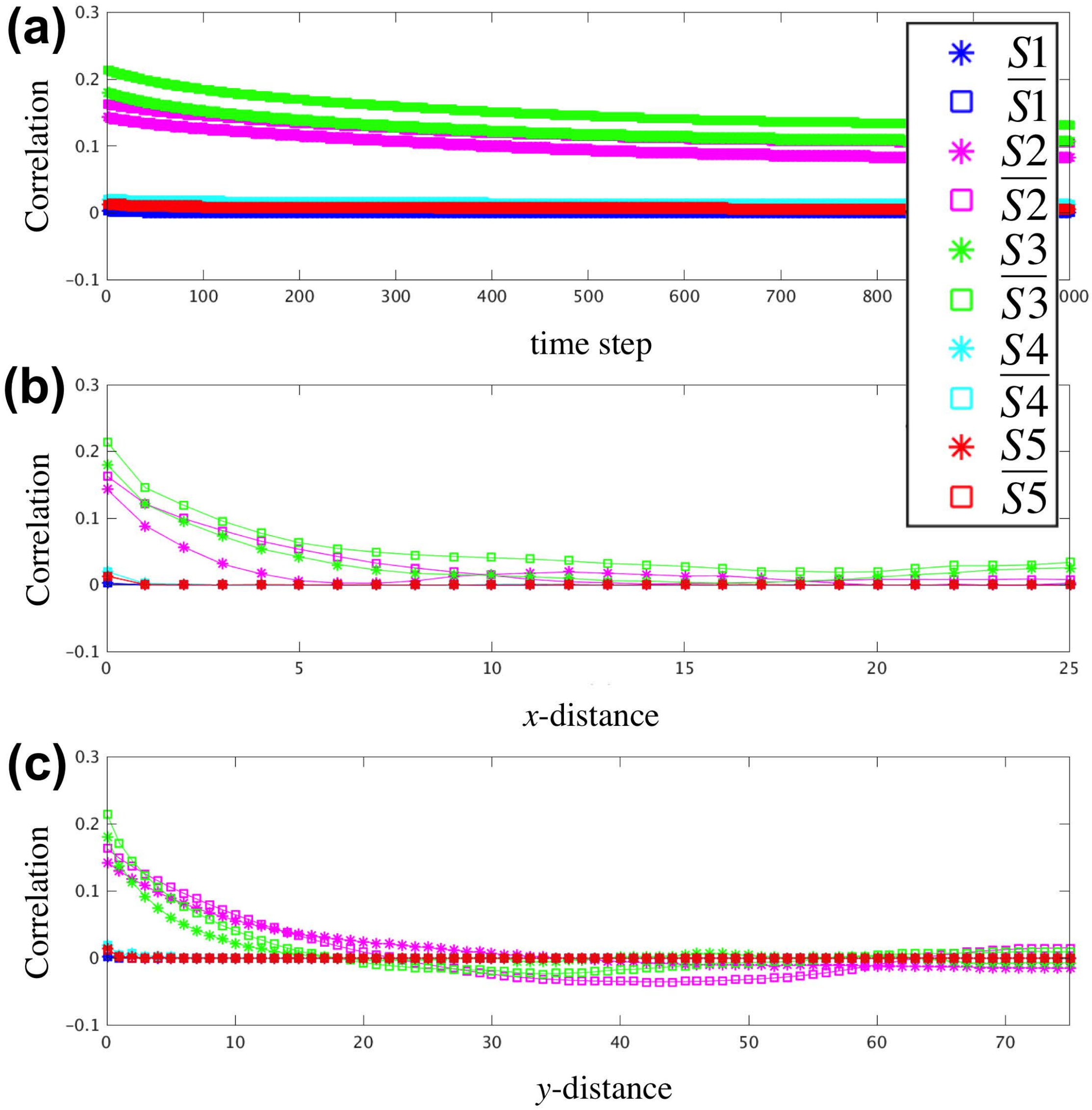


Figure 5.16: Temporal and spatial correlation functions of all 10 types of states for a Monte Carlo simulation for $k_\mathrm{B}T = 0.016$ eV. (a) Temporal correlation (autocorrelation) function defined in equation (5.5), (b) correlation function along the x-direction as per equation (5.6) and (c) correlation function along the y-direction as per equation (5.7).

are representative of large volumes of the configuration space sampled with the Boltzmann distribution at $k_{\mathrm{B}}T = 0.016$ eV (186 K): namely, while the lattice system has not fully equilibrated, it is not very far from equilibrium either. Please notice how domains $S2$ and $\overline{S2}$ are increasing in size at the expense of other states as the simulation goes on. Hence, despite its shortcomings, these results show that at this low temperature, the lattice system should be dominated by large domains of $S2$ and $\overline{S2}$ followed by smaller domains of $S3$ and $\overline{S3}$.

We now return to the mean field prediction that at temperatures lower than 1.4 eV the system should be dominated by either one of the ground states: this is not supported by the Monte Carlo simulations. Our Monte Carlo simulations show that for $k_BT \gtrsim 0.5$ eV, there is no long range order. The system gradually becomes more ordered as the temperature is lowered, but we have not yet found a sign of a sharp phase transition in our simulations. In 5.18, we plot the correlation lengths ξ_x and ξ_y along the x- and y-directions, respectively. The correlation lengths are calculated through the correlation functions $C_x(\Delta x)$ and $C_y(\Delta y)$ by fitting them to an exponential form $A\exp(-\Delta\alpha/\xi_\alpha)$. We calculate the correlation lengths (averaged over all states) for each run and then average over all runs at the same temperature. We find no indication of a second-order phase transition where one would expect the correlation lengths to diverge approaching to the critical temperature from both sides. Instead, the long range order increases gradually as the temperature decreases.

Finally, we comment on broad characteristics of the multi-domain structure of these films based on the lattice model. In Figure 5.19, we display the probability for a site to be occupied by each state as a function of temperature, where the probability values are averaged over the last quarter of each run, and then further averaged over 10 runs. Notice that the system is dominated by the second and the third lowest energy configurations $(S2, S3, \overline{S2}, \overline{S3})$. As discussed above, we believe that this is due to the rather low couplings $J_x\left(S2, \overline{S3}\right)$ and $J_y\left(S3, \overline{S3}\right)$ when compared to the

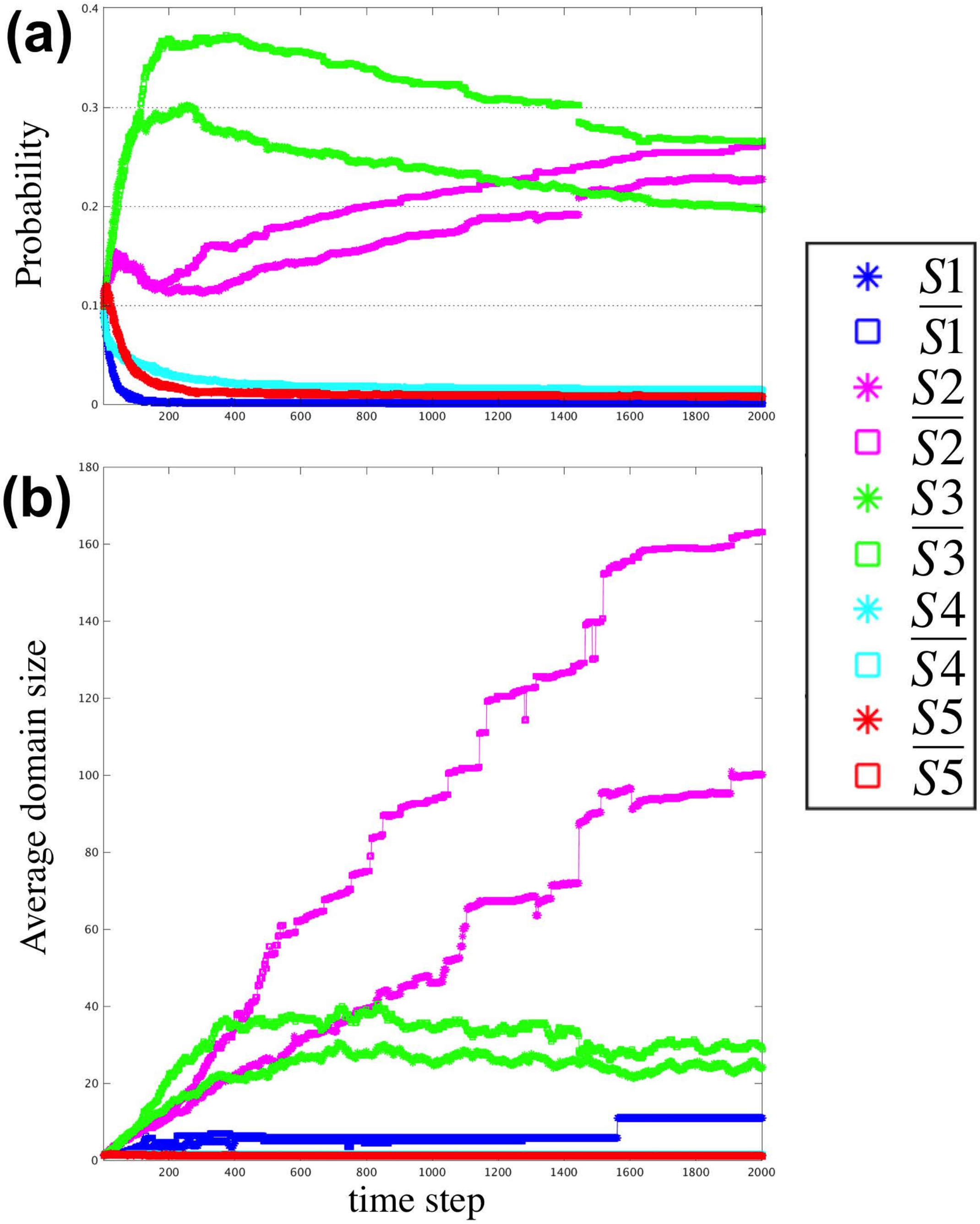


Figure 5.17: Probabilities of finding a state at an arbitrary site (a) and average domain sizes of each state (b), as they evolve during a Monte Carlo simulation for $k_{\mathrm{B}}T = 0.016$ eV.

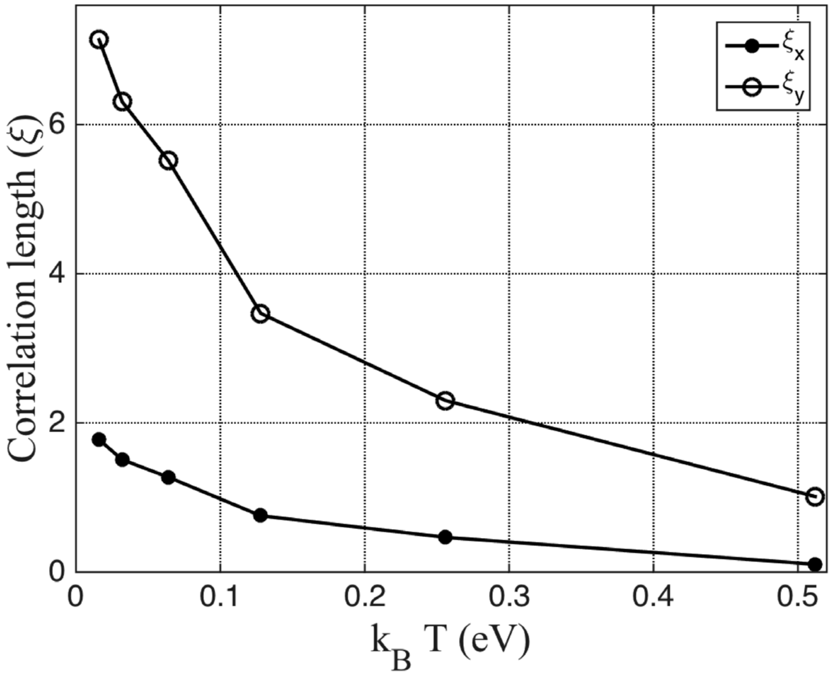


Figure 5.18: Correlation lengths along the x- and y-directions vs temperature. Each data point is obtained by fitting an exponential decay function to spatial correlation functions for each run at a given temperature, and then averaging the results of the fit for all the runs at that temperature.

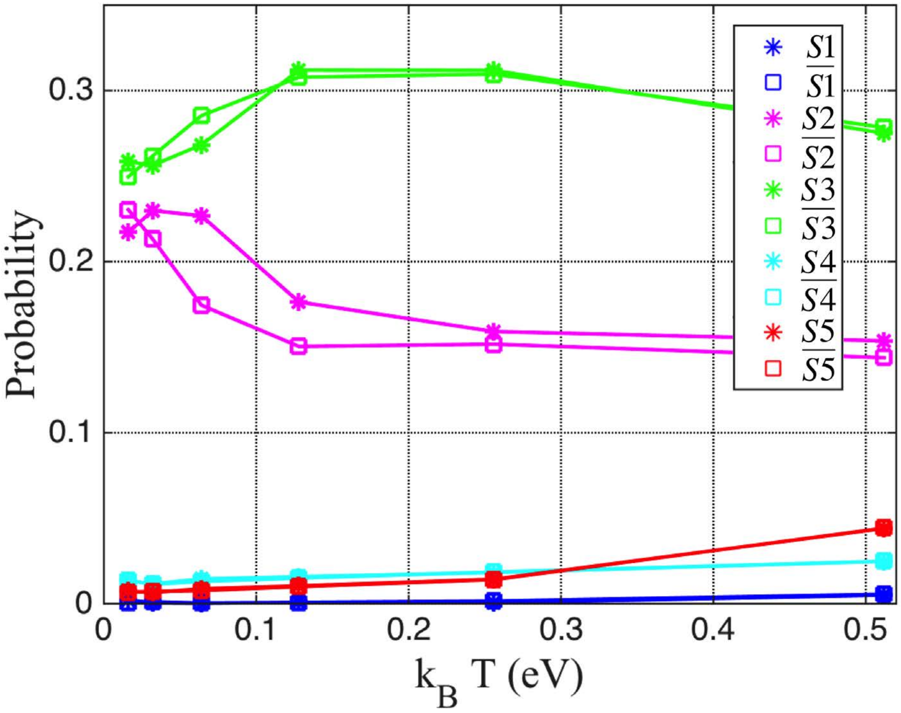


Figure 5.19: Probabilities of finding a state at an arbitrary site vs temperature, as computed by Monte Carlo simulations. For each temperature, the probabilities are averaged over the last quarter of each run, and then further averaged over 10 runs.

other couplings in Table 5.4. Namely, these domain walls are not very energetically costly, so their entropic contribution is significant even at low temperatures.

In Figure 5.20 we display the average domain size of each state vs temperature, again averaged over 10 runs for each temperature. We find that, on average, the domains of states $S2$ and $\overline{S2}$ are larger than the domains of states $S3$ and $\overline{S3}$, even though the latter occupy more of the simulation cell. This may be because $J_y\left(S3, \overline{S3}\right) = 0.002$ eV so the $S3$ and $\overline{S3}$ easily form vertical stacks of domains at essentially no energetic cost, as in Figure 5.15.

To sum up, according to our discrete lattice model, for 2×1 ordered ZrO_2 monolayers on the Si(001) surface, for the temperature range $200 - 800$ K, domains of

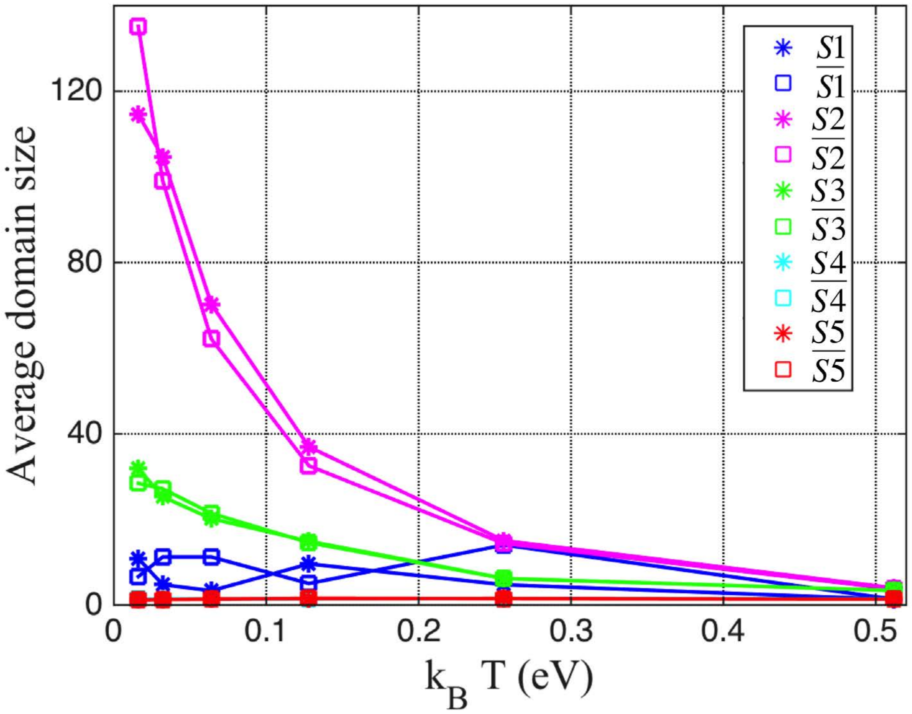


Figure 5.20: Average domain size for each type of state vs temperature, as computed by Monte Carlo simulations. For each temperature, the domain sizes are averaged over the last quarter of each run, and then further averaged over 10 runs.

$S2$, $\overline{S2}$, $S3$ and $\overline{S3}$ should be expected to occur with sizes of a few to a few dozen unit cells. This supports our claim that reaching epitaxy for these films should be challenging. However, we expect that locally the structure should be approximated by $S2$ and $S3$, and so the ferroelectric switching is understood largely as a transition between these two states. Time scales and dynamics of domain switching are subjects of potential further studies, as are much longer Monte Carlo simulations to fully explore the phase space and compute accurate thermodynamic quantities.

5.4.4 Oxygen content of the films

With the particular growth procedure described in Section §3.2, we expect the films to have an O:Zr ratio of $2:1$, with no oxygenation of the Si substrate. However, we know that the oxygen content of oxide thin films can be highly dependent on the growth conditions. In order to provide a comprehensive survey of these monolayers that may be experimentally realized through different growth methods, we have investigated ZrO_2 monolayers with variable amounts of oxygen. To this end, we have simulated 2×1 systems with O:Zr ratios of 1.0, 1.5, 2.5 and 3.0.

