From ab initio properties of the Si-SiO₂ interface, to electrical characteristics of metal-oxide-semiconductor devices

S Markov¹², P Sushko³, C Fiegna², E Sangiorgi², A Shluger³, A Asenov⁴
¹ Dept. of Electronics & Electrical Eng., Rankine Building, Glasgow, G12 8LT, UK
² ARCES-IUNET, University of Bologna, via Venezia 52, Cesena 47023, Italy
³ Dept. of Physics and Astronomy, University College of London, WC1E 6BT, UK
E-mail: s.markov@elec.gla.ac.uk

Abstract. The gradual transition of the band-gap at the Si-SiO₂ interface affects quantisation and leakage characteristics of MOS inversion layer. We establish a link between first principles DFT simulations of the interface, and continuum simulations in the effective mass approximation, in order to obtain a realistic description of the band-gap transition for device modelling. The simplistic approach of obtaining real-space-dependent band-gap profile from the ab initio calculated electronic structure results in uncertainty of the simulated device characteristics. This uncertainty is small however, when compared to the magnitude of the simulated impact of the transition layer. A linear transition of the band-gap over 6 – 7 Å in the oxide approximates well the effects simulated with the realistic band-gap profile from DFT.

1. Introduction
The change of the atomic structure at the Si-SiO₂ interface leaves additional electronic states, energetically aligned closer to the Si conduction band, in the first 2 – 6 Å of the SiO₂ away from the interface, which implies a non-abrupt transition of the band-gap within the oxide [1]. This transition is of great importance for the characterization of gate oxide (or high permittivity gate stacks) and gate leakage in metal-oxide-semiconductor (MOS) field effect transistors (FETs), since the width of the interface transition (a few Å) is of the order of the oxide thickness itself [2]. Although physical understanding of the involved phenomena could be obtained with the help of ab initio modelling methods (e.g. density functional theory), DFT modelling is impractical for device simulations (of model systems of 10⁸ atoms, at a finite bias), and even a tight-binding approach presents a limiting computational cost for that purpose [3]. Accordingly, we established a hierarchical simulation approach, allowing DFT-calculated electronic properties of the interface to be incorporated as real-space-dependent macroscopic parameters (e.g. band-gap/offsets) in our device simulations [4]. The translation of ab initio electronic structure to macroscopic material parameters is not without ambiguity however, owing to certain freedoms in setting up the DFT simulations (e.g. choice of functional), and the procedure of band-gap determination and scaling. In this paper we explore the extent to which such uncertainties influence the magnitude of the simulated impact on device characteristics, due to the band-gap transition layer.

Corresponding author; SM acknowledges financial support by European Community via “Marie Curie” Fellowship (EDITH: MEST-CT-2004-504195)
2. Methodology

We aim to account for certain microscopic properties of the Si-SiO$_2$ interface in macroscopic, continuum device simulations. Our hierarchical simulation approach consists of three steps: 1) calculation of the atomic and electronic structure of the oxide interface, based on DFT simulations of a 3D-periodic SiO$_2$-Si-SiO$_2$ super-cell, adopting α-quartz model for SiO$_2$; 2) translation of the electronic structure to continuous, one-dimensional space-dependent conduction and valence band offsets (relative to the conduction band edge of bulk Si); 3) one-dimensional device simulation of a metal/oxide/p-Si(001) structure, using a self-consistent Poisson-Schrödinger (PS) solver. The details of the atomic structure model and optimisation during step 1), and of the PS-solver of step 3), are described elsewhere [4]. Here, we elaborate on step 2).

The starting point for this is the electronic structure calculated for the selected super-cell with CRYSTAL-2003 computer code, using gradient-corrected density functionals and Gaussian-type basis set. The total density of states (DOS) is obtained using the one-electron $E(k)$ dispersion, calculated at a pre-defined set of $k$-points in reciprocal space. The corresponding one-electron states are represented as linear combination of atomic orbitals, allowing the projection of the total DOS on such an atomic basis, and therefore, the calculation of the atom-projected DOS (pDOS). Results for pDOS, obtained in this way depend on the chosen density functional and the choice of pre-defined $k$-points. For the super-cell used in our study, we have three sets of pDOS data: SET1 (PBE functional), SET2 (B3LYP functional), SET3 (B3LYP functional, and 2 times denser $k$-grid, compared to other two sets).

