Modelling of oxygen vacancy aggregates in monoclinic HfO$_2$: can they contribute to conductive filament formation?

To cite this article: Samuel R Bradley et al 2015 J. Phys.: Condens. Matter 27 415401

View the article online for updates and enhancements.

Related content
- Diffusion and aggregation of oxygen vacancies in amorphous silica
  Manveer S Munde, David Z Gao and Alexander L Shluger
- Optimal stoichiometry for nucleation and growth of conductive filaments in HfO$_x$
  K P McKenna
- Effects of interaction between defects on the uniformity of dopa HfO$_2$-based RRAM: a first principle study
  Zhao Qiang, Zhou Maoxiu, Zhang Wei et al.

Recent citations
- Filament-to-dielectric band alignments in TiO$_2$ and HfO$_2$ resistive RAMs
  Ze-Han Wu et al
- Modeling resistive switching materials and devices across scales
  Stefano Ambrogio et al
- The Role of Ti Buffer Layer Thickness on the Resistive Switching Properties of Hafnium Oxide-Based Resistive Switching Memories
  Sk, Ziaur Rahaman et al
Modelling of oxygen vacancy aggregates in monoclinic HfO$_2$: can they contribute to conductive filament formation?

Samuel R Bradley$^1$, Gennadi Bersuker$^2$ and Alexander L Shluger$^1$

$^1$ Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
$^2$ The Aerospace Corporation, Los Angeles, CA 90009-2957, USA

E-mail: samuel.bradley.10@ucl.ac.uk

Received 2 June 2015, revised 15 July 2015
Accepted for publication 4 August 2015
Published 28 September 2015

Abstract

Formation of metal rich conductive filaments and their rearrangements determine the switching characteristics in HfO$_2$ based resistive random access memory (RRAM) devices. The initiation of a filament formation process may occur either via aggregation of pre-existing vacancies randomly distributed in the oxide or via generation of new oxygen vacancies close to the pre-existing ones. We evaluate the feasibility of vacancy aggregation processes by calculating the structures and binding energies of oxygen vacancy aggregates consisting of 2, 3 and 4 vacancies in bulk monoclinic (m)-HfO$_2$ using density functional theory (DFT). We demonstrate that formation of neutral oxygen vacancy aggregates is accompanied by small energy gain, which depends on the size and shape of the aggregate. In the most strongly bound configurations, vacancies are unscreened by Hf cations and form voids within the crystal, with the larger aggregates having larger binding energy per vacancy ($-0.11$ to $-0.18$ eV). The negatively charged di-vacancy was found to have similar binding energies to the neutral one, while the positively charged di-vacancy was found to be unstable. Thus aggregation process of either neutral or negatively charged oxygen vacancies is energetically feasible.

Keywords: HfO$_2$, conductive filament, oxygen vacancy aggregation, DFT

(Some figures may appear in colour only in the online journal)
HfO₂ crystal grains, suggesting that mobile vacancies tend to aggregate there [19]. Conductive-AFM measurements of the electron conductance through HfO₂ have shown that switchable conductive paths are initiated in the grain boundary regions, and then grow extending into the surrounding crystal grains (see e.g. [9, 17, 20, 21]). However, the mechanisms of creation of oxygen deficient conductive paths through monocristalline and amorphous hafnia films have not yet been fully examined on atomistic level and is the focus on this study.

In monoclinic (m) HfO₂, planes of 3 coordinated (3C) and 4 coordinated (4C) O atoms alternate. Several theoretical studies have demonstrated that individual O vacancies in m-HfO₂ are stable in the neutral, positively and negatively charged states [22–24]. Different computational methods predict that the 4C vacancy has a lower formation energy than the 3C vacancy by ~0.1 eV (see e.g. [22, 24]). These vacancies create localized states in the middle of the HfO₂ band gap [25–27]. The mobility of the vacancies depends on their charge state, with the doubly positively charged vacancy having a low diffusion barrier of 0.5–0.7 eV compared to a much higher barrier of 2.4 eV for a neutral vacancy [28]. It is likely that the applied forming voltage will stabilize the vacancies in the neutral or negative charge states, so the initial stage of a CF growth process may proceed via neutral or negative O vacancies randomly diffusing and aggregating. Taking into account that activation barriers for such diffusion are high, effective aggregation is possible only if there is attraction between vacancies.