5.4.4.1 Low energy structures of ZrO_x films

To generate initial configurations, we have taken all the metastable configurations of the ZrO_2 monolayers, and either removed or added oxygens appropriately. For ZrO_x where $x = 1.0\,(1.5)$, we have removed 2 (1) O per cell, which has yielded 6 (4) initial configurations for each metastable configuration of ZrO_2. For ZrO_x where $x = 2.5$ or 3.0, we have added oxygen atoms to one or two of the following four positions: between surface Si atoms along the $x-$direction (midpoint of a dimer and halfway between successive dimers for non-1×1 systems), and between a surface Si atom and its neighbor in the subsurface layer (two inequivalent positions for non-1×1 systems). These choices are suggested by previous studies of O adsorption to

Structure	E (eV)	Structure	E (eV)
$ZrO_{1.0}$ $S1$	$\equiv 0.00$	$ZrO_{1.5}$ $S5$	0.88
$ZrO_{1.0}$ $S2$	0.02	$ZrO_{2.5}$ $S1$	$\equiv 0.00$
$ZrO_{1.0}$ $S3$	0.18	$ZrO_{2.5}$ $S2$	0.02
$ZrO_{1.0}$ $S4$	0.19	$ZrO_{2.5}$ $S3$	0.32
$ZrO_{1.0}$ $S5$	0.49	$ZrO_{2.5}$ $S4$	0.33
$ZrO_{1.0}$ $S6$	0.83	$ZrO_{2.5}$ $S5$	0.50
$ZrO_{1.5}$ $S1$	$\equiv 0.00$	$ZrO_{2.5}$ $S6$	0.90
$ZrO_{1.5}$ $S2$	0.12	$ZrO_{3.0}$ $S1$	$\equiv 0.00$
$ZrO_{1.5}$ $S3$	0.45	$ZrO_{3.0}$ $S2$	0.00
$ZrO_{1.5}$ $S4$	0.63	$ZrO_{3.0}$ $S3$	0.29

Table 5.5: Energies of ZrO_x monolayers on Si(001), where $x = 1.0, 1.5, 2.5, 3.0$ (reported in eV per 2×1 cell).

the bare Si(001) surface [105, 106] (and our confirmation of those works). We find that the most favorable position for an O atom on the Si(001) surface is the midpoint of a dimer, and the next most favorable position (at 0.23 eV per O) is between a surface atom and its subsurface neighbor (known as a back bond). Therefore, we have four positions to add an oxygen to a 2×1 Si/ZrO_2 interface, and hence 4 (6) initial configurations of $ZrO_{2.5(3.0)}$ for each metastable configuration of ZrO_2.

In Table 5.5 we list the energies of the metastable configurations for ZrO_x for $x = 1.0, 1.5, 2.5$ and 3.0. We have restricted the table to structures that are 1 eV or less per 2×1 cell higher than the ground state for all x. We notice that the multiplicity of structures at low energies is a feature of this system independent of the oxygen content. For $x = 1.0, 2.5$ and 3.0, the second lowest energy structure is within 0.02 eV of the ground state. See Figures 5.21, 5.22, 5.23, 5.24 and 5.25 for illustrations of these listed structures.

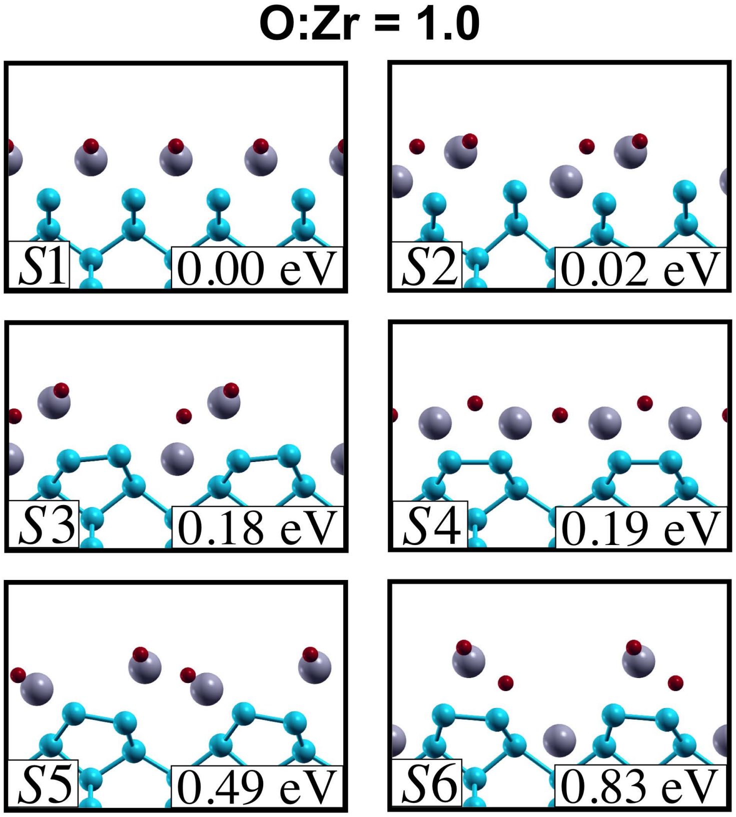


Figure 5.21: Low energy relaxed configurations of monolayers with O:Zr = 1.0. Reported energies are per 2 × 1 in-plane cell.

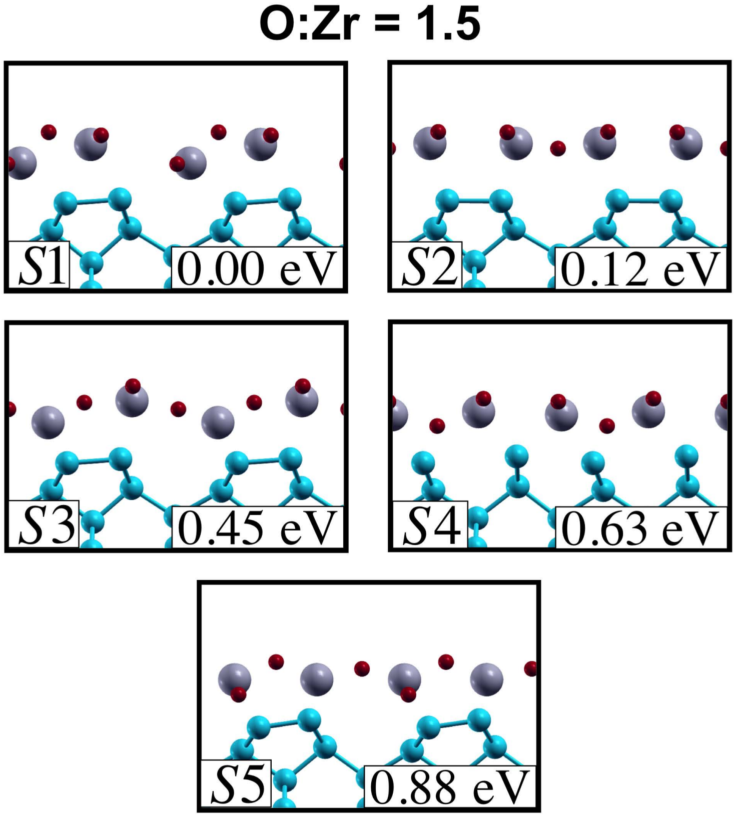


Figure 5.22: Low energy relaxed configurations of monolayers with O:Zr = 1.5. Reported energies are per 2 × 1 in-plane cell.

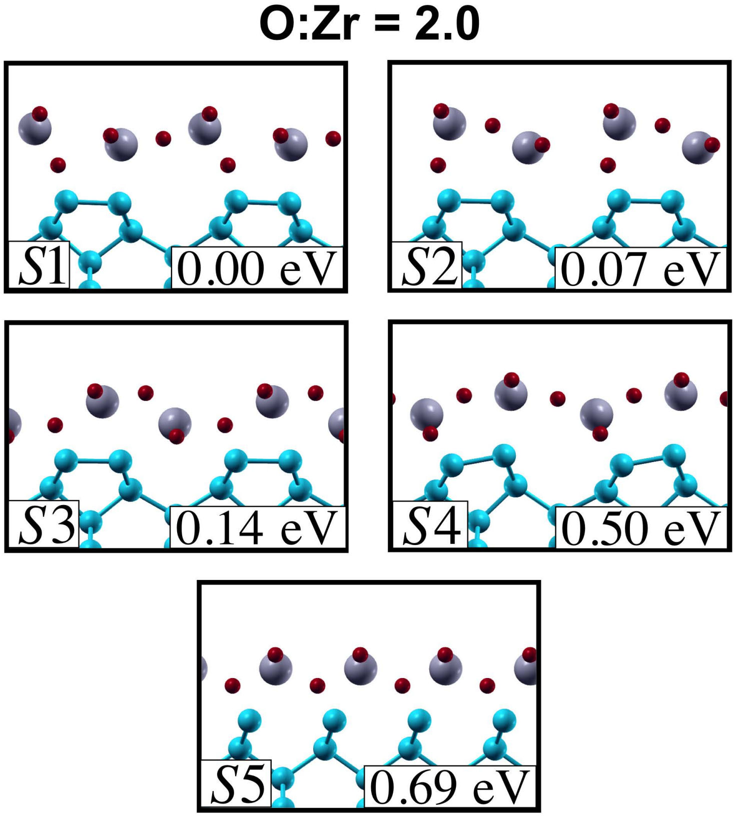


Figure 5.23: Low energy relaxed configurations of monolayers with O:Zr = 2.0. Reported energies are per 2 × 1 in-plane cell.

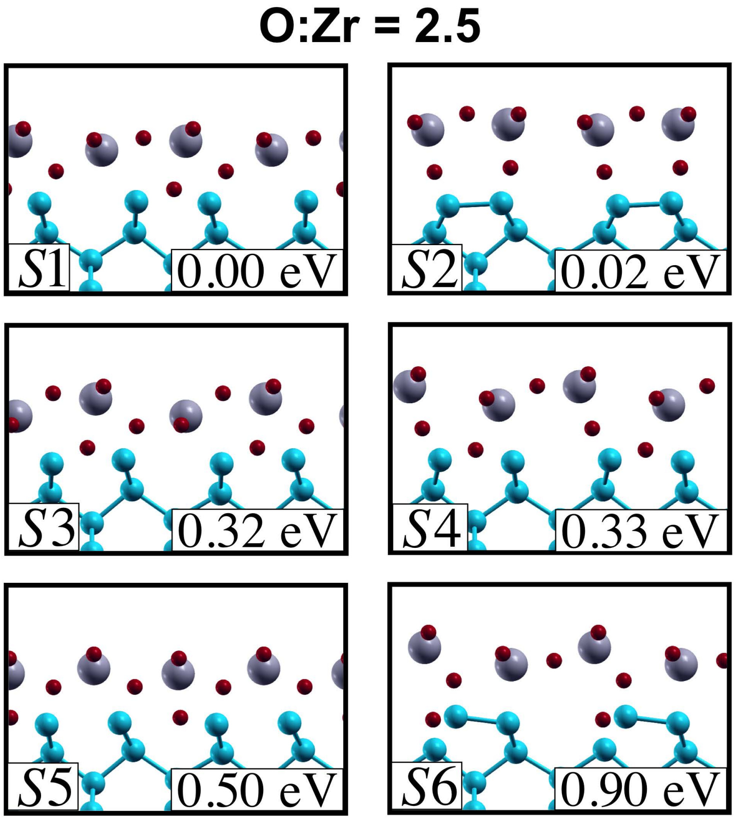


Figure 5.24: Low energy relaxed configurations of monolayers with O:Zr = 2.5. Reported energies are per 2 × 1 in-plane cell.

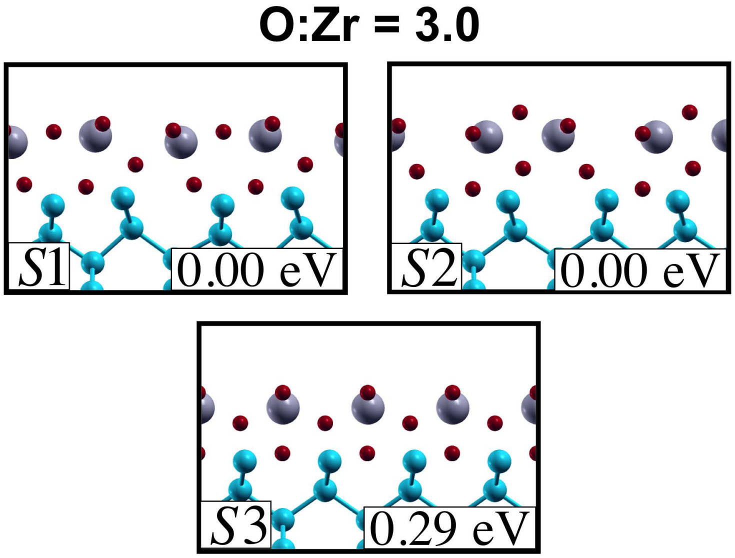


Figure 5.25: Low energy relaxed configurations of monolayers with O:Zr = 3.0. Reported energies are per 2 × 1 in-plane cell.

5.4.4.2 Polarization and Zr-O coordination

Given the large number of structures, we are only able to describe overall statistical trends for this dataset. Generally, we find that for the under-oxygenated cases ($x = 1.0, 1.5$), structures with lower energy have fewer Si-O bonds, and on average the oxygens are farther from the Si surface than the Zr: this represents the fact that O prefers to bond to Zr over Si. Separately, for low oxygen content, we find more Si-Zr bonds forming: this happens because the Zr are not fully oxidized and prefer to donate electrons to the more electronegative Si. On the other hand, in the over-oxygenated cases ($x = 2.5, 3.0$), we find that the extra oxygens bind to the available sites on the Si surface and the remaining oxygen distribute themselves among the Zr atoms such that the coordination of the Zr by O is maximized.

The above descriptions are supported by the data shown in Figure 5.26. On the left hand size, we plot energy E vs δz for all of the structures in our library. We find that negative ionic polarization is preferred for $x = 1.0, 1.5$, and that the energy increases as the polarization increases. Hence, the O anions prefer to be farther from the Si surface, and the Zr cations prefer to be closer. We find the opposite for $x = 2.5, 3.0$, where positive polarization is preferred, and the lowest energy structures have the highest Zr-O out-of-plane separation. On the right hand side, we plot E vs the coordination number of Zr by O. We calculate the coordination number for a Zr atom by counting the number of O atoms within 3 Å. We then average the coordination numbers of the two inequivalent Zr atoms, and report it in the figure. We find that for all O:Zr ratios, higher C.N. tends to lower the energy, with the exception of O:Zr = 2, where one special low-energy configuration ($S2$ as shown in Figure 5.7) has C.N. = 3.5 while the others have C.N. = 4. Note that for these plots, we count symmetry related structures such as $S1$ and $\overline{S1}$ in $ZrO_{2.0}$ as a single structure.

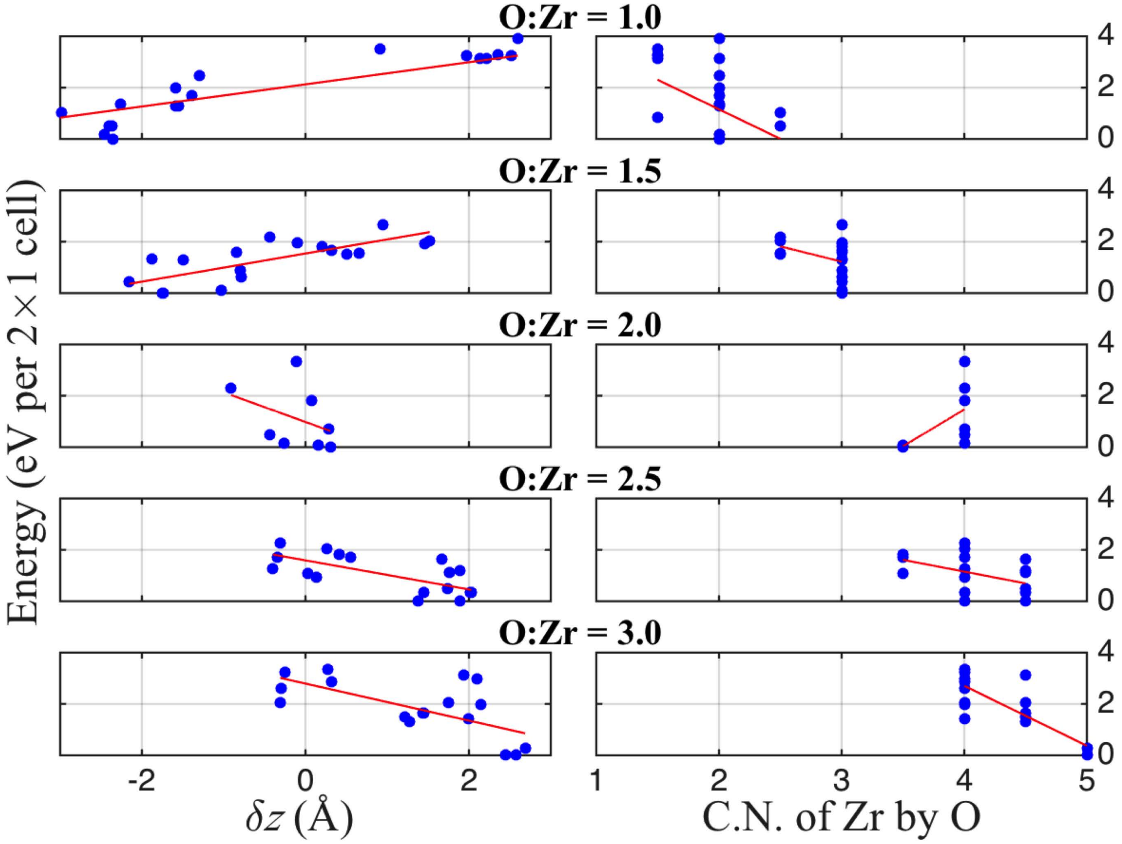


Figure 5.26: E vs δz and E vs the coordination number (C.N.) of Zr by O scatter plots for the relaxed structures obtained in this study, for ZrO_x monolayers on Si(001), where $x = 1.0, 1.5, 2.0, 2.5, 3.0$. Linear fits computed by the method of least squares are also displayed. $\delta z \equiv \overline{z\,(\mathrm{Zr})} - \overline{z\,(\mathrm{O})}$ measures the ionic out-of-plane polarization. C. N. of Zr by O is defined as the number of oxygen atoms that are within 3 Å of a zirconium atom. For each structure, the C.N. is calculated separately for the two generally inequivalent Zr atoms, and then averaged. Energies are listed in eV per 2×1 in-plane cell measured with respect to the lowest energy structure for each O:Zr ratio.