Figure 1 shows the conduction and valence band-edge (CBE and VBE) profiles extracted from the pDOS of SET1. The band-edges are determined from the highest occupied and lowest unoccupied states associated with each atom. Different criteria of the minimum pDOS value result in slightly different band-edge profile. For $0.05 < \text{min}(\text{pDOS}) < 0.1$ (values in a.u.), the extracted profiles nearly overlap. We find the same qualitative picture for the other two sets, and similar profiles presented in other numerical and experimental studies [5]. Note that oxygen atoms within the sub-stoichiometric oxide (delimited by dotted, vertical line in figure 1) contribute states that are energetically in the same range as the ones contributed from Si-atoms. Analogous effect is observed for the valence-band edge.

Figure 1. CBE and VBE profiles for SET1, for three band-edge determination criteria, as shown. Results for $0.05<\text{min}(\text{pDOS})<0.1$ nearly overlap. Si(SiO$_2$) is along the positive(negative) direction.

Figure 2. CBE and VBE profiles for the three different sets. The band gap strongly depends on the chosen functional. The number of $k$-points affects mostly SiO$_2$ VBE.

Figure 2 compares the band-edge profiles extracted for each of the three sets, at the same band-edge determination criterion of $(\text{pDOS})>0.1$. It highlights a quantitative discrepancy, arising from the use of different functionals. Both PBE and B3LYP underestimate the experimental (bulk) SiO$_2$ band-gap of approximately 8.9 eV, and overestimate (significantly for the B3LYP sets) the bulk Si band-gap (1.12 eV). The latter effect is partly due to the small thickness of ordered Si (5 Å simulated). The former is attributed to the DFT approach, as applied here, and requires scaling of the band gaps prior
to incorporation of band-edge profiles in the device simulator. Figure 2 also shows the effect of a denser k-point grid, to flatten the band-edge in the “bulk”-like parts, most notably for the SiO₂ VBE.

For device simulation, not only the band-gap $E_G$, but also the conduction and valence band offsets $\Delta E_C$ and $\Delta E_V$, are important. As we are interested in the inversion and gate leakage characteristics of a metal/SiO₂/p-Si(100) structure, the penetration of electrons in the oxide becomes the main impact factor, and it depends on the potential barrier for electrons at the interface. Therefore, we choose $\Delta E_C$ to be the leading parameter for scaling of the oxide region. The accuracy of the oxide band-gap is relaxed in this case, leading to an overestimation of the valence band offset. This has no effect in our study, since hole density only in the bulk-Si is relevant. For the region of the Si substrate $E_G$ is the leading scaling parameter.

The scaled conduction and valence band-edge profiles, $E_C$ and $E_V$, respectively, are obtained from:

$$E_C(z) = E_0 + \alpha(z) \Delta E_C^{DFT}(z),$$

and

$$E_V(z) = E_C(z) - \alpha(z) E_G^{DFT}(z),$$

where $E_0$ is our energy reference, taken as the CBE of bulk Si. The scaling coefficient $\alpha$ has a different value in SiO₂ and in Si, as discussed above:

$$\alpha(SiO₂) = \frac{\Delta E_C^{NOM}}{\Delta E_C^{DFT}},$$

and

$$\alpha(Si) = \frac{E_G^{NOM}}{E_G^{DFT}},$$

where $\Delta E_C^{NOM}$=3.15 eV, and $E_G^{NOM}$=1.12 eV. In the sub-stoichiometric oxide, $\alpha$ is assumed to change linearly from the value in Si, to that in SiO₂:

$$\alpha(SiOX) = \left(\frac{\alpha(SiO₂) - \alpha(Si)}{t_{SiOx}}\right) z + \alpha(Si),$$

where $z$ is the distance from the interface, and $t_{SiOx}$ is the thickness of the sub-stoichiometric oxide (~2 Å for the α-quartz SiO₂).