The formation of oxygen vacancy clusters has been studied in cerium, vanadium and titanium oxides [10]. The preliminary studies of the interaction between neutral as well as charged vacancies in m-HfO₂ have predicted these to be either weakly attractive or repulsive with the binding energy about 0.1 eV, depending on the coordination and the distance between vacancies [29, 30]. However, these calculations considered only selected vacancy pairs and have been carried out in small periodic cells.

In this work, we evaluate the feasibility of O vacancy aggregation process by calculating the structure and binding energies of aggregates consisting of 2, 3 and 4 O vacancies in m-HfO₂. The calculations confirm that there is attraction between two neutral vacancies as well as between individual vacancies and their larger aggregates so that diffusing vacancies can aggregate efficiently.

2. Methods of calculations

Most of the density functional theory (DFT) calculations were carried out using the projector augmented wave (PAW) method as implemented in the Vienna *ab initio* simulation package, [31, 32] and using the Perdew-Wang 91 (PW91) GGA functional [33, 34]. To optimize the geometries of the multiple vacancy configurations 320–324 atom, 3 × 3 × 3, supercells of m-HfO₂ were used. All calculations were carried out in the gamma point and geometries were optimized using the conjugate-gradient algorithm until the maximum residual forces were less than 0.06 eV Å⁻¹.

The formation energies of the 4C and 3C single oxygen vacancies in HfO₂ were calculated in this way as 6.52 and 6.60 eV, respectively.

Since vacancy aggregates are supposed to serve as percolation paths for trap-assistant tunnelling between electrodes, [26, 35] we also considered negatively charged di-vacancies. However, the size of the band gap is underestimated using the PW91 functional and any attempt to calculate a negatively charged vacancy simply gives a neutral vacancy with the added electrons in the HfO₂ conduction band. This remains true for the di-vacancy case. The hybrid functional HSE06 [36] gives a much more accurate band gap (5.6 eV) and was used here to calculate the negatively charged vacancies in HfO₂. Due to computational constraints, a 94–96 atom unit cells were used in these calculations.

A Lany–Zunger (LZ) correction based on an anisotropic point charge was applied to remove the spurious interactions between the charged defects and their periodic images [37, 38].

The formation energy of an aggregate of *n* neutral O vacancies, $E_{\text{form},A}$, was calculated as

$$E_{\text{form},A} = E_{\text{perf}} - (E_A + nE_O).$$

Here $E_A$ is the total energy of the relaxed defective supercell, $E_{\text{perf}}$ is the energy of the perfect cell, $n$ is the number of removed O atoms, and $E_O$ is the energy of half an oxygen molecule calculated in an asymmetric unit cell (~4.89 eV). In these calculations we assume that the Fermi level is suitably high that the neutral vacancies are stable.

The stability of the vacancy aggregate is described by the binding energy between all the constituent vacancies. This is calculated by taking the formation energy of the aggregate using the equation (1) and deducting the formation energies of the constituent vacancies.

$$E_{\text{bind},A} = E_{\text{form},A} - E_{\text{for},V} - E_{\text{for},V} - ...$$

Here $A$ is a vacancy aggregate and $V$ is a vacancy within the aggregate with coordination number $r, r$ etc. Since the vacancy formation energies are calculated individually in separate calculations, the binding energy is the change in energy due to vacancies being simultaneously brought from an infinite separation into the aggregate. A negative $E_{\text{bind}}$ would indicate that the aggregate is more stable than its parts, whereas a positive $E_{\text{bind}}$ would describe an unstable aggregate which is unlikely to form at all. This method effectively describes an aggregation process, regardless of the size of the aggregate and the way in which vacancies come together via diffusion.

3. Neutral O vacancy aggregates

Table 1 shows the binding energies of various multiple vacancy clusters. Each cluster is a different configuration of the neutral O vacancies within the m-HfO₂ lattice. A negative $E_{\text{bind}}$ indicates a gain in energy from bringing infinitely separated vacancies together into the vacancy aggregate. These results show that all neutral vacancy clusters are stable with a binding energy which varies depending on the aggregate’s size and shape. The neutral di-vacancy binding energies are
Table 1. Binding energies of the calculated neutral vacancy aggregates. D, T and Q indicate di-, tri- and tetra-vacancies respectively. Each configuration is labelled by a letter in parentheses.