5.5 Conclusion

We have conducted a computational study of ZrO_2 monolayers on Si(001) using DFT. We have found a multiplicity of (meta)stable structures with a large variation in ionic polarization but no large differences in energy, atomic structure and chemistry. This suggests that achieving epitaxy in the experiment should be challenging. We have also summarized the experimental results on amorphous ZrO_2 monolayers on Si(001) with an atomically abrupt interface. These monolayers are shown with C-V measurements to be ferroelectric with $\Delta V = 0.4$ eV. In order to understand the finite-temperature behavior of these ultrathin films, we have developed a two dimensional Ising-like discrete lattice model of the domains in these thin films with DFT-derived parameters. We have employed mean-field and Monte Carlo simulation approaches to study this lattice model and concluded that two oppositely polarized structures, namely $S2$, $S3$ and their counterparts $\overline{S2}$ and $\overline{S3}$, dominate the films at temperatures of interest. Ferroelectric switching observed in the experiment can be understood as locally switching between these two structures. We have found that for epitaxial films this switching would lead to a VBE shift in silicon of $\Delta V = 0.6$ eV, which is comparable to the experimental value. Finally we have investigated the monolayers with varying oxygen content, i.e. ZrO_x with $x = 1.0, 1.5, 2.0, 2.5, 3.0$. We have reported the main trends in these films that may guide future theoretical and experimental research into monolayer oxides.

Appendix: List of all domain wall energies

We list all the domain wall energies in Table 5.6, which includes the couplings between $S1$, $S2$, $S3$ and their barred counterparts, also reported above in Table 5.4.

A few of the couplings that involve the higher-energy $S4$ and $S5$ structures are negative, which can be understood in some cases as the domain boundary region re-

Domain boundary	J_x (eV)	J_y (eV)
$S1, \overline{S1}$	0.26	1.35
$S1, S2$	0.76	1.13
$S1, \overline{S2}$	0.96	0.99
$S1, S3$	0.61	4.81
$S1, \overline{S3}$	0.44	1.75
$S1, S4$	0.55	2.79
$S1, \overline{S4}$	-0.20	2.37
$S1, S5$	0.35	1.35
$S1, \overline{S5}$	0.40	0.56
$S2, \overline{S2}$	0.38	1.64
$S2, S3$	0.17	0.98
$S2, \overline{S3}$	0.01	0.91
$S2, S4$	-0.17	2.23
$S2, \overline{S4}$	0.86	2.28
$S2, S5$	0.34	0.59
$S2, \overline{S5}$	-0.12	1.31
$S3, \overline{S3}$	0.73	0.002
$S3, S4$	0.10	1.89
$S3, \overline{S4}$	0.23	1.84
$S3, S5$	-0.24	0.75
$S3, \overline{S5}$	0.69	1.21
$S4, \overline{S4}$	0.55	-0.33
$S4, S5$	0.71	0.65
$S4, \overline{S5}$	-0.26	1.86
$S5, \overline{S5}$	0.56	0.30

Table 5.6: Domain boundary energies computed from first principles. These energies serve as the couplings of nearest neighbors in our lattice model.

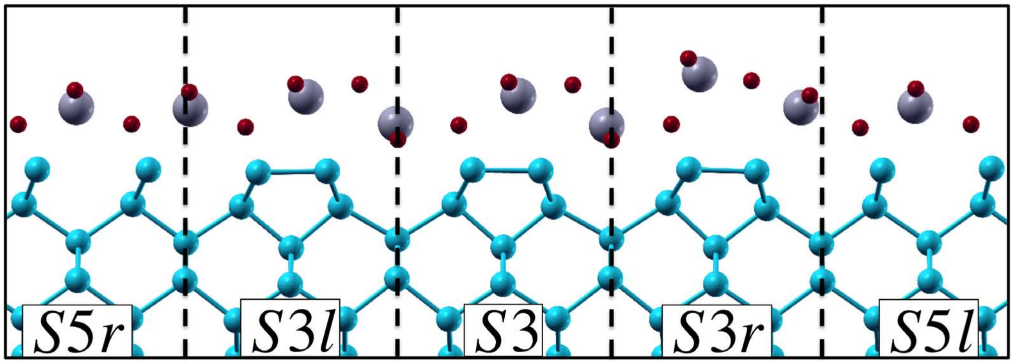


Figure 5.27: The domain boundaries along the y-direction between $S3$ and $S5$, computed by stacking 3 unit cells of each structure along the x-direction. The domain energy, computed to be $J_x\,(S5, S5) = -0.24$ eV per unit length, is negative partly due to the fact that the vicinity of one of the boundary walls (labelled $S3r$) resembles a lower energy configuration $S2$ (see Figure 5.7).

sembling a lower energy structure. In Figure 5.27 we illustrate the domain boundaries that correspond to $J_x\,(S3, S5) = -0.24$ eV. The structure immediately to the left of the right domain boundary ($S3r$) closely resembles the $S2$ structure (see Figure 5.7). However, the fact that the higher energy $S4$ and $S5$ structures have negative domain wall energies with the lower energy structures in some cases is not enough to generate antiferroelectric patterns in our Monte Carlo simulations. This may be due to the separation of scale in the energies of the lowest three structures and $S4$ and $S5$ (see Table 5.2). Hence the energy gain by making these domain boundaries (0.26 eV or less) is not enough to compensate the high energy of these two structures.

Chapter 6

Ferroelectricity in HfO_2 thin films

6.1 Introduction

Achieving ferroelectricity in thin films is the subject of decades long research endeavor because of potential technological applications, e.g., the ferroelectric field effect transistors (FeFET) and ferroelectric memory (FeRAM). We have discussed thin film ferroelectricity on semiconductors in Chapters 4 and 5. In this Chapter, we turn to HfO_2 which is the widely used dielectric gate oxide in today's field effect transistor devices. HfO_2 is thermodynamically stable on silicon up to high temperatures, allowing abrupt HfO_2/Si interfaces to be grown without formation of silica in the interfacial region [107]. In addition, with a large band gap of 5.3 - 5.7 eV and a dielectric constant of $\varepsilon_r \approx 20$ in its bulk form under ordinary conditions [108], HfO_2 is a widely used gate insulator and a replacement material for SiO_2 [4]. The recent discovery of ferroelectricity in HfO_2-based thin films has further multiplied the research interest in this material [9, 10]. It has been shown in various experimental studies that ferroelectricity in HfO_2-based films arises from the creation of the polar orthorhombic phase (space group: $Pca2_1$) of HfO_2 during a rapid annealing process in conjunction with the presence of a capping electrode (typically TiN). It has also been demonstrated that the ferroelectric properties of these films strongly depend on factors such as the doping species, doping concentration, annealing temperature and film thickness [108]. Even though devices based on these films are already being designed and developed, a microscopic understanding of the effects of the aforementioned factors is presently lacking. This is, in part, due to the polycrystalline and complex nature of the HfO_2 films that have been grown to date, as well as due to the relative newness of this field.

To the best of our knowledge, the ferroelectric hafnia-based thin films are polycrystalline with differently oriented grains of monoclinic (space group: $P2_1/s$), tetragonal ($P4_2/nmc$) and orthorhombic ($Pca2_1$) phases in various ratios. The monoclinic and the tetragonal phases are non-polar, and they are the observed bulk phases of HfO_2 at room temperature and at high temperature, respectively. Experiments demonstrate that the orthorhombic phase arises during the rapid annealing with a capping electrode. In addition, the concentration of dopants is crucial in determining the ferroelectric properties. Because the volume fraction of the orthorhombic phase compared to the other non-polar phases is what ultimately decides the robustness of ferroelectricity in the HfO_2 films, a structural understanding of the favorable conditions for this phase is crucial in order to optimize the growth procedure.

To this end, in our *ab initio* studies we investigate the energetics of different phases of HfO_2 with varying amounts of Si and Zr doping and subject to a range of epitaxial strain states. In Section §6.2 we summarize the experimental findings on ferroelectric thin films of hafnia to date. In Section §6.4 we describe our computational study of doping and strain on hafnia with a particular focus on Si:HfO_2 and $Hf_xZr_{1-x}O_2$. We have found that at certain doping levels the transformation of the high temperature tetragonal phase to the out-of-plane polarized orthorhombic phase is favored. These results, together with additional analysis, help explain the common experimental observations as well as some of the underlying causes from a microscopic viewpoint. We also describe, in Section §6.4, preliminary results on simulated HfO_2 thin films including the interface with electrodes in order to find a comprehensive understanding of the energetics in thin films.

6.2 Summary of experiments to date

Since their discovery in 2011, ferroelectric hafnia thin films have garnered tremendous experimental attention. In Table 6.1 we list some of the studies that have investigated factors such as doping species, doping concentration, thickness of the film, and top and bottom electrodes (TE/BE).

These studies have found that when the HfO_2 films are doped with a few percent of a wide range of dopants, or with $\sim$50% Zr, and grown to be $\sim 5-20$ nm thick between two metal electrodes, and annealed at $\sim 800-1000$ °C, they can display ferroelectricity. Some of the studies [109, 9, 110, 114] have shown that for $TiN/HfO_2/TiN$ stacks, if the film is annealed before the deposition of the top electrode, then the ferroelectric hysteresis is significantly suppressed. Given that TaN and Ir top electrodes have also been found to be favorable for ferroelectricity, this effect is thought not to be due to the particular structure of TiN, but to be due to the confinement the top electrode provides to the film during annealing.

In terms of the atomic structure of the films, many of these studies have performed XRD analyses to show that ferroelectricity is intimately related with the presence of the orthorhombic $Pca2_1$ phase of HfO_2. The path to the stabilization of this phase is believed to be via the creation of grains of the tetragonal $P4_2/nmc$ phase which are present when the film is first deposited [10]. The tetragonal phase is known to be stabilized by surface effects [120], and recent studies have shown doping to be a stabilizer for this phase as well [121, 122]. Hence, it appears that for ferroelectric hafnia-based films, the orthorhombic phase is obtained from the tetragonal grains during rapid thermal annealing under the confinement of a top electrode. Crystallization without

Reference	Dopant	TE/BE	Thickness	Observations
Böscke [9]	Si	TiN/TiN	7-10 nm	FE at 2.6%-4.3% Si; ∼AFE at 5.6% Si.
Müller [109]	Y	TiN/TiN	10 nm	FE at 2.3%-5.2% Y.
Mueller [110]	Al	TiN/TiN	16 nm	FE at 4.8% Al; AFE at 8.5% Al.
Müller [10]	Zr	TiN/TiN	9 nm	PE at $<$30% Zr; FE at %30-%60 Zr; AFE at $>$70% Zr.
Yurchuk [111]	Si	TiN/TiN	9, 27 nm	FE at 4.4% Si, 9 nm; ∼PE at 27 nm.
Park [112]	50% Zr	TiN/TiN	5-25 nm	FE at 5-17 nm; PE at 25 nm. $Pca2_1$ phase confirmed.
Park [113]	50% Zr	TiN/TiN TiN/Pt	5-27 nm	FE at 8-19 nm for TiN; Less FE for Pt at 8 nm; PE for Pt at $>$ 13 nm.
Lomenzo [114]	Si	TiN/Si Ir/Si Ir/Ir	10 nm	FE similar for TiN/Si and Ir/Si. Smaller P_r for Ir/Ir.
Schroeder [115]	Si, Al, Y, Gd, La, Sr	TiN/TiN	10 nm	FE at 4.4% Si; AFE at $>$5.6% Si. Similar for Al. No AFE for other dopants.
Park [116]	50% Zr	TiN/TiN TiN/Ir	9-24 nm	FE at 9-19 nm for TiN; FE at 12-15 nm for Ir.
Sang [117]	Gd	TiN/TiN	27 nm	FE; $Pca2_1$ phase confirmed.
Hoffmann [118]	Gd	TiN/TiN TiN/TaN TaN/TaN	10-27 nm	FE similar for all stacks; TaN/TaN $>$ TiN/TiN $\simeq$ TiN/TaN in terms of P_r.
Chernikova [119]	50% Zr	TiN/TiN	2.5 nm	FE; $Pca2_1$ phase confirmed.

Table 6.1: Selected experimental studies of ferroelectric hafnia thin films published between 2011 and 2016, listed chronologically. FE means ferroelectric, AFE means antiferroelectric.

a top electrode leads to the formation of the monoclinic P2_1/s phase, which is the non-polar ground state at relevant temperatures. The causes for the favorability of the tetragonal → orthorhombic transition over the tetragonal → monoclinic transition under these doping/thickness/temperature/confinement conditions are not well understood.

6.3 Computational methods

We compute minimum energy structures using density functional theory (DFT) in the Perdew–Burke–Ernzerhof Generalized Gradient Approximation (PBE GGA) [18] with ultrasoft pseudopotentials [19]. We use the QUANTUM ESPRESSO software package [65]. Plane wave energy cutoffs of 55 Ry and 35 Ry are used for bulk and thin film simulations, respectively. We sample the Brillouin zone with an $8 \times 8 \times 8$ Monkhorst–Pack k-point mesh (per 12 atom unit cells) and a 0.02 Ry Marzari–Vanderbilt smearing [64] for bulk samples; slab (thin film) systems are sampled at a density of $6 \times 6 \times 1$. For thin film simulations, periodic copies of the slab are separated by ~ 12Å of vacuum in the $z-$direction (see Figure 6.1 for a representative unit cell), and the in-plane lattice constants of the slab are fixed to the bulk computed lattice constants of the HfO_2 for the plane in consideration. In general, a slab may have an overall dipole moment that can artificially interact with its periodic copies. We eliminate this effect by introducing a fictitious dipole in the vacuum region of the cell which cancels out the electric field in vacuum [66]. All atoms are relaxed until the forces on the atoms are less than 10^{-4} Ry/a_0 in all axial directions, where a_0 is the Bohr radius. We use both direct substitution of atoms as well as the virtual crystal approximation (VCA) to model doping [123]. In the VCA, to approximate the mixing of two elements A and B with ratios x and $1 - x$, a virtual element is created by linearly interpolating the pseudopotentials of A and B such that the

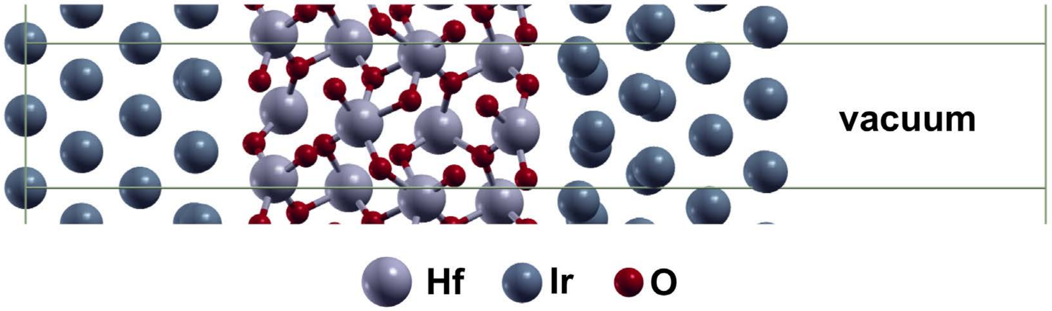


Figure 6.1: A sample simulation cell for a thin film relaxation of an Ir/HfO_2/Ir stack. HfO_2 is in the monoclinic phase with [001] orientation. The in-plane lattice is fixed to the lattice parameters of this phase and orientation of HfO_2 (see section 6.4.4 for details of these thin film simulations). Periodic copies of the stack are separated by vacuum.

resulting pseudopotential is $V = xV_A + (1-x)\,V_B$. The VCA is known to be a good approximation when alloying chemically similar elements with the same valence state. A detailed comparison of these two approaches for some test cases is reported in section 6.4.2.

6.4 Results

6.4.1 Bulk phases of HfO_2

HfO_2 can be observed in three phases in its bulk form. The monoclinic phase (space group: $P2_1/c$) is stable all the way up to ∼2000 K. Between ∼2000 K and ∼2900 K, the tetragonal phase (space group: $P4_2/nmc$) is observed. The highest symmetry cubic phase (space group: $Fm\overline{3}m$) is observed between ∼2900 K and the melting temperature of ∼3100 K [124]. The cubic phase of HfO_2 is depicted in Figure 6.2. It is a face centered cubic structure with one formula unit per lattice point. The tetragonal and the monoclinic phases are obtained from the cubic phase through consecutive symmetry breaking operations.

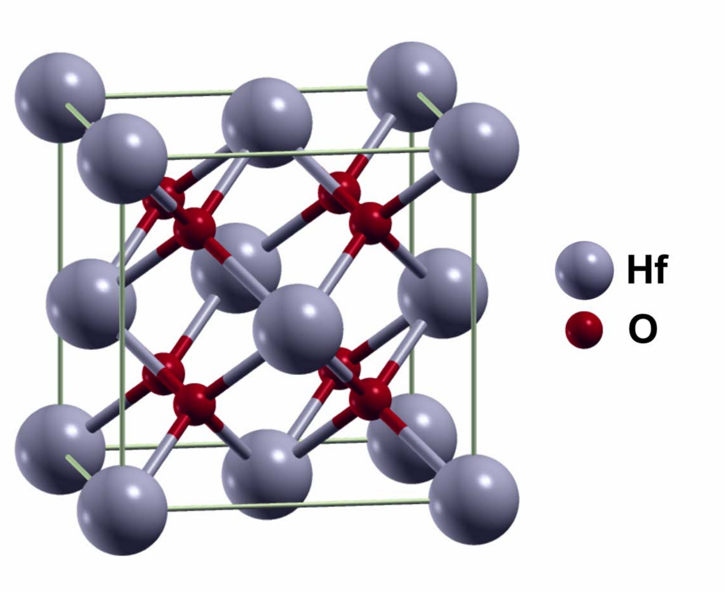


Figure 6.2: Cubic phase of HfO_2.