Figure 3 shows the CBE and VBE profiles around the oxide interface, for the three data sets, after scaling according to the above equations. Although the three profiles look quite similar, the relative difference between them exceeds twice the thermal potential (~26 meV), in the transition region. In the next section, we show the implication of this discrepancy, on the simulated device characteristics.

3. Device simulation results

We refer to the CBE and VBE profiles obtained as discussed above as realistic, despite their quantitative discrepancy between each other. Comparison of these profiles with a linearly approximated band-gap transition is shown in figure 3. The main feature of the realistic profiles is the rather slow initial development of the band-gap, followed by a steeper rise. This feature is reported in other DFT and experimental studies of the Si-SiO₂ interface [5], and is in contrast to the linear profile of a similar total transition width (~6 Å).

Figure 4 shows conduction band and electron density (inset) in a MOS structure biased in strong inversion. The different curves reflect the different barrier models from figure 3 (realistic, DFT profile corresponds to SET1). The lowest two subbands (hosting almost the entire electron population in the
substrate) are energetically close to the bottom of the quantum well. For the realistic profile, the well is wider, and the potential barrier is thinner. This induces more inversion charge in proximity to the oxide relative to the simulated linear and abrupt cases. Therefore, the impact of the non-abrupt band-gap transition is enhanced, due to the non-linear band-edge profile, and the exact profile of the band evolution, rather than the total transition width, determines the magnitude of the observed effect.

Figures 5 shows the energy level and occupancy (percentage of the total electron density) of the lowest two subbands, as a function of gate bias, for all simulated interface barrier profiles. The most important observation here is that the uncertainty in the definition of the realistic profiles (resulting from SET1 to SET3) is reflected in a spread of the device characteristics (note the group of dashed curves). Quantitatively however, this spread is small compared to the magnitude of the simulated impact, due to band-gap transition. The same conclusion is drawn from figure 6, showing the direct tunnelling gate current simulated for all simulated interface barrier profiles.

![Figure 5](image-url) Energy (top) and occupancy (bottom) of the lowest two subbands. The spread of realistic curves is small compared to their shift from abrupt line.

![Figure 6](image-url) Gate tunnelling current density. Similar impact of the transition layer is observed, as in figure 6.

4. Conclusions

We elaborated on a simple procedure for obtaining a real-space-dependent band-gap and band-offsets from DFT calculations of the electronic structure of the oxide interface. Despite the inherent uncertainties associated with the translation of the microscopic data into continuum, macroscopic parameters, a more realistic picture of the band-gap transition is obtained, suitable for device modelling and simulation. Any of the simulated realistic barrier models exaggerates the magnitude of the impact that we have previously simulated with linear barrier model (up to 6 Å transition width). The spread in the characteristics, due to the uncertainty in the realistic profile, may be regarded as an uncertainty in the transition width $t_{TR}$, of a linear barrier model between 6 – 7 Å.

5. References

[1] Muller D A, Sorsch T, Moccio S, Baumann F H, Evans-Lutterodt K and Timp G 2000 Nature 399 758; Hirose K, Nohira H, Azuma K and Hattori T 2007 Prog Surf Sci 82 3

[2] Watanabe H, Matsushita D and Muraoka K 2006 IEEE Trans Elec Dev 53 1323

[3] Stadele M, Sacconi F, Di Carlo A and Lugli P 2003 J. Appl Phys 93 2681

[4] Markov S, Sushko P V, Roy S, Fiegna C, Sangiorgi E, Shluger A and Asenov A 2008 Phys Stat Sol (A) 205 1290

[5] Yamasaki T, Kaneta C, Uchiyama T, Uda T and Terakura K 2001 Phys Rev B 63 115314; Giustino F and Pasquarello A 2005 Surf Sci 586 183; Xue K, Ho H P and Xu J B 2007 J Phys D: Appl Phys 40 2886