These three bulk phases are all centrosymmetric, which makes the oxide dielectric under ordinary conditions. However, as described in Section §6.2, the recent discovery of ferroelectricity in HfO_2 thin films indicates that a non-centrosymmetric orthorhombic phase (space group: $Pca2_1$) is stabilized under certain growth conditions that gives rise to a switchable polarization. The four phases of HfO_2 that are the focus of this study are shown in Figure 6.3. The orthorhombic and monoclinic phases are obtained from the tetragonal phase by symmetry breaking operations, indicated by a blue arrow in the figure.

In Table 6.2 we compare our computed energies of each phase (relative to the monoclinic phase) with the results from previous computational studies. We find that our results agree well with Ref. [125] which uses the GGA to approximate the exchange-correlation functional; and Refs. [126] and [127] which use the Local Density Approximation (LDA) agree with each other.

In Table 6.3 we report the lattice parameters we have found compared with the results of previous studies. We find that our results generally lie within the range of

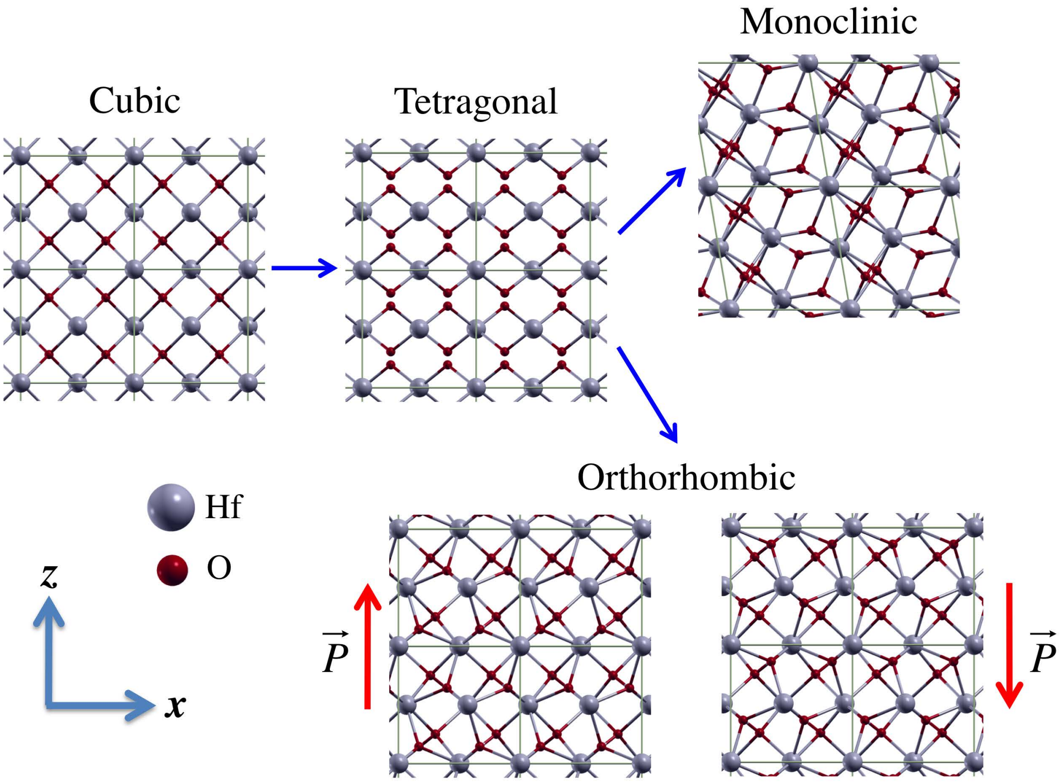


Figure 6.3: Bulk phases of HfO_2 considered in this work. Subgroup relations are shown by blue arrows. Cubic, tetragonal and monoclinic phases are experimentally observed bulk phases. The non-centrosymmetric orthorhombic phase is observed in some thin films of HfO_2 and gives rise to ferroelectricity. For each phase, the 12-atom simulation cell that we use in this study is also displayed by thin lines.

Bulk phase	This work (GGA)	[125] (GGA)	[126] (LDA)	[127] (LDA)
mono $P2_1/c$	$\equiv 0.00$	$\equiv 0.00$	$\equiv 0.00$	$\equiv 0.00$
ortho $Pca2_1$	0.08	0.06	0.05	0.06
tetra $P4_2/nmc$	0.16	0.14	0.07	0.09
cubic $Fm\overline{3}m$	0.27	0.21	0.09	0.14

Table 6.2: Energies of the bulk phases of HfO_2 (relative to the monoclinic phase) considered in this work, compared with previous computational work.

agreement among the previous work, where there is a difference of <5% for a given lattice parameter between studies.

6.4.2 Effects of doping on bulk HfO_2

Due to its importance in the ferroelectric properties of hafnia thin films, we have investigated the role of doping in stabilizing the various phases of HfO_2 with respect to each other. We list the energies of bulk phases with respect to the monoclinic phase for various dopants in Table 6.4. These simulations are done with $2 \times 2 \times 2$ simulation cells with 96 atoms, where one Hf per cell (1 in 32) is replaced by the dopant. All atomic positions and cell parameters are then relaxed. This leads to a doping ratio of 3.125% where the dopants are equally spaced in the three lattice directions. For elements with a different number of valence electrons than Hf, such as N, Al, Sr, Y and La, we have additionally computed relaxed energies with compensating (added) electrons or holes and compared these with the neutral relaxations.

We find that doping does not change the energy difference between the orthorhombic and the monoclinic phases significantly. However, in some cases (C, N*, Si and Ge) it reduces the energy of the tetragonal phase, and in some cases (C, N, N*, Al*, Si, Ge and Sr*) it increases the energy of the cubic phase. We also find that changing the number of electrons in the cell does not significantly modify the energies of the orthorhombic and the tetragonal phases (with the exception of N), but increases the energy of the cubic phase (again with the exception of N).

The most significant observation of this survey of dopants is that, for non-metal dopants (C, N*, Si and Ge), the tetragonal phase experiences a significant stabilization. If we focus on the monoclinic, orthorhombic and tetragonal phases, which are the phases that participate in the thin film processes, we conclude that, apart from the cases of C, N*, Si and Ge, no significant change occurs in terms of pure phase energetics.

Bulk phase	Parameters	This work (Å)	Ref. [125]	Ref. [128]	Ref. [127]
Monoclinic	$a, b, c,$ β	5.12, 5.18, 5.30, 99.6°	5.09, 5.16, 5.26, 99.7°	5.14, 5.20, 5.31, 99.8°	5.11, 5.18, 5.29, 99.7°
Orthorhombic	a, b, c	5.25, 5.04, 5.07	5.11, 4.90, 4.92	5.29, 5.01, 5.08	5.23, 5.04, 5.06
Tetragonal	a, c	5.07, 5.19	5.03, 5.15	5.06, 5.28	5.05, 5.14
Cubic	a	5.06	5.03	-	5.04

Table 6.3: Lattice parameters (in Å) of the bulk phases of HfO_2 compared with previous computational studies. For the monoclinic phase, β is the angle between $\vec{a}$ and $\vec{c}$ which is the only non-perpendicular angle for this phase.

Bulk phase	ortho $Pca2_1$	tetra $P4_2/nmc$	cubic $Fm\overline{3}m$
No doping	0.09	0.16	0.26
C	0.07	0.08	0.46
N	0.10	0.15	0.41
N* (-1e)	0.09	0.10	0.41
Al	0.09	0.16	0.28
Al* (+1e)	0.09	0.15	0.33
Si	0.07	0.11	0.33
Ti	0.09	0.15	0.27
Ge	0.09	0.12	0.30
Sr	0.08	0.16	0.23
Sr* (+2e)	0.10	0.18	0.34
Y	0.08	0.17	0.24
Y* (+1e)	0.09	0.17	0.29
La	0.08	0.16	0.24
La* (+1e)	0.08	0.16	0.29

Table 6.4: Energies of the bulk phases of HfO_2 (relative to the monoclinic $P2_1/c$ phase) with 3.125% doping with various elements. The elements with a star (*) denote simulations with added electrons or holes to compensate for the difference in the number of valence electrons of that element and Hf. The number of extra electrons per simulation cell is shown next to each element in parentheses. For the case of N, one electron is taken out of the system, which is denoted as (-1e).

Phase	Nearest O neighbor distances (Å)								C. N.
mono Hf	2.05	2.14	2.14	2.16	2.18	2.28	2.30		7
ortho Hf	2.04	2.12	2.13	2.14	2.14	2.24	2.27		7
tetra Hf	2.07	2.07	2.07	2.07	2.39	2.39	2.39	2.39	8
mono Si	1.76	1.79	1.80	1.83	1.87	2.36	2.92		5
ortho Si	1.75	1.77	1.81	1.81	1.92	2.27	3.07		5
tetra Si	1.69	1.69	1.69	1.69	2.75	2.75	2.75	2.75	4

Table 6.5: List of (Hf,Si) - O bond lengths for each of the monoclinic, orthorhombic and tetragonal phases, for the undoped case (top three rows) and the 3.125% Si doped case. The number of oxygen neighbors to Hf or Si (coordination number) is reported in the rightmost column.

6.4.2.1 Doping by Si

In order to gain a comprehensive understanding of the effects of doping by one of the elements, we focus on Si which is one of the most widely used dopants in hafnia-based thin films (aside from Zr). In Figure 6.4, we present the environment of a hafnium atom by showing its bonds with the neighboring oxygen atoms for the (a) monoclinic, (b) orthorhombic, and (c) tetragonal phases of HfO_2. We also present the environment of a silicon atom that replaces a hafnium atom after a full relaxation, for the the (d) monoclinic, (e) orthorhombic, and (f) tetragonal phases.

As seen by the number of Hf-O bonds drawn in Figure 6.4, a hafnium atom is seven-fold coordinated by oxygen atoms in the monoclinic and orthorhombic phases and eight-fold coordinated in the tetragonal phase. A silicon dopant becomes five-fold coordinated in the monoclinic and orthorhombic phases, and four-fold coordinated in the tetragonal phase. We list the (Hf,Si)-O distances for all cases in Table 6.5.

We find that the monoclinic and the orthorhombic phases have the same coordination configuration in HfO_2, and the coordination of Si dopant is approximately the

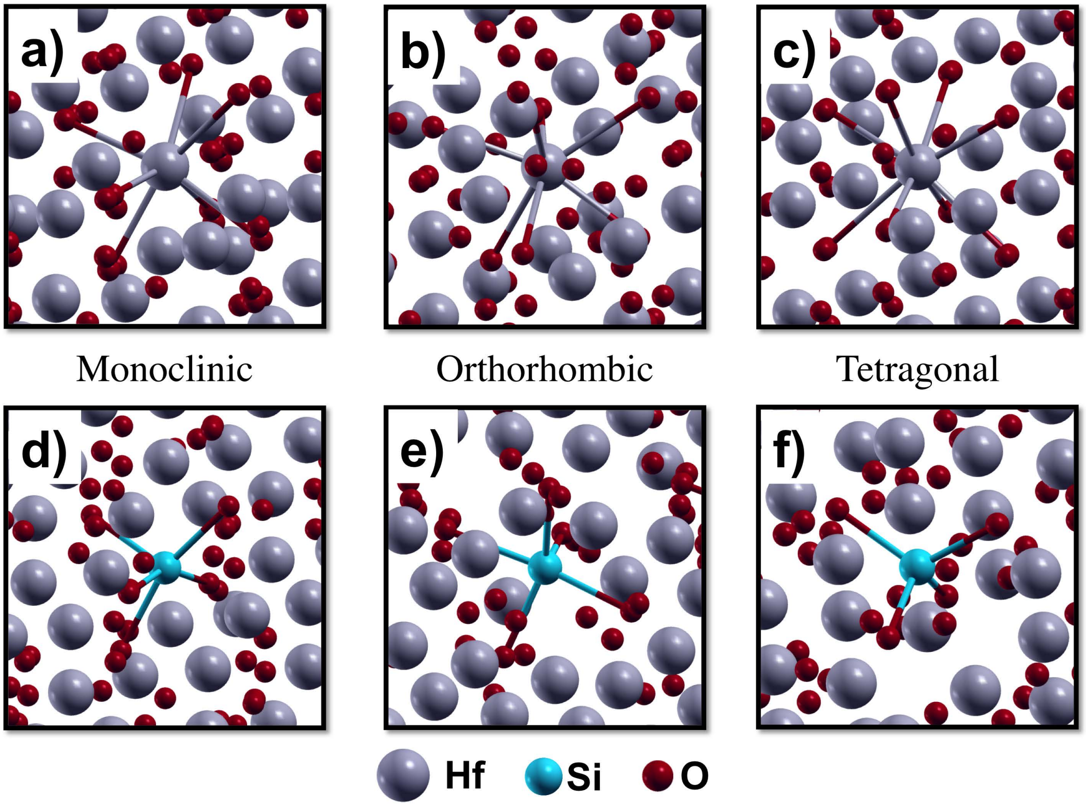


Figure 6.4: The environment of a Hf atom in the (a) monoclinic, (b) orthorhombic, and (c) tetragonal phases of HfO_2 compared with the environment of a Si dopant in the (d) monoclinic, (e) orthorhombic and (f) tetragonal phases. For each case, the bonds between the atom in question and its nearest oxygen neighbors are drawn. See Table 6.5 for the list of (Hf,Si)-O bond lengths in each case and the number of nearest oxygen neighbors (coordination number).

	HfO_2	2% Si (VCA)	3.125% Si (AS)	3.125% Si (VCA)	4% Si (VCA)
ortho	0.08	0.07	0.07	0.05	0.03
tetra	0.16	0.15	0.11	0.11	0.08
cubic	0.26	0.17	0.33	0.12	0.08

Table 6.6: Energies of the orthorhombic, tetragonal and cubic phases with respect to the monoclinic phase, for pure and Si doped HfO_2, as computed by direct atomic substitution (AS) for 3.125% and the virtual crystal approximation (VCA) for 2%, 3.125% and 4%. The energies are listed in eV per unit formula.

same for these two phases. Hence the energy difference between the orthorhombic and the monoclinic phases does not significantly change. However, in the tetragonal phase, the hafnium atom is coordinated by 8 oxygens. A closer inspection reveals that Hf is surrounded by two concentric oxygen tetrahedra with Hf-O distances of 2.07 Å and 2.39 Å. After replacement of this hafnium with a silicon atom, the closer tetrahedron is pulled in and the farther tetrahedron is pushed out so that the distances become 1.69 Å and 2.75 Å. This oxygen environment for silicon is almost identical to its environment in the ground state of bulk SiO_2. In the $P3_121$ (α-quartz) phase of SiO_2, silicon atoms lie in the centers of oxygen tetrahedra with a Si-O distance of 1.63 Å (our calculation). We note that among the low-energy polymorphs of SiO_2, five out of six lowest energy structures feature tetrahedral cages [125]. Therefore, we conclude that the favorable tetrahedral environment of Si dopant in the tetragonal phase causes its significant stabilization with respect to the orthorhombic and the monoclinic phases.

We conclude the discussion on Si doping by comparing our results that are obtained by atomic substitution (AS) and virtual crystal approximation (VCA). In Table 6.6, we list the energies of the orthorhombic, tetragonal and cubic phases with respect to the monoclinic phase, for 2% and 4% Si-doped cases computed by VCA, and the 3.125% Si-doped case computed by AS and VCA. We find that VCA is in agreement with AS for the tetragonal phase, and gives an acceptable result for the orthorhombic phase. We have investigated the disagreement in the cubic phase by first inspecting the environment of the Si dopant in the case of AS. The eightfold coordination of hafnium persists for the silicon dopant. We have then relaxed the structure again after slightly displacing one of the neighboring oxygens, which has resulted in the transformation of the cell into a tetragonal cell, indicating that the cubic phase is unstable toward silicon doping. Hence, for the remainder of our study, we do not discuss the behavior of the cubic phase, which is also not observed in hafnia-based thin films.

6.4.2.2 Hf/Zr mixing

As the other most widely used dopant in hafnia thin films, we now turn to Zr. In Table 6.7 we list the energies of the orthorhombic, tetragonal and cubic phases with respect to the monoclinic phase for bulk $Hf_xZr_{1-x}O_2$, where $x = 1.00, 0.75, 0.50, 0.25, 0.00$. For each case we compare the AS and the VCA results. For AS computations, we used 4-unit formula (12-atom) cells and replaced 0 - 4 Hf atoms in the cell with Zr. For the 50% mixing case, where 2 Hf atoms per cell are substituted by Zr, we compute the energies for all possible 2-atom substitutions in the cell. These differently chosen pairs of atoms lead to relaxed energies within 0.02 eV per unit formula, and the lowest such energy is reported for each phase.

We find that AS and VCA are in very good agreement for the orthorhombic and tetragonal phases, as in the Si-doped case. We have not investigated the cause of the differences between these two methods for the cubic case, but since the cubic phase does not appear to participate in ferroelectricity of hafnia thin films, we have decided to leave this question for future research.

	HfO_2	$Hf_{0.75}Zr_{0.25}O_2$	$Hf_{0.50}Zr_{0.50}O_2$	$Hf_{0.25}Zr_{0.75}O_2$	ZrO_2
ortho (AS)	0.08	0.08	0.08	0.07	0.08
ortho (VCA)	0.08	0.07	0.07	0.07	0.08
tetra (AS)	0.16	0.14	0.14	0.14	0.12
tetra (VCA)	0.16	0.14	0.13	0.12	0.12
cubic (AS)	0.26	0.25	0.23	0.22	0.21
cubic (VCA)	0.26	0.20	0.19	0.19	0.21

Table 6.7: Energies of the orthorhombic, tetragonal and cubic phases with respect to the monoclinic phase, for $Hf_xZr_{1-x}O_2$ where x =1.00, 0.75, 0.50, 0.25 and 0.00, as computed by atomic substitution (AS) and virtual crystal approximation (VCA). The energies are listed in eV per unit formula.

6.4.3 Effects of strain on doped HfO_2

6.4.3.1 Matching planes for bulk phases

It has been observed that ferroelectric hafnia thin films have large numbers of tetragonal grains in the early stages of growth [121, 122]. This is understood to be caused by a reduction of the relative energy of the tetragonal phase through its low energy surface [120] and doping. We have shown that for 3 - 4% Si doping, the tetragonal phase is significantly stabilized, which agrees with the findings of previous computational studies [129, 130] . We have also shown that for Zr mixing above 25% the tetragonal phase is also stabilized, again in agreement with previous studies [130, 127]. Hence it is reasonable to think that a significant ratio of the initial grains during film growth are tetragonal. Then after deposition of the top electrode, some portion of these tetragonal grains transform into orthorhombic grains. We will now investigate this possible scenario in more detail. Our main physical assumption will be that during the potential transformation of the tetragonal gain into other phases, the grain is

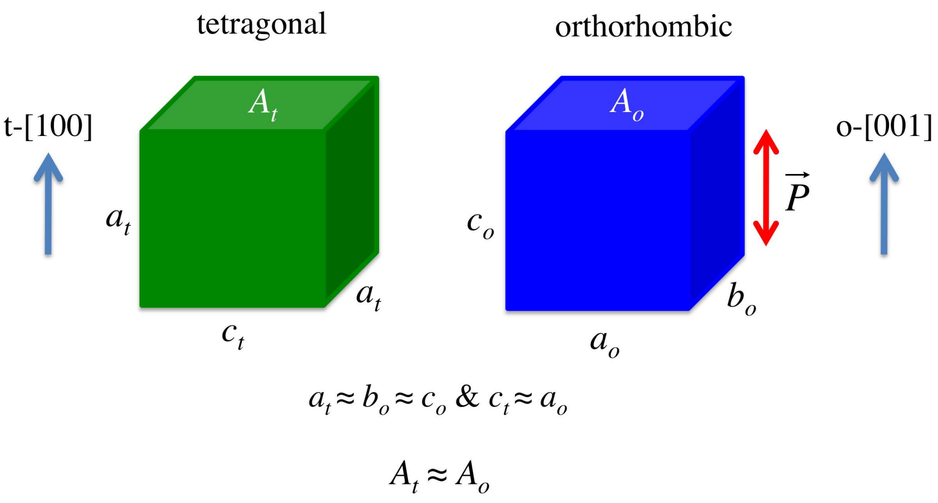


Figure 6.5: Conventional unit cells of the tetragonal and the orthorhombic phases, and the approximate equalities between their lattice parameters and lattice planes. The polarization vector lies in the [001] direction of the orthorhombic phase. A [100] oriented tetragonal grain can transform into a [001] oriented out-of-plane polarized orthorhombic grain by a set of small changes in the lattice parameters.

confined within the film and in approximate epitaxial relation to a substrate: that it can not change its in-plane area significantly during the transformation.

In order for an out-of-plane polarized orthorhombic grain (orientation [001]) to form without a large change in the in-plane lattice parameters, the parent tetragonal grain needs to have the orientation [100] or [010] (which are physically equivalent). We demonstrate these matchings in Figure 6.5. The short sides of the tetragonal phase (a_t) and the orthorhombic phase (b_o, c_o) are similar in length; and the long sides of these two phases (c_t and a_o) are also similar in length (see Table 6.3 for the computed values). Therefore the tetragonal cell can transform into the orthorhombic cell by slightly elongating c_t and slightly contracting one of the a_t.

Repeating the same analysis for the monoclinic phase, we eliminate the m [001] $\longleftrightarrow$ o [001] matching because of the mismatch in lengths, and the m [010] $\longleftrightarrow$ o [001] matching because of the mismatch in the angles between the in-plane lattice vectors.

Therefore the set of matchings that leads to a [001] oriented orthorhombic phase is:

$$\mathrm{m}\,[100] \longleftrightarrow \mathrm{t}\,[100] \longleftrightarrow \mathrm{o}\,[001]\,.$$

6.4.3.2 Effects of strain on undoped grains

To investigate the likelihood of the tetra → mono and the tetra → ortho transformations with these matchings, we have simulated epitaxially strained phases of hafnia via computational relaxations of bulk hafnia strained to pre-specified lattice parameters. For each of the three phases, we have fixed the in-plane lattice parameters to -4%, -2%, 0%, 2% and 4% strained values and have relaxed the third lattice parameter as well as all the atomic positions. In Figure 6.6, we plot the energies of the three phases of HfO_2 vs the area of the matching plane based on this method. For each phase, we fit a third degree polynomial to the five data points we have obtained to draw a smooth curve. The tetragonal, orthonormal and monoclinic phases are shown as green, blue and black curves, respectively.

A t [100] grain of optimized in-plane area may transform into the orthorhombic and the monoclinic phases without changing its area, which would be represented in Figure 6.6 as a downward jump from the bottom of the green curve to a point on either the blue or the black curve. Because at the optimized area of the t [100] grain the monoclinic phase is 0.18 eV lower than the orthorhombic phase, the likelihood of the tetra → mono transformation should be much higher than the likelihood of the tetra → ortho transformation. We also circle in Figure 6.6 the point at which the curves that correspond to the monoclinic and the orthorhombic phases cross. That point corresponds to a 3% compressive biaxial strain with respect to the t [100] grain. Therefore, in the absence of a mechanism that generates such a compressive strain, the grain is expected to transform into a m [100] grain during annealing.

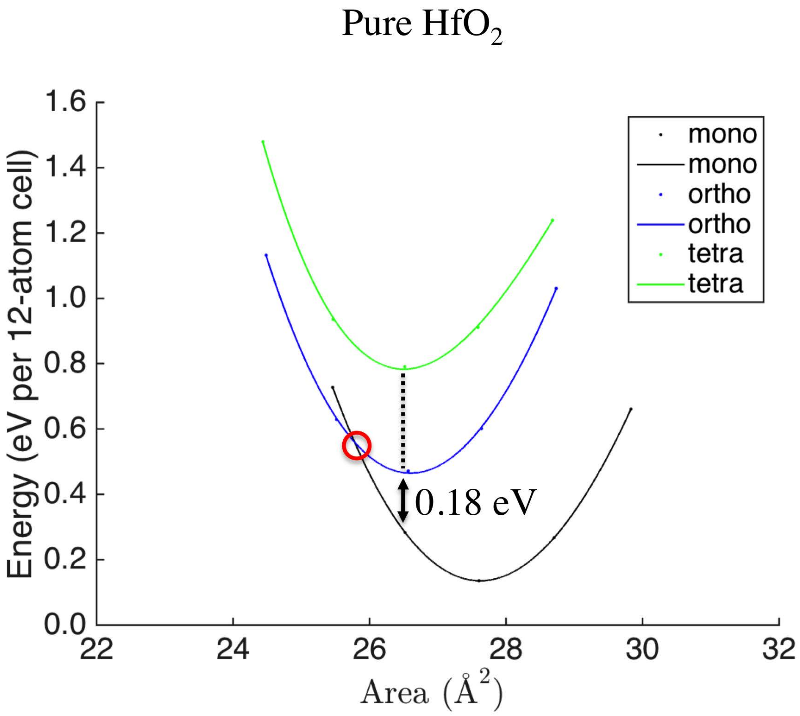
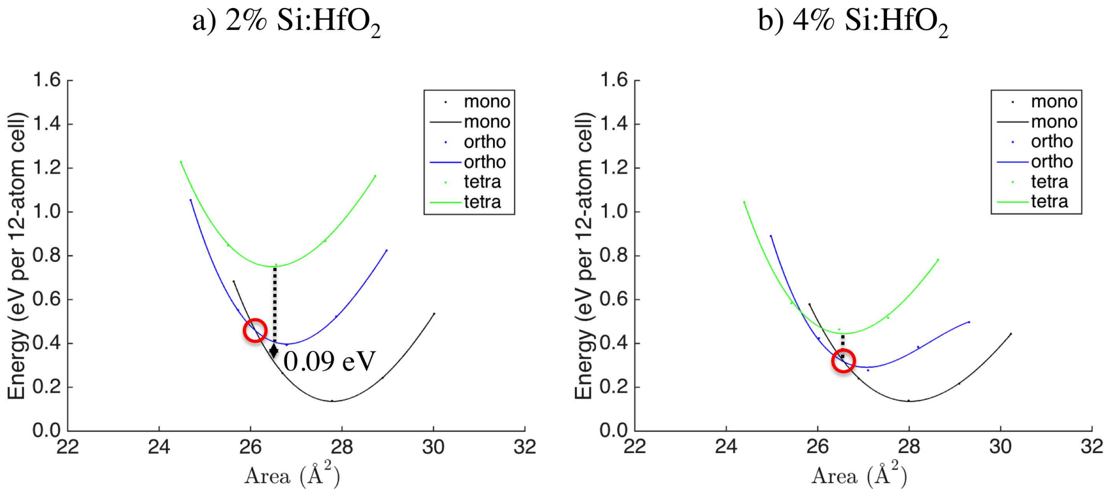


Figure 6.6: Energies of the monoclinic, orthorhombic and tetragonal phases vs in-plane matching area for epitaxially strained bulk simulations of pure HfO_2. For each phase, five data points at -4%, -2%, 0%, 2% and 4% strain are chosen and computed (small dots). The curves are obtained by fitting cubic polynomials to these five data points. The energy difference between the orthorhombic and the monoclinic phases at the optimized area of the t [100] grain is equal to 0.18 eV per 12-atom cell, and labelled in the figure. The zero of energy is chosen arbitrarily.



a) 2% Si:HfO$_2$

b) 4% Si:HfO$_2$

Figure 6.7: Energies of the monoclinic, orthorhombic and tetragonal phases vs in-plane area for epitaxially strained bulk simulations of (a) 2% Si doped and (b) 4% Si doped HfO$_2$. For each composition and phase, five data points at -4%, -2%, 0%, 2% and 4% strain are chosen and computed. The curves are obtained by fitting cubic polynomials to these five data points. The energy difference between the orthorhombic and the monoclinic phases at the optimized area of the t [100] grain is labelled in the figure in (a), and is equal to zero in (b). The zero of energy is chosen arbitrarily.

6.4.3.3 Effects of strain on Si:HfO$_2$

We repeat the same set of simulations for the 2% and the 4% Si doped HfO$_2$ as modeled by the VCA. We present the results in Figure 6.7.

We observe in Figure 6.7 (a) that for the 2% Si doped HfO$_2$, the energy difference between the orthorhombic and the monoclinic phases at the optimized area of t [100] grains is 0.09 eV, which is lower than the undoped value of 0.18 eV. The mono/ortho crossing occurs at 1% compressive strain as opposed to 3% in the undoped case. Hence the formation of the o [001] grains through the t [100] grains is favored by Si doping. In the case of 4% Si doping as shown in Figure 6.7 (b), the mono/ortho crossing occurs at no strain. Thus from a purely energetic point of view, an optimized t [100] grain has equal chance of transforming into a o [001] grain or an m [100] grain.

In Table 6.8 we summarize our findings on the epitaxially strained Si-doped HfO$_2$.

	HfO_2	2% Si	4% Si
% strain where $E\,(\mathrm{o}) = E\,(m)$	-3%	-1%	0%
$E\,(\mathrm{t}) - E\,(\mathrm{o})$ (eV)	0.32	0.34	0.12
$E\,(\mathrm{t}) - E\,(\mathrm{m})$ (eV)	0.50	0.43	0.12

Table 6.8: Summary of the results described in Figures 6.6 and 6.7. The strain values reported are biaxial strain with respect to the optimized in-plane area of the [100] oriented tetragonal phase when the energies of the orthorhombic and the monoclinic phases coincide. The energies are reported in eV per 12-atom cell.

As the doping concentration increases, the energy difference between the orthorhombic and the monoclinic phases at the optimized in-plane area of t [100] grains decreases. At 4% doping, the energies of the o [001] and m [100] grains coincide for the in-plane area that is optimized for t [100] grains. However, the tetra → mono transformation that keeps the area fixed increases the volume of the cell by 5%, whereas the tetra → ortho transformation that keeps the area fixed decreases the volume by 1%. Therefore, in the presence of a top electrode that provides confinement, the tetra → ortho transformation may have a further advantage compared to the tetra → mono transformation. Our findings offer an explanation for the experimental observation that 3 - 4% Si doped films that are subjected to high temperature annealing with a top electrode have ferroelectric properties; i.e. >2% doping (in the case of silicon) and pre-annealing deposition of the top electrode are necessary conditions for ferroelectricity (see Section §6.2).

6.4.3.4 Effects of strain on $Hf_xZr_{1-x}O_2$

As one of the most common hafnia derivatives that has successfully been used as a ferroelectric thin film, we now investigate strain effects in $Hf_xZr_{1-x}O_2$. We present our results in Figure 6.8. We find that the energy difference between the orthorhombic

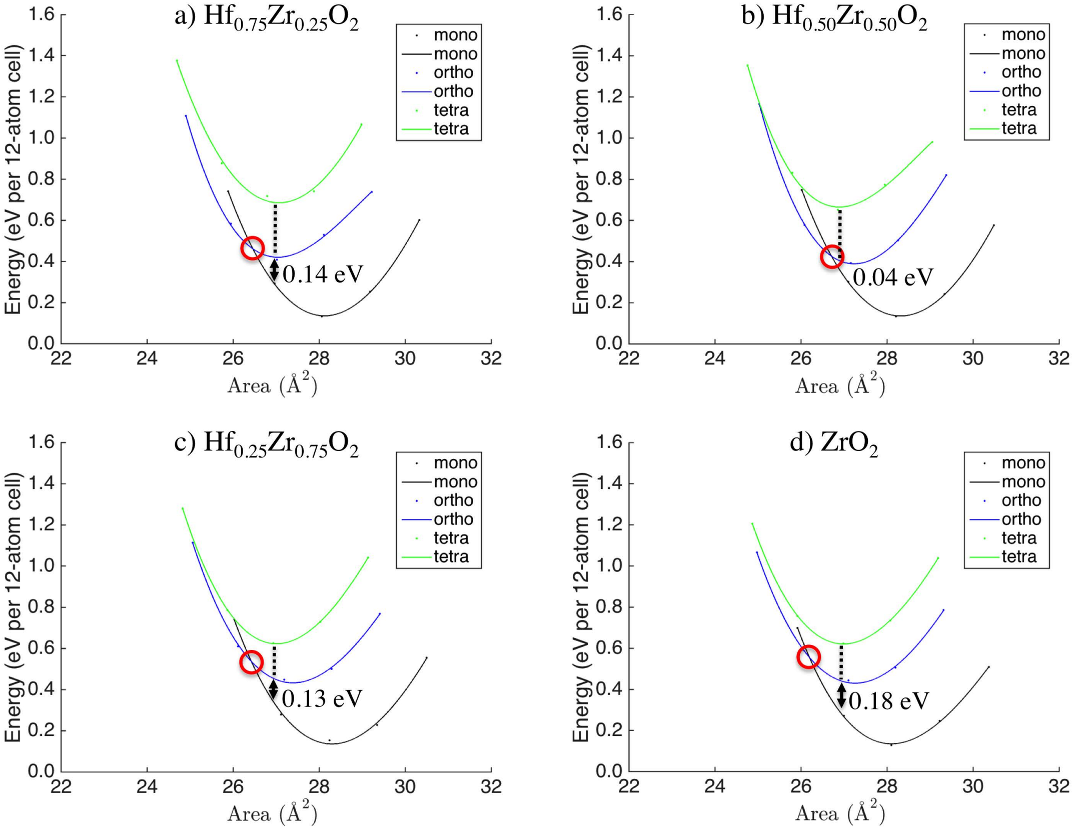


Figure 6.8: Energies of the monoclinic, orthorhombic and tetragonal phases vs in-plane area for epitaxially strained bulk simulations of (a) $Hf_{0.75}Zr_{0.25}O_2$, (b) $Hf_{0.50}Zr_{0.50}O_2$, (c) $Hf_{0.25}Zr_{0.75}O_2$ and (d) pure ZrO_2. For each composition and phase, five data points at -4%, -2%, 0%, 2% and 4% strain are chosen and computed. The curves are obtained by fitting cubic polynomials to these five data points. The energy difference between the orthorhombic and the monoclinic phases at the optimized area of the t [100] grain is labelled in the figure in for each case. The zero of energy is chosen arbitrarily.

and the monoclinic phases at the optimized area of t [100] grains is (a) 0.14 eV for $x = 0.75$, (b) 0.04 eV for $x = 0.50$, (c) 0.13 eV for $x = 0.25$ and (d) 0.18 eV for pure ZrO_2. Hence $HfZrO_4$ ($x = 1/2$) presents the most suitable case for the tetra → ortho transformation.

In Table 6.9 we report our relevant results for $Hf_xZr_{1-x}O_2$. The energy difference between the orthorhombic and the monoclinic phases at the optimized in-plane area of t [100] grains is minimized at $x = 0.50$. For this case, the energies of the o [001]

	HfO_2	$Hf_{0.75}Zr_{0.25}O_2$	$Hf_{0.50}Zr_{0.50}O_2$	$Hf_{0.25}Zr_{0.75}O_2$	ZrO_2
% strain where $E\,(\mathrm{o}) = E\,(m)$	-3%	-2%	-1%	-2%	-3%
$E\,(\mathrm{t}) - E\,(\mathrm{o})$ (eV)	0.32	0.27	0.26	0.17	0.18
$E\,(\mathrm{t}) - E\,(\mathrm{m})$ (eV)	0.50	0.43	0.30	0.30	0.36

Table 6.9: Summary of the results described in Figures 6.6 and 6.8. The strain values reported are biaxial strain with respect to the optimized in-plane area of the [100] oriented tetragonal phase when the energies of the orthorhombic and the monoclinic phases coincide. The energies are reported in eV per 12-atom cell.

and m [100] grains coincide for the in-plane area that is 1% compressively strained with respect to the optimized area for t [100] grains. Without the strain, the tetra → mono transformation is preferred to the tetra → ortho transformation by 0.04 eV per 12-atom cell. However, the former increases the volume by 3%, whereas the latter increases the volume by 1%. Therefore the confinement effects provided by the top electrode may favor the tetra → ortho transformation over the tetra → mono transformation. As in the case of Si doping, our findings explain the experimental observation that 30 - 60% Zr doped films that are annealed with a capping electrode present ferroelectric properties (see Section §3.2).

In addition, we observe that for $Hf_{0,25}Zr_{0.75}O_2$ and ZrO_2, the energy versus strain curves that represent the tetragonal and the orthorhombic phases lie closer to each other (see Figure 6.8). In Table 6.9 we notice that the energy difference between the optimized t [100] grains and the o [100] grains with the same area is lowest for high Zr:Hf ratios. Because the orthorhombic $Pca2_1$ space group is a subgroup of the tetragonal $P4_2/nmc$ space group, the proximity in their energies promotes antiferroelectricity [126]. This supports the experimental observation of antiferroelectric behavior in thin films with higher Zr content (see Section §3.2).

6.4.4 Thin film simulations

In addition to the combined effects of doping and strain, we have performed some preliminary investigations of interface effects. In hafnia-based thin films, ferroelectricity occurs when the film is ∼8 - 24 nm thick, and the grains are generally a few nanometers in size. This makes finite-size effects potentially important. Surfaces of ZrO_2 and HfO_2 have been studied experimentally [131, 132] and theoretically [120, 133, 134, 135] prior to the discovery of ferroelectricity in these films, with a focus on the monoclinic and tetragonal phases. A recent study has included the polar orthorhombic $Pca2_1$ phase into a first principles investigation of surfaces of hafnia [136].

Our goal is to compute the energies of the interfaces between relevant phases of hafnia and typical electrodes such as TiN or Ir. For that purpose, we have begun a computational survey of hafnia thin films. In Figure 6.9, we demonstrate an $Ir/HfO_2/Ir$ stack schematically. In order to isolate thin film effects from strain effects, we fix the in-plane lattice parameters to the lattice parameters of the HfO_2 phase and orientation that we choose to study. We have found that the lattice parameters of HfO_2 are in the range of 5.04 - 5.30 Å (see Table 6.3). On the other hand, typical electrodes used with hafnia thin films, e.g. TiN and Ir have lattice constants of 4.24 Å and 3.90 Å, respectively. Therefore, very large simulation cells would be needed to accurately represent epitaxial interfaces between the hafnia and the electrode with a small lattice mismatch. However, epitaxy is not actually observed in the experimental systems. Therefore, we study the interfaces using much more reasonably sized $\sqrt{2} \times \sqrt{2}$ cells of TiN (a = 6.00 Å) and Ir (a = 5.52 Å) with HfO_2, strained to match various phases and orientations of HfO_2. We emphasize that these are *model* calculations: the theoretical simulation has periodic boundary conditions and thus is always epitaxial, while the experimental interfacial structure is much more complex,

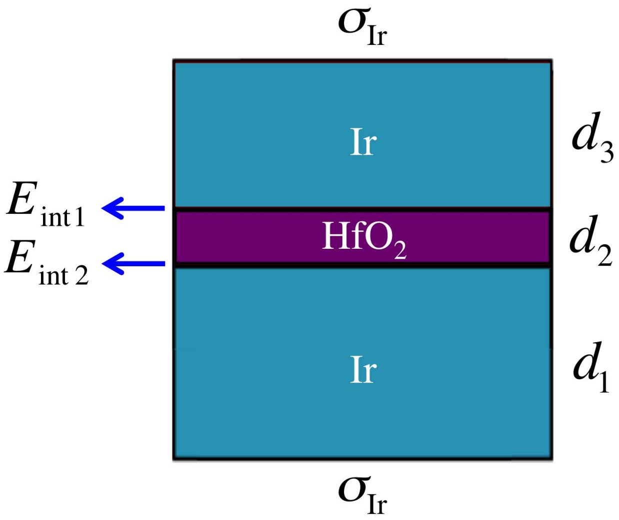


Figure 6.9: Schematic demonstration of a Ir/HfO$_2$/Ir stack simulation. d_1, d_2 and d_3 denote the thicknesses of the components of the stack, $\sigma_{\rm Ir}$ denotes the surface energy of Ir, and $E_{\rm int\ 1}$ and $E_{\rm int\ 2}$ denote interface energies of the two Ir/HfO$_2$ interfaces.

not epitaxial, and unknown with any precision at present. Our aim is to begin a first-principles approach to the interfacial energetics to gauge their approximate size and potential importance.

For the rest of this section, we focus on Ir electrodes. To compute the interfacial energies shown in Figure 6.9, we first compute the surface energy of the free-standing Ir thin film as

$$2\sigma_{\rm Ir} = E^{\rm (Ir)}_{\rm film} - d \times E^{\rm (Ir)}_{\rm bulk}, \tag{6.1}$$

where $E^{\rm (Ir)}_{\rm film}$ is the computed energy of the free-standing strained Ir film, $E^{\rm (Ir)}_{\rm bulk}$ is the energy per u.c. of the bulk strained Ir, and d is the thickness of the film in number of u.c. $E^{\rm (Ir)}_{\rm bulk}$ is computed by straining a plane of bulk Ir to the choice of hafnia phase and orientation, and relaxing the third lattice parameter and atomic positions.

Once we have found $\sigma_{\rm Ir}$, we calculate the interfacial energies E_{int} as

$$E_{\rm int\ 1} + E_{\rm int\ 2} = E_{\rm stack} - (d_1 + d_3)\, E^{\rm (Ir)}_{\rm bulk} - d_2 E^{\rm (HfO_2)}_{\rm bulk} - 2\sigma_{\rm Ir}, \tag{6.2}$$

where E_{stack} is the computed energy of the final materials stack, $E_{\text{bulk}}^{(\text{Ir})}$ is the energy per u.c. of the bulk strained Ir, $E_{\text{bulk}}^{(\text{HfO}_2)}$ is the energy per u.c. of bulk HfO_2, and the thicknesses are shown in Figure 6.9. We note that we can only compute the sum of the two interfacial energies using this approach (if the interfaces are physically identical, then a single interface energy is available).

The relaxed configuration of the Ir/HfO_2/Ir stack with $d_1 = d_2 = d_3 = 2$ u.c. where HfO_2 is in the monoclinic-[001] configuration is shown in Figure 6.1. To find the lowest energy interfaces for a given phase and orientation, we have first simulated possible surface terminations of HfO_2. In order to include an integer number of unit cells in the thin film (i.e., stoichiometric hafnia), we have restricted the terminations to be Hf on one end and O on the other (Hf-. . .-OO) or O terminated on both ends (O-. . .-O). Prior work on zirconia has shown that the energy of an O-. . .-O terminated slab is lower in energy than a Zr-. . .-OO terminated slab by 13.0 eV per in-plane cell for t [001] films for free-standing ZrO_2 [137]. We have found a very similar (and huge) value of 13.4 eV for free -standing HfO_2 by simulating 2 u.c. thick t [001] films. To check that the Hf-. . .-OO termination remains high-energy for oxide/metal interfaces, we have simulated Ir/HfO_2 stacks with the cubic phase for (Ir)-Hf-. . .-OO, (Ir)-OO-. . .-Hf and (Ir)-O-. . .-O terminations. We have found that the (Ir)-Hf-. . .-OO and (Ir)-OO-. . .-Hf stacks are 6.3 and 7.8 eV per in-plane cell higher in energy than the (Ir)-O-. . .-O stack, respectively. Therefore, we have decided to restrict our studies below to the O-. . .-O terminated HfO_2 films for all phases of hafnia.

We have begun studying the interfaces of Ir and pure HfO_2 in the monoclinic, tetragonal, orthorhombic and cubic phases, in all possible inequivalent principle orientations. The only exception is the m [010] orientation, which we exclude because of the non-orthogonal in-plane lattice vectors that do not match the (simulated) substrate structure. To extract interface energies via equation (6.2), we have simulated stacks with $d_1 = d_2 = d_3 = 2$ u.c. We have found that all phases and orientations are

	$E_{\text{int 1}} + E_{\text{int 2}}$ (eV per in-plane cell)
monoclinic-[001]	3.1
monoclinic-[100]	5.1
orthorhombic-[001]	4.0
orthorhombic-[100]	4.5
orthorhombic-[010]	7.1
tetragonal-[001]	4.5
tetragonal-[100]	3.0
cubic-[001]	3.8

Table 6.10: Interface energies for Ir/HfO_2/Ir stacks for each HfO_2 phase and orientation, computed via the equation (6.2), where $d_1 = d_2 = d_3 = 2$ u.c.

mechanically stable in thin film form, with modest relaxations at Ir/HfO_2 interfaces. We list the interface energies we have found in Table 6.10. Further simulations with thicker slabs are in progress in order to check the validity of the interface energies that we have computed. However, the current results suggest that t [100] and m [001] films compete for dominance for very thin films where the interface energy effects may be determinative.

With the interface energies we have obtained via equation (6.2) and listed in Table 6.10, we build an energetic model for variable thickness films based on our interfacial energies and bulk hafnia energies using the formula:

$$E_{\text{film}}^{(\text{ph-or})} = E_{\text{int 1}}^{(\text{ph-or})} + E_{\text{int 2}}^{(\text{ph-or})} + \frac{h}{a_{\text{lattice}}^{(\text{ph-or})}} E_{\text{bulk}}^{(\text{ph})}, \tag{6.3}$$

where $E_{\text{film}}^{(\text{ph-or})}$ is the energy of the thin film of HfO_2 in a given phase and orientation

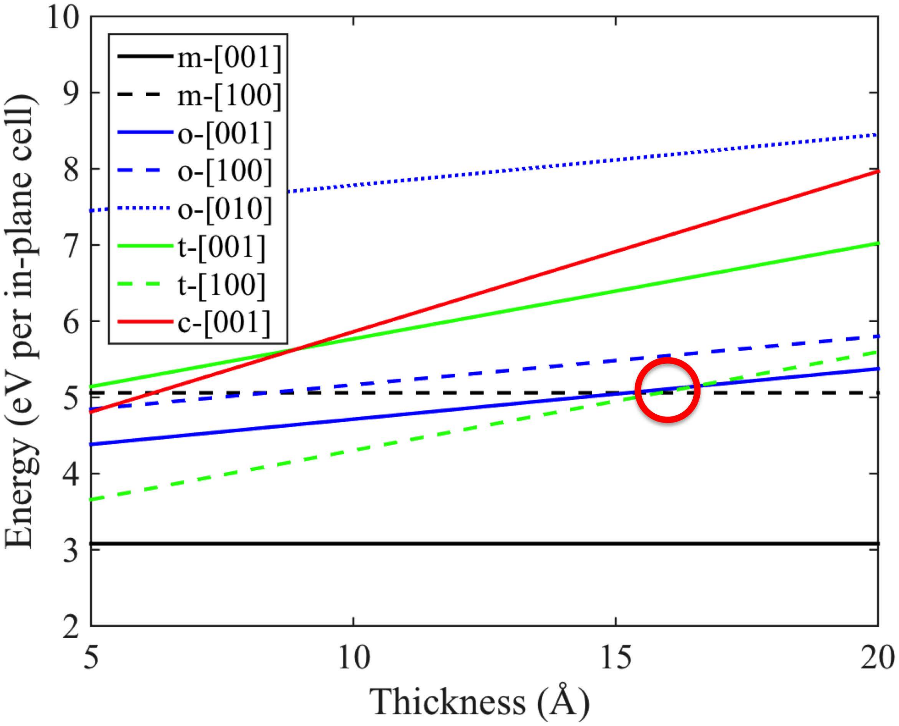


Figure 6.10: Thin film energies computed via the model described in equation (6.3) as a function of film thickness for Ir/HfO_2/Ir stacks. The zero of the bulk energies is taken as the monoclinic phase, causing the lines corresponding to this phase (black) to be flat. The orthorhombic, tetragonal and cubic phases are represented by blue, green and red lines, respectively. The crossing in the energies of t [100], o [001] and m [100] films is marked with a red circle.

“ph-or”, h is the thickness of the oxide in Å, $a_{\text{lattice}}^{(\text{ph-or})}$ is the out-of-plane lattice parameter of HfO_2 in “ph-or” in Å, and $E_{\text{bulk}}^{(\text{ph})}$ is the energy of a u.c. of HfO_2 in phase “ph”.

We summarize the results of this model in Figure 6.10 as a plot of energy vs. thickness for each of the phase-orientation pairs we have studied. We define $E_{\text{bulk}}^{(\text{mono})}$ to be zero, so the lines that correspond to the monoclinic phase are flat. We observe that, because of variations in the interface energies that are of the order of a few eV per in-plane cell, several crossings occur. For the purpose of understanding the stabilization of the o [001] phase, we focus on t [100], o [001] and m [100] films as in section 6.4.3. We find that during the early stages of growth, m [001] and t [100] grains are close in energy (also see Table 6.10 for the interface energy values). Some of the initial t [100] grains may become kinetically trapped as the film grows further. The energy of the t [100] film crosses the energies of the o [001] and the m [100] films approximately at a thickness of 16 Å. For thicker films, o [001] grains are more favorable than t [100] grains, but less favorable than m [100]. Hence this simple model predicts that t [100] grains may initially form during growth and then transform into o [001] and m [100] grains during annealing.

Our modeling thus far has a number of limitations. Two that can be addressed relatively easily are: (1) Further DFT calculations of thicker stacks are in progress that will refine interface energies that are likely to modify the positions of crossings. (2) The simulation for every phase and orientation is done at the unstrained lattice parameters of that phase and orientation. The electrode is assumed to be unaffected by strain in any significant way. This assumption can be tested by applying small strain to each phase and orientation and re-computing the interface energies to check that they do not change in irregular ways.

6.5 Conclusion

We have conducted a first-principles study of doped hafnia with the goal of understanding some of the experimental observations from a structural point of view. We have described the effects of doping with various dopants on the energetics of bulk phases of HfO_2. We have discussed in detail the structural changes that are caused by Si doping. We have compared the two methods of modeling doping: atomic substitution (AS) and virtual crystal approximation (VCA). We have decided that VCA compares well with AS, and used VCA to simulate the effects of epitaxial strain on doped HfO_2. We have found that among 0%, 2% and %4 Si doping, 4% doping provides the best conditions for initial tetragonal [100] grains to transform into orthorhombic [001] grains. We have also found that for $Hf_xZr_{1-x}O_2$, where $x = 1.00, 0.75, 0.50, 0.25, 0.00$, $x = 0.50$ provides the most favorable conditions for the tetragonal [100] → orthorhombic [001] transformation. However, for this transformation to be preferred over the tetragonal [100] → monoclinic [100] transformation, some confinement needs to be present. In experiments, this confinement is provided by a top electrode (typically TiN). Our findings provide an explanation for common experimental observations for the optimal ranges of doping in Si:HfO_2 and $Hf_xZr_{1-x}O_2$. Finally, we have described a model to estimate the interface effects for thin films of HfO_2 based on *ab initio* simulations of Ir/HfO_2/Ir stacks. Our results are preliminary, but offer interesting clues for how the interface effects may be in play for the stabilization of the ferroelectric phase in these films.

Chapter 7

Summary and Outlook

In this Thesis, we have studied three oxides as potential candidates for ferroelectric thin films, using first-principles methods: $BaTiO_3$ epitaxial on Ge (Chapters 3 and 4), ZrO_2 monolayers on Si (Chapter 5), and thin films of doped HfO_2 (Chapter 6). Summaries and future research prospects for each of these systems follow.

7.1 Epitaxial $BaTiO_3$/Ge interface

We have presented our *ab initio* study of the epitaxial $BaTiO_3$/Ge interface. We have obtained two (meta)stable configurations for stoichiometric interfaces with 2×1 periodicity, one of which agrees well with the atomic positions derived from experimental measurements [11]. We have found that the choice of the top electrode determines the relative energies of the two interfaces and can in principle be used to cause them to be degenerate. Because the two interfaces are differently polarized, this interface may in principle show ferroelectric properties. This is in contrast with the previously studied analogous system of $SrTiO_3$/Si, where the polarization is found to be fixed [83]. We have also found that the two (meta)stable configurations have different interfacial chemistries and different band alignments. The difference in the band alignments is so significant that the promoted charge carrier type of the semiconductor changes, which may have novel technological applications. We have performed an analysis of atomic displacements due to the interface via eigenmodes of the force constant matrix of the bulk oxide. We have shown that these eigenmodes can be used to engineer interfaces with desired properties in order to induce physical phenomena such as charge ordering or multiferroicity [11]. We have analyzed the polarization profile of the BTO film with its different interfaces and found that the general behavior agrees with the previous work on STO/Si [63]. Finally, we have studied the thermodynamics and kinetics of oxygen for these films in order to understand the growth procedure that leads to abrupt and stoichiometric films. We have found that kinetic trapping of oxygen in the interfacial layer is crucial, as in the case of STO/Si [93].

Our current addition to this work is the study of the Rashba effect at these interfaces. The Rashba effect arises when the spin-orbit coupling is important and inversion symmetry is broken, such as at surfaces and interfaces [138, 139, 140]. Our preliminary results on the band structure of these interfaces suggest a non-zero Rashba splitting, but further work is in progress. Rashba physics is linked to a rich set of phenomena and technological applications [141, 142].

7.2 ZrO_2 monolayers on Si(001)

Recognizing the limitations of the interface chemistry between a semiconductor such as Ge and a perovskite oxide such as BTO, which restrict the possible modulation of film polarization, we have proposed an ultrathin binary oxide, namely a monolayer of ZrO_2, as a candidate for a polarized thin film that is switchable on a semiconductor, namely Si. We have found multiple (meta)stable structures with various ionic polarizations but similar energy, atomic structure and chemistry. We have briefly summarized the experimental results on amorphous ZrO_2 monolayers on Si(001), which have been shown with C-V measurements to be ferroelectric. To explore the finite-temperature behavior of these films, we have presented a two dimensional discrete lattice model of the domains whose parameters are derived from DFT results. Through cluster-based Monte Carlo simulations, we have concluded that two oppositely polarized structures dominate the films at temperatures of interest. We have proposed that the experimentally observed ferroelectric switching can be modeled as locally switching between these two structures. We have found that for epitaxial films this switching would lead to a VBE shift in silicon that is comparable to the experimental measurements. Finally we have examined monolayers with varying oxygen content, i.e. ZrO_x with $x = 1.0, 1.5, 2.0, 2.5, 3.0$ and presented the main structural trends in these films.

There are several possible avenues to follow to extend this work. First, we are studying epitaxial $SrTiO_3/ZrO_2/Si$ stacks in order to check whether the multiplicity of polarizations we have achieved is retained. Our preliminary results on stacks with 3 atomic layers of STO and 1 ML of ZrO_2 are promising in that we have obtained six (meta)stable configurations with different polarization profiles. This indicates that ZrO_2 may also serve as a buffer layer between Si and STO to induce ferroelectricity in STO, which has been a long-standing research goal [83]. Second, we are working to improve the performance of our Wolff-cluster-type algorithm for the Monte Carlo simulations we run in order to study the domain dynamics in these monolayers. In its current form, our algorithm performs slowly at low temperatures, which is due to low acceptance ratios (see Chapter 2). Modifying the parameter p_{add} in a temperature dependent way should help us overcome this issue. Third, it is possible to expand the discrete lattice model to include low-energy structures of the ZrO_x monolayers where $x = 1.0, 1.5, 2.0, 2.5, 3.0$. This is a longer term endeavor which would require a grand canonical ensemble treatment of the lattice model with varying oxygen content, but would lead to a more complete understanding of the behavior of these monolayers at finite temperatures and oxygen partial pressures.

7.3 Ferroelectricity in HfO_2 thin films

We have conducted a computational study of doped HfO_2 in order to provide an atomistic understanding to some of the experimental observations on hafnia-based ferroelectric thin films. We have described the effects of doping on the relative energies of the bulk phases of hafnia with a number of dopants. We have then performed an analysis of biaxial strain on differently doped HfO_2 in order to determine the doping conditions that are optimal for the formation of the out-of-plane polarized orthorhombic phase. We have found that for Si-doped films, 4% doping, and for Hf-Zr oxides, 50% Zr mixing are the optimal doping conditions. Moreover, we find that annealing after the top electrode is deposited, which is a necessary element for ferroelectricity according to experiments, may provide the confinement which is necessary for the suppression of the monoclinic phase. These findings are in agreement with available experimental results. Finally, we have presented a simple model based on DFT simulations of thin films to estimate the effects of the interfaces between the oxide and the electrode. Our preliminary results offer potential additions to the mechanisms that create the pathway to the ferroelectric phase in hafnia-based thin films.

In the short term, we are working to improve the thin film simulations with the purpose of improving on the interface energies that enter our model for thin film energetics. Specifically, we are testing the dependence of interface energies on the thicknesses of the elements of the electrode/HfO_2/electrode stack. Then, we will study these energies for different strain values to ascertain the importance of in-plane strain, which is bound to come into play during experimental growth. In the longer term, it may be useful to study the effects of doping in thin films in order to test whether combining these two factors lead to unexpected results. On a separate course, we are repeating the analysis of biaxial strain with HfO_2 that is doped with different elements, in order to propose alternate dopants for experiments.

Bibliography

[1] P. Ghosez, X. Gonze, and J. P. Michenaud, Ferroelectrics **220**, 1 (1999), ISSN 0015-0193, URL http://www.tandfonline.com/doi/abs/10.1080/00150199908007992.

[2] J. W. Reiner, F. J. Walker, and C. H. Ahn, Science **323**, 1018 (2009), ISSN 0036-8075, 1095-9203, URL http://www.sciencemag.org/content/323/5917/1018.

[3] J. W. Reiner, A. M. Kolpak, Y. Segal, K. F. Garrity, S. Ismail-Beigi, C. H. Ahn, and F. J. Walker, Advanced Materials **22**, 2919 (2010), ISSN 1521-4095, URL http://onlinelibrary.wiley.com/doi/10.1002/adma.200904306/abstract.

[4] G. D. Wilk, R. M. Wallace, and J. M. Anthony, Journal of Applied Physics **89**, 5243 (2001), ISSN 00218979, URL http://jap.aip.org/resource/1/japiau/v89/i10/p5243_s1.

[5] J. Robertson, Reports on Progress in Physics **69**, 327 (2006), ISSN 0034-4885, URL http://iopscience.iop.org/0034-4885/69/2/R02.

[6] R. A. McKee, F. J. Walker, and M. F. Chisholm, Science **293**, 468 (2001), ISSN 0036-8075, 1095-9203, URL http://www.sciencemag.org/content/293/5529/468.

[7] M. M. Vijatović, J. D. Bobić, and B. D. Stojanović, Science of Sintering **40**, 235 (2008), URL http://www.doiserbia.nb.rs/Article.aspx?id=0350-820X0803235V#.

[8] J. W. Reiner, F. J. Walker, R. A. McKee, C. A. Billman, J. Junquera, K. M. Rabe, and C. H. Ahn, physica status solidi (b) **241**, 2287 (2004), ISSN 1521-3951, URL http://onlinelibrary.wiley.com/doi/10.1002/pssb.200404932/abstract.

[9] T. Boscke, J. Muller, D. Brauhaus, U. Schroder, and U. Bottger, in *Electron Devices Meeting (IEDM), 2011 IEEE International* (2011), pp. 24.5.1–24.5.4.

[10] J. Müller, T. S. Böscke, U. Schröder, S. Mueller, D. Bräuhaus, U. Böttger, L. Frey, and T. Mikolajick, Nano Letters **12**, 4318 (2012), ISSN 1530-6984, URL http://dx.doi.org/10.1021/nl302049k.

[11] D. P. Kumah, M. Dogan, J. H. Ngai, D. Qiu, Z. Zhang, D. Su, E. D. Specht, S. Ismail-Beigi, C. H. Ahn, and F. J. Walker, Physical Review Letters **116**, 106101 (2016), URL http://link.aps.org/doi/10.1103/PhysRevLett.116.106101.

[12] P. Hohenberg and W. Kohn, Physical Review **136**, B864 (1964), URL http://link.aps.org/doi/10.1103/PhysRev.136.B864.

[13] W. Kohn and L. J. Sham, Physical Review **140**, A1133 (1965), URL http://link.aps.org/doi/10.1103/PhysRev.140.A1133.

[14] J. P. Perdew and A. Zunger, Physical Review B **23**, 5048 (1981), URL http://link.aps.org/doi/10.1103/PhysRevB.23.5048.

[15] D. M. Ceperley and B. J. Alder, Physical Review Letters **45**, 566 (1980), URL http://link.aps.org/doi/10.1103/PhysRevLett.45.566.

[16] R. O. Jones and O. Gunnarsson, Reviews of Modern Physics **61**, 689 (1989), URL http://link.aps.org/doi/10.1103/RevModPhys.61.689.

[17] J. P. Perdew and W. Yue, Physical Review B **33**, 8800 (1986), URL http://link.aps.org/doi/10.1103/PhysRevB.33.8800.

[18] J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77**, 3865 (1996), URL http://link.aps.org/doi/10.1103/PhysRevLett.77.3865.

[19] D. Vanderbilt, Physical Review B **41**, 7892 (1990), URL http://link.aps.org/doi/10.1103/PhysRevB.41.7892.

[20] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Reviews of Modern Physics **73**, 515 (2001), URL http://link.aps.org/doi/10.1103/RevModPhys.73.515.

[21] C. Wert and C. Zener, Physical Review **76**, 1169 (1949), URL http://link.aps.org/doi/10.1103/PhysRev.76.1169.

[22] G. H. Vineyard, Journal of Physics and Chemistry of Solids **3**, 121 (1957), ISSN 0022-3697, URL http://www.sciencedirect.com/science/article/pii/0022369757900598.

[23] J. K. Norskov, F. Studt, F. Abild-Pedersen, and T. Bligaard, *Fundamental Concepts in Heterogeneous Catalysis* (Wiley, 2014), URL http://www.wiley.com/WileyCDA/WileyTitle/productCd-1118888952.html.

[24] G. Henkelman, G. Jóhannesson, and H. Jónsson, in *Theoretical Methods in Condensed Phase Chemistry*, edited by S. D. Schwartz (Springer Netherlands, 2002), no. 5 in Progress in Theoretical Chemistry and Physics, pp. 269–302, ISBN 978-0-7923-6687-4 978-0-306-46949-7, dOI: 10.1007/0-306-46949-9_10, URL http://link.springer.com/chapter/10.1007/0-306-46949-9_10.

[25] G. Henkelman, B. P. Uberuaga, and H. Jónsson, The Journal of Chemical Physics **113**, 9901 (2000), ISSN 0021-9606, 1089-7690, URL http://scitation.aip.org/content/aip/journal/jcp/113/22/10.1063/1.1329672.

[26] E. Luijten, in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, edited by M. F. Professor, G. C. Professor, and K. B. Professor (Springer Berlin Heidelberg, 2006), no. 703 in Lecture Notes in Physics, pp. 13–38, ISBN 978-3-540-35270-9 978-3-540-35273-0, dOI: 10.1007/3-540-35273-2_1, URL http://link.springer.com/chapter/10.1007/3-540-35273-2_1.

[27] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller, The Journal of Chemical Physics **21**, 1087 (1953), ISSN 0021-9606, URL http://aip.scitation.org/doi/abs/10.1063/1.1699114.

[28] G. T. Barkema and M. E. J. Newman, arXiv:cond-mat/9703179 (1997), arXiv:cond-mat/9703179, URL http://arxiv.org/abs/cond-mat/9703179.

[29] R. H. Swendsen and J.-S. Wang, Physical Review Letters **58**, 86 (1987), URL http://link.aps.org/doi/10.1103/PhysRevLett.58.86.

[30] U. Wolff, Physical Review Letters **62**, 361 (1989), URL http://link.aps.org/doi/10.1103/PhysRevLett.62.361.

[31] E. J. Moon, J. M. Rondinelli, N. Prasai, B. A. Gray, M. Kareev, J. Chakhalian, and J. L. Cohn, Physical Review B **85**, 121106 (2012), URL http://link.aps.org/doi/10.1103/PhysRevB.85.121106.

[32] J. M. Rondinelli, S. J. May, and J. W. Freeland, MRS Bulletin **37**, 261 (2012), ISSN 1938-1425, 0883-7694, URL https://doi.org/10.1557/mrs.2012.49.

[33] M. K. Stewart, J. Liu, M. Kareev, J. Chakhalian, and D. N. Basov, Physical Review Letters **107**, 176401 (2011), URL http://link.aps.org/doi/10.1103/PhysRevLett.107.176401.

[34] S. J. May, C. R. Smith, J.-W. Kim, E. Karapetrova, A. Bhattacharya, and P. J. Ryan, Physical Review B **83**, 153411 (2011), URL http://link.aps.org/doi/10.1103/PhysRevB.83.153411.

[35] J. H. Lee, L. Fang, E. Vlahos, X. Ke, Y. W. Jung, L. F. Kourkoutis, J.-W. Kim, P. J. Ryan, T. Heeg, M. Roeckerath, et al., Nature **466**, 954 (2010), ISSN 0028-0836, URL http://www.nature.com/nature/journal/v466/n7309/full/nature09331.html.

[36] C. Ederer and N. A. Spaldin, Physical Review Letters **95**, 257601 (2005), URL http://link.aps.org/doi/10.1103/PhysRevLett.95.257601.

[37] D. G. Schlom, L.-Q. Chen, C.-B. Eom, K. M. Rabe, S. K. Streiffer, and J.-M. Triscone, *Strain Tuning of Ferroelectric Thin Films** (2007), URL http://www.annualreviews.org/doi/10.1146/annurev.matsci.37.061206.113016.

[38] E. Bousquet, M. Dawber, N. Stucki, C. Lichtensteiger, P. Hermet, S. Gariglio, J.-M. Triscone, and P. Ghosez, Nature **452**, 732 (2008), ISSN 0028-0836, URL http://www.nature.com/nature/journal/v452/n7188/full/nature06817.html.

[39] P. Zubko, S. Gariglio, M. Gabay, P. Ghosez, and J.-M. Triscone, *Interface Physics in Complex Oxide Heterostructures* (2011), URL http://www.annualreviews.org/doi/10.1146/annurev-conmatphys-062910-140445.

[40] N. A. Spaldin and M. Fiebig, Science **309**, 391 (2005), ISSN 0036-8075, 1095-9203, URL http://science.sciencemag.org/content/309/5733/391.

[41] N. A. Benedek and C. J. Fennie, Physical Review Letters **106**, 107204 (2011), URL http://link.aps.org/doi/10.1103/PhysRevLett.106.107204.

[42] J. M. Tranquada, J. D. Axe, N. Ichikawa, A. R. Moodenbaugh, Y. Nakamura, and S. Uchida, Physical Review Letters **78**, 338 (1997), URL http://link.aps.org/doi/10.1103/PhysRevLett.78.338.

[43] Y. Segal, J. W. Reiner, A. M. Kolpak, Z. Zhang, S. Ismail-Beigi, C. H. Ahn, and F. J. Walker, Physical Review Letters **102**, 116101 (2009), URL http://link.aps.org/doi/10.1103/PhysRevLett.102.116101.

[44] M. B. Nardelli, F. J. Walker, and R. A. McKee, physica status solidi (b) **241**, 2279 (2004), ISSN 1521-3951, URL http://onlinelibrary.wiley.com/doi/10.1002/pssb.200404942/abstract.

[45] A. Lin, X. Hong, V. Wood, A. A. Verevkin, and C. H. Ahn, Applied Physics Letters **78**, 2034 (2001), ISSN 0003-6951, URL http://aip.scitation.org/doi/10.1063/1.1358848.

[46] R. A. McKee, F. J. Walker, and M. F. Chisholm, Physical Review Letters **81**, 3014 (1998), URL http://link.aps.org/doi/10.1103/PhysRevLett.81.3014.

[47] M. K. Hudait, Y. Zhu, N. Jain, D. Maurya, and Y. Zhou, Journal of Applied Physics **114**, 024303 (2013), ISSN 0021-8979, URL http://aip.scitation.org/doi/10.1063/1.4813226.

[48] C. Merckling, G. Saint-Girons, C. Botella, G. Hollinger, M. Heyns, J. Dekoster, and M. Caymax, Applied Physics Letters **98**, 092901 (2011), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/98/9/10.1063/1.3558997.

[49] K. D. Fredrickson, P. Ponath, A. B. Posadas, M. R. McCartney, T. Aoki, D. J. Smith, and A. A. Demkov, Applied Physics Letters **104**, 242908 (2014), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/104/24/10.1063/1.4883883.

[50] Y. Liang, J. Kulik, T. C. Eschrich, R. Droopad, Z. Yu, and P. Maniar, Applied Physics Letters **85**, 1217 (2004), ISSN 0003-6951, URL http://aip.scitation.org/doi/10.1063/1.1783016.

[51] R. Kile, Y. Zhu, E. I. Altman, and Y. Liang, Applied Physics Letters **87**, 143106 (2005), ISSN 0003-6951, URL http://aip.scitation.org/doi/10.1063/1.2077837.

[52] J. L. Wang, A. Pancotti, P. Jégou, G. Niu, B. Gautier, Y. Y. Mi, L. Tortech, S. Yin, B. Vilquin, and N. Barrett, Physical Review B **84**, 205426 (2011), URL http://link.aps.org/doi/10.1103/PhysRevB.84.205426.

[53] A. A. Demkov and A. B. Posadas, *Integration of Functional Oxides with Semiconductors* (Springer New York, 2014), ISBN 978-1-4614-9319-8 978-1-4614-9320-4, dOI: 10.1007/978-1-4614-9320-4_1, URL http://link.springer.com/book/10.1007%2F978-1-4614-9320-4.

[54] G. Shirane, H. Danner, and R. Pepinsky, Physical Review **105**, 856 (1957), URL http://link.aps.org/doi/10.1103/PhysRev.105.856.

[55] H. J. W. Zandvliet, Physics Reports **388**, 1 (2003), ISSN 0370-1573, URL http://www.sciencedirect.com/science/article/pii/S037015730300334X.

[56] P. Krüger and J. Pollmann, Physical Review Letters **74**, 1155 (1995), URL http://link.aps.org/doi/10.1103/PhysRevLett.74.1155.

[57] B. J. Sternlieb, J. P. Hill, U. C. Wildgruber, G. M. Luke, B. Nachumi, Y. Moritomo, and Y. Tokura, Physical Review Letters **76**, 2169 (1996), URL http://link.aps.org/doi/10.1103/PhysRevLett.76.2169.

[58] D. Reznik, L. Pintschovius, M. Ito, S. Iikubo, M. Sato, H. Goka, M. Fujita, K. Yamada, G. D. Gu, and J. M. Tranquada, Nature **440**, 1170 (2006), ISSN 0028-0836, URL http://www.nature.com/nature/journal/v440/n7088/full/nature04704.html.

[59] A. Cammarata and J. M. Rondinelli, Physical Review B **87**, 155135 (2013), URL http://link.aps.org/doi/10.1103/PhysRevB.87.155135.

[60] S.-H. Baek and C.-B. Eom, Acta Materialia **61**, 2734 (2013), ISSN 1359-6454, URL http://www.sciencedirect.com/science/article/pii/S1359645412007069.

[61] D. P. Kumah, J. W. Reiner, Y. Segal, A. M. Kolpak, Z. Zhang, D. Su, Y. Zhu, M. S. Sawicki, C. C. Broadbridge, C. H. Ahn, et al., Applied Physics Letters **97**, 251902 (2010), ISSN 00036951, URL http://apl.aip.org/resource/1/applab/v97/i25/p251902_s1.

[62] M. Björck and G. Andersson, Journal of Applied Crystallography **40**, 1174 (2007), ISSN 0021-8898, URL http://scripts.iucr.org/cgi-bin/paper?aj5091.

[63] A. M. Kolpak and S. Ismail-Beigi, Physical Review B **85**, 195318 (2012), URL http://link.aps.org/doi/10.1103/PhysRevB.85.195318.

[64] N. Marzari, D. Vanderbilt, A. De Vita, and M. C. Payne, Physical Review Letters **82**, 3296 (1999), URL http://link.aps.org/doi/10.1103/PhysRevLett.82.3296.

[65] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, Davide Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al., Journal of Physics: Condensed Matter **21**, 395502 (2009), ISSN 0953-8984, URL http://stacks.iop.org/0953-8984/21/i=39/a=395502.

[66] L. Bengtsson, Physical Review B **59**, 12301 (1999), URL http://link.aps.org/doi/10.1103/PhysRevB.59.12301.

[67] P. Ghosez, E. Cockayne, U. V. Waghmare, and K. M. Rabe, Physical Review B **60**, 836 (1999), URL http://link.aps.org/doi/10.1103/PhysRevB.60.836.

[68] W. Zhong, D. Vanderbilt, and K. M. Rabe, Physical Review Letters **73**, 1861 (1994), URL http://link.aps.org/doi/10.1103/PhysRevLett.73.1861.

[69] M. P. Marder, *Condensed matter physics* (John Wiley, 2000), ISBN 978-0-471-17779-1, google-Books-ID: hs7vAAAAMAAJ, URL http://www.wiley.com/WileyCDA/WileyTitle/productCd-0470617985.html.

[70] H. Y. Hwang, Y. Iwasa, M. Kawasaki, B. Keimer, N. Nagaosa, and Y. Tokura, Nature Materials **11**, 103 (2012), ISSN 1476-1122, URL http://www.nature.com/nmat/journal/v11/n2/abs/nmat3223.html.

[71] J. Mannhart and D. G. Schlom, Science **327**, 1607 (2010), ISSN 0036-8075, 1095-9203, URL http://science.sciencemag.org/content/327/5973/1607.

[72] M. Buongiorno Nardelli, F. J. Walker, and R. A. McKee, physica status solidi (b) **241**, 2279 (2004), ISSN 1521-3951, URL http://onlinelibrary.wiley.com/doi/10.1002/pssb.200404942/abstract.

[73] R. Droopad, R. Contreras-Guerrero, J. P. Veazey, Q. Qiao, R. F. Klie, and J. Levy, Microelectronic Engineering **109**, 290 (2013), ISSN

0167-9317, URL http://www.sciencedirect.com/science/article/pii/S0167931713003535.

[74] A. Posadas, M. Berg, H. Seo, A. d. Lozanne, A. A. Demkov, D. J. Smith, A. P. Kirk, D. Zhernokletov, and R. M. Wallace, Applied Physics Letters **98**, 053104 (2011), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/98/5/10.1063/1.3549301.

[75] R. P. Laughlin, D. A. Currie, R. Contreras-Guererro, A. Dedigama, W. Priyantha, R. Droopad, N. Theodoropoulou, P. Gao, and X. Pan, Journal of Applied Physics **113**, 17D919 (2013), ISSN 0021-8979, 1089-7550, URL http://scitation.aip.org/content/aip/journal/jap/113/17/10.1063/1.4796150.

[76] C. Dubourdieu, J. Bruley, T. M. Arruda, A. Posadas, J. Jordan-Sweet, M. M. Frank, E. Cartier, D. J. Frank, S. V. Kalinin, A. A. Demkov, et al., Nature Nanotechnology **8**, 748 (2013), ISSN 1748-3387, URL http://www.nature.com/nnano/journal/v8/n10/full/nnano.2013.192.html.

[77] Z. Li, X. Guo, H.-B. Lu, Z. Zhang, D. Song, S. Cheng, M. Bosman, J. Zhu, Z. Dong, and W. Zhu, Advanced Materials **26**, 7185 (2014), ISSN 1521-4095, URL http://onlinelibrary.wiley.com/doi/10.1002/adma.201402527/abstract.

[78] J. H. Ngai, D. P. Kumah, C. H. Ahn, and F. J. Walker, Applied Physics Letters **104**, 062905 (2014), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/104/6/10.1063/1.4864648.

[79] P. Ponath, K. Fredrickson, A. B. Posadas, Y. Ren, X. Wu, R. K. Vasudevan, M. Baris Okatan, S. Jesse, T. Aoki, M. R. McCartney, et al., Nature Communi-

cations **6**, 6067 (2015), URL http://www.nature.com/ncomms/2015/150114/ncomms7067/full/ncomms7067.html.

[80] C. J. Först, C. R. Ashman, K. Schwarz, and P. E. Blöchl, Nature **427**, 53 (2004), ISSN 0028-0836, URL http://www.nature.com/nature/journal/v427/n6969/abs/nature02204.html.

[81] S. A. Chambers, Y. Liang, Z. Yu, R. Droopad, J. Ramdani, and K. Eisenbeiser, Applied Physics Letters **77**, 1662 (2000), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/77/11/10.1063/1.1310209.

[82] X. Zhang, A. A. Demkov, H. Li, X. Hu, Y. Wei, and J. Kulik, Physical Review B **68**, 125323 (2003), URL http://link.aps.org/doi/10.1103/PhysRevB.68.125323.

[83] A. M. Kolpak, F. J. Walker, J. W. Reiner, Y. Segal, D. Su, M. S. Sawicki, C. C. Broadbridge, Z. Zhang, Y. Zhu, C. H. Ahn, et al., Physical Review Letters **105**, 217601 (2010), URL http://link.aps.org/doi/10.1103/PhysRevLett.105.217601.

[84] F. Amy, A. S. Wan, A. Kahn, F. J. Walker, and R. A. McKee, Journal of Applied Physics **96**, 1635 (2004), ISSN 0021-8979, 1089-7550, URL http://scitation.aip.org/content/aip/journal/jap/96/3/10.1063/1.1766417.

[85] G. Niu, S. Yin, G. Saint-Girons, B. Gautier, P. Lecoeur, V. Pillard, G. Hollinger, and B. Vilquin, Microelectronic Engineering **88**, 1232 (2011), ISSN 0167-9317, URL http://www.sciencedirect.com/science/article/pii/S0167931711002875.

[86] L. Kornblum, M. D. Morales-Acosta, E. N. Jin, C. H. Ahn, and F. J. Walker,

Advanced Materials Interfaces **2**, n/a (2015), ISSN 2196-7350, URL http://onlinelibrary.wiley.com/doi/10.1002/admi.201500193/abstract.

[87] G. H. Kwei, A. C. Lawson, S. J. L. Billinge, and S. W. Cheong, The Journal of Physical Chemistry **97**, 2368 (1993), ISSN 0022-3654, URL http://dx.doi.org/10.1021/j100112a043.

[88] M. K. Hudait, Y. Zhu, N. Jain, D. Maurya, Y. Zhou, R. Varghese, and S. Priya, ACS Applied Materials & Interfaces **5**, 11446 (2013), ISSN 1944-8244, URL http://pubs.acs.org/doi/abs/10.1021/am4036866.

[89] M. Needels, M. C. Payne, and J. D. Joannopoulos, Physical Review Letters **58**, 1765 (1987), URL http://link.aps.org/doi/10.1103/PhysRevLett.58.1765.

[90] M. Needels, M. C. Payne, and J. D. Joannopoulos, Physical Review B **38**, 5543 (1988), URL http://link.aps.org/doi/10.1103/PhysRevB.38.5543.

[91] Y. Yoshimoto, Y. Nakamura, H. Kawai, M. Tsukada, and M. Nakayama, Physical Review B **61**, 1965 (2000), URL http://link.aps.org/doi/10.1103/PhysRevB.61.1965.

[92] P. Tipler and R. Llewellyn, *Modern Physics* (W. H. Freeman, 2003), ISBN 978-0-7167-4345-3, URL https://books.google.com/books?id=tpU18JqcSNkC.

[93] A. M. Kolpak and S. Ismail-Beigi, Physical Review B **83**, 165318 (2011), URL http://link.aps.org/doi/10.1103/PhysRevB.83.165318.

[94] *NIST Chemistry WebBook*, URL http://webbook.nist.gov.

[95] K. Suzuki and K. Kijima, Japanese Journal of Applied Physics **44**, 2081 (2005), ISSN 1347-4065, URL http://iopscience.iop.org/1347-4065/44/4R/2081.

[96] C. Kittel, *Introduction to Solid State Physics* (Wiley, 2004), ISBN 978-0-471-41526-8, google-Books-ID: kym4QgAACAAJ.

[97] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Physical Review B **44**, 943 (1991), URL http://link.aps.org/doi/10.1103/PhysRevB.44.943.

[98] Z. H. Lu, M. J. Graham, D. T. Jiang, and K. H. Tan, Applied Physics Letters **63**, 2941 (1993), ISSN 0003-6951, URL http://aip.scitation.org/doi/10.1063/1.110279.

[99] S.-Y. Wu, IEEE Transactions on Electron Devices **21**, 499 (1974), ISSN 0018-9383.

[100] A. Ramstad, G. Brocks, and P. J. Kelly, Physical Review B **51**, 14504 (1995), URL http://link.aps.org/doi/10.1103/PhysRevB.51.14504.

[101] Ó. Paz, A. J. R. da Silva, J. José Sáenz, and E. Artacho, Surface Science **482–485, Part 1**, 458 (2001), ISSN 0039-6028, URL http://www.sciencedirect.com/science/article/pii/S0039602800010220.

[102] A. V. Bune, V. M. Fridkin, S. Ducharme, L. M. Blinov, S. P. Palto, A. V. Sorokin, S. G. Yudin, and A. Zlatkin, Nature **391**, 874 (1998), ISSN 0028-0836, URL http://www.nature.com/nature/journal/v391/n6670/full/391874a0.html.

[103] F. Y. Wu, Reviews of Modern Physics **54**, 235 (1982), URL http://link.aps.org/doi/10.1103/RevModPhys.54.235.

[104] M. A. Neto, R. A. dos Anjos, and J. R. de Sousa, Physical Review B **73**, 214439 (2006), URL http://link.aps.org/doi/10.1103/PhysRevB.73.214439.

[105] Y. Miyamoto and A. Oshiyama, Physical Review B **41**, 12680 (1990), URL http://link.aps.org/doi/10.1103/PhysRevB.41.12680.

[106] T. Uchiyama and M. Tsukada, Surface Science **357–358**, 509 (1996), ISSN 0039-6028, URL `http://www.sciencedirect.com/science/article/pii/0039602896800768`.

[107] Y.-K. Chiou, C.-H. Chang, and T.-B. Wu, Journal of Materials Research **22**, 1899 (2007), ISSN 2044-5326, 0884-2914, URL `https://doi.org/10.1557/jmr.2007.0242`.

[108] M. H. Park, Y. H. Lee, H. J. Kim, Y. J. Kim, T. Moon, K. D. Kim, J. Müller, A. Kersch, U. Schroeder, T. Mikolajick, et al., Advanced Materials **27**, 1811 (2015), ISSN 1521-4095, URL `http://onlinelibrary.wiley.com/doi/10.1002/adma.201404531/abstract`.

[109] J. Müller, U. Schröder, T. S. Böscke, I. Müller, U. Böttger, L. Wilde, J. Sundqvist, M. Lemberger, P. Kücher, T. Mikolajick, et al., Journal of Applied Physics **110**, 114113 (2011), ISSN 0021-8979, 1089-7550, URL `http://scitation.aip.org/content/aip/journal/jap/110/11/10.1063/1.3667205`.

[110] S. Mueller, J. Mueller, A. Singh, S. Riedel, J. Sundqvist, U. Schroeder, and T. Mikolajick, Advanced Functional Materials **22**, 2412 (2012), ISSN 1616-3028, URL `http://onlinelibrary.wiley.com/doi/10.1002/adfm.201103119/abstract`.

[111] E. Yurchuk, J. Müller, S. Knebel, J. Sundqvist, A. P. Graham, T. Melde, U. Schröder, and T. Mikolajick, Thin Solid Films **533**, 88 (2013), ISSN 0040-6090, URL `http://www.sciencedirect.com/science/article/pii/S0040609012016264`.

[112] M. H. Park, H. J. Kim, Y. J. Kim, W. Lee, T. Moon, and C. S. Hwang, Applied Physics Letters **102**, 242905 (2013), ISSN 0003-6951, 1077-

3118, URL http://scitation.aip.org/content/aip/journal/apl/102/24/10.1063/1.4811483.

[113] M. H. Park, H. J. Kim, Y. J. Kim, T. Moon, and C. S. Hwang, Applied Physics Letters **104**, 072901 (2014), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/104/7/10.1063/1.4866008.

[114] P. D. Lomenzo, P. Zhao, Q. Takmeel, S. Moghaddam, T. Nishida, M. Nelson, C. M. Fancher, E. D. Grimley, X. Sang, J. M. LeBeau, et al., Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena **32**, 03D123 (2014), ISSN 2166-2746, URL http://avs.scitation.org/doi/10.1116/1.4873323.

[115] U. Schroeder, E. Yurchuk, J. Müller, D. Martin, T. Schenk, P. Polakowski, C. Adelmann, M. I. Popovici, S. V. Kalinin, and T. Mikolajick, Japanese Journal of Applied Physics **53**, 08LE02 (2014), ISSN 0021-4922, 1347-4065, URL http://stacks.iop.org/1347-4065/53/i=8S1/a=08LE02?key=crossref.ad3265465d10aa004e56539e1d4feb68.

[116] M. H. Park, H. J. Kim, Y. J. Kim, W. Lee, T. Moon, and C. S. Hwang, Applied Physics Letters **105**, 072902 (2014), ISSN 0003-6951, URL http://aip.scitation.org/doi/10.1063/1.4893376.

[117] X. Sang, E. D. Grimley, T. Schenk, U. Schroeder, and J. M. LeBeau, Applied Physics Letters **106**, 162905 (2015), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/106/16/10.1063/1.4919135.

[118] M. Hoffmann, U. Schroeder, T. Schenk, T. Shimizu, H. Funakubo, O. Sakata, D. Pohl, M. Drescher, C. Adelmann, R. Materlik, et al.,

Journal of Applied Physics **118**, 072006 (2015), ISSN 0021-8979, 1089-7550, URL http://scitation.aip.org/content/aip/journal/jap/118/7/10.1063/1.4927805.

[119] A. Chernikova, M. Kozodaev, A. Markeev, D. Negrov, M. Spiridonov, S. Zarubin, O. Bak, P. Buragohain, H. Lu, E. Suvorova, et al., ACS Applied Materials & Interfaces **8**, 7232 (2016), ISSN 1944-8244, URL http://dx.doi.org/10.1021/acsami.5b11653.

[120] R. C. Garvie, The Journal of Physical Chemistry **82**, 218 (1978), ISSN 0022-3654, URL http://dx.doi.org/10.1021/j100491a016.

[121] K. Tomida, K. Kita, and A. Toriumi, Applied Physics Letters **89**, 142902 (2006), ISSN 0003-6951, URL http://aip.scitation.org/doi/10.1063/1.2355471.

[122] T. S. Böscke, P. Y. Hung, P. D. Kirsch, M. A. Quevedo-Lopez, and R. Ramirez-Bon, Applied Physics Letters **95**, 052904 (2009), ISSN 0003-6951, URL http://aip.scitation.org/doi/10.1063/1.3195623.

[123] L. Bellaiche and D. Vanderbilt, Physical Review B **61**, 7877 (2000), URL http://link.aps.org/doi/10.1103/PhysRevB.61.7877.

[124] J. Wang, H. P. Li, and R. Stevens, Journal of Materials Science **27**, 5397 (1992), ISSN 0022-2461, 1573-4803, URL http://link.springer.com/article/10.1007/BF00541601.

[125] Q. Zeng, A. R. Oganov, A. O. Lyakhov, C. Xie, X. Zhang, J. Zhang, Q. Zhu, B. Wei, I. Grigorenko, L. Zhang, et al., Acta Crystallographica Section C Structural Chemistry **70**, 76 (2014), ISSN 2053-2296, URL http://scripts.iucr.org/cgi-bin/paper?S2053229613027861.

[126] S. E. Reyes-Lillo, K. F. Garrity, and K. M. Rabe, Physical Review B **90**, 140103 (2014), URL http://link.aps.org/doi/10.1103/PhysRevB.90.140103.

[127] R. Materlik, C. Künneth, and A. Kersch, Journal of Applied Physics **117**, 134109 (2015), ISSN 0021-8979, 1089-7550, URL http://scitation.aip.org/content/aip/journal/jap/117/13/10.1063/1.4916707.

[128] T. D. Huan, V. Sharma, G. A. Rossetti, and R. Ramprasad, Physical Review B **90**, 064111 (2014), URL http://link.aps.org/doi/10.1103/PhysRevB.90.064111.

[129] D. Fischer and A. Kersch, Journal of Applied Physics **104**, 084104 (2008), ISSN 0021-8979, 1089-7550, URL http://scitation.aip.org/content/aip/journal/jap/104/8/10.1063/1.2999352.

[130] C.-K. Lee, E. Cho, H.-S. Lee, C. S. Hwang, and S. Han, Physical Review B **78**, 012102 (2008), URL http://link.aps.org/doi/10.1103/PhysRevB.78.012102.

[131] M. W. Pitcher, S. V. Ushakov, A. Navrotsky, B. F. Woodfield, G. Li, J. Boerio-Goates, and B. M. Tissue, Journal of the American Ceramic Society **88**, 160 (2005), ISSN 1551-2916, URL http://onlinelibrary.wiley.com/doi/10.1111/j.1551-2916.2004.00031.x/abstract.

[132] J. Chevalier, L. Gremillard, A. V. Virkar, and D. R. Clarke, Journal of the American Ceramic Society **92**, 1901 (2009), ISSN 1551-2916, URL http://onlinelibrary.wiley.com/doi/10.1111/j.1551-2916.2009.03278.x/abstract.

[133] A. B. Mukhopadhyay, J. F. Sanz, and C. B. Musgrave, Physical Review B **73**, 115330 (2006), URL http://link.aps.org/doi/10.1103/PhysRevB.73.115330.

[134] X. Luo, A. A. Demkov, D. Triyoso, P. Fejes, R. Gregory, and S. Zollner, Physical Review B **78**, 245314 (2008), URL http://link.aps.org/doi/10.1103/PhysRevB.78.245314.

[135] X. Luo, W. Zhou, S. V. Ushakov, A. Navrotsky, and A. A. Demkov, Physical Review B **80**, 134119 (2009), URL http://link.aps.org/doi/10.1103/PhysRevB.80.134119.

[136] R. Batra, H. D. Tran, and R. Ramprasad, Applied Physics Letters **108**, 172902 (2016), ISSN 0003-6951, 1077-3118, URL http://scitation.aip.org/content/aip/journal/apl/108/17/10.1063/1.4947490.

[137] A. Eichler and G. Kresse, Physical Review B **69**, 045402 (2004), URL http://link.aps.org/doi/10.1103/PhysRevB.69.045402.

[138] L. Petersen and P. Hedegård, Surface Science **459**, 49 (2000), ISSN 0039-6028, URL http://www.sciencedirect.com/science/article/pii/S0039602800004416.

[139] G. Bihlmayer, Y. M. Koroteev, P. M. Echenique, E. V. Chulkov, and S. Blügel, Surface Science **600**, 3888 (2006), ISSN 0039-6028, URL http://www.sciencedirect.com/science/article/pii/S0039602806004195.

[140] K. V. Shanavas, Z. S. Popović, and S. Satpathy, Physical Review B **90**, 165108 (2014), URL http://link.aps.org/doi/10.1103/PhysRevB.90.165108.

[141] D. Bercioux and P. Lucignano, Reports on Progress in Physics **78**, 106001 (2015), ISSN 0034-4885, URL http://stacks.iop.org/0034-4885/78/i=10/a=106001.

[142] A. Manchon, H. C. Koo, J. Nitta, S. M. Frolov, and R. A. Duine, Nature Ma-

terials **14**, 871 (2015), ISSN 1476-1122, URL http://www.nature.com/nmat/journal/v14/n9/abs/nmat4360.html.