<table>
<thead>
<tr>
<th>Defect</th>
<th>(E_{\text{bind},A}) (eV)</th>
<th>Defect</th>
<th>(E_{\text{bind},A}) (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D(A)_{444}^0)</td>
<td>-0.01</td>
<td>(T(A)_{444}^0)</td>
<td>-0.15</td>
</tr>
<tr>
<td>(D(B)_{444}^0)</td>
<td>-0.05</td>
<td>(T(B)_{444}^0)</td>
<td>-0.10</td>
</tr>
<tr>
<td>(D(C)_{444}^0)</td>
<td>-0.11</td>
<td>(T(C)_{444}^0)</td>
<td>-0.09</td>
</tr>
<tr>
<td>(D(D)_{444}^0)</td>
<td>-0.03</td>
<td>(T(D)_{444}^0)</td>
<td>-0.45</td>
</tr>
<tr>
<td>(D(E)_{444}^0)</td>
<td>-0.17</td>
<td>(T(E)_{444}^0)</td>
<td>-0.22</td>
</tr>
<tr>
<td>(D(F)_{444}^0)</td>
<td>-0.16</td>
<td>(T(F)_{444}^0)</td>
<td>-0.18</td>
</tr>
<tr>
<td>(D(G)_{444}^0)</td>
<td>-0.21</td>
<td>(T(G)_{444}^0)</td>
<td>-0.34</td>
</tr>
<tr>
<td>(D(H)_{444}^0)</td>
<td>-0.18</td>
<td>(T(H)_{444}^0)</td>
<td>-0.39</td>
</tr>
<tr>
<td>(Q(A)_{44444}^0)</td>
<td>-0.14</td>
<td>(Q(E)_{4444}^0)</td>
<td>-0.42</td>
</tr>
<tr>
<td>(Q(B)_{44443}^0)</td>
<td>-0.43</td>
<td>(Q(F)_{44443}^0)</td>
<td>-0.45</td>
</tr>
<tr>
<td>(Q(C)_{44443}^0)</td>
<td>-0.31</td>
<td>(Q(G)_{44443}^0)</td>
<td>-0.71</td>
</tr>
<tr>
<td>(Q(D)_{44443}^0)</td>
<td>-0.54</td>
<td>(Q(H)_{44443}^0)</td>
<td>-0.59</td>
</tr>
</tbody>
</table>

\[\text{Note:}\] Subscripts denote the Hf coordination of the constituent vacancies and superscripts indicate the overall charge of the unit cell.

Comparable to those calculated by Zhang et al despite their using the PBE functional [30].

The lowest energy di-vacancy structures for each of the vacancy coordination combinations is shown in figure 1. In each case the lowest energy structure is one in which the two vacancies are not screened from one another by a Hf cation and therefore, form the largest void possible within the structure.

Figure 2 shows that there is no correlation between the binding energies of the di-vacancies and the vacancy–vacancy separation, instead, the leading factors seem to be the vacancy configuration and coordination. The lowest energy \(D_{44}\) in particular does not consist of the closest vacuum pair. The di-vacancies containing a 3C vacancy site are lowest in energy. This is due to the closer packing of the 4C sites and so a \(D_{44}\) vacancy cannot create as big a void in the lattice.

The neutral oxygen tri-vacancies demonstrate behaviour similar to the di-vacancies and have generally small binding energies. However, all the energies are slightly more negative than those in the di-vacancies, showing that the addition of a third vacancy to the aggregate is more favourable than the second. This depends, however, on the overall structure of the aggregate as there is still a 0.36 eV variation of \(E_{\text{bind}}\) in the considered structures.

Figure 3 shows the tri-vacancy structure with the most negative \(E_{\text{bind},T}\) of approximately 0.45 eV. The vacancies in this tri-vacancy, like in the di-vacancy case, are not screen from one another by a Hf cation and form a trigonal structure and the largest void possible within the \(m\)-HfO\(_2\) lattice.

In the case of neutral tetra-vacancies, the trends shown by di- and tri-vacancies are further enhanced. Figure 4 shows one of the lowest energy neutral \(m\)-HfO\(_2\) tetra-vacancy structures, which has an \(E_{\text{bind}}\) of approximately 0.54 eV. Again this configuration forms a large void and the vacancy sites are not screened from one another by Hf cations. Other favourable tetra-vacancy configurations show similar characteristics whereby the constituent vacancies are largely unscreened from one another. Configurations forming a three-dimensional void have lower energies than those forming two-dimensional structures within a single sublattice. This suggests that, despite the layered structure of HfO\(_2\), there is no clear directionality to how the vacancy clusters will form.

The largest binding energy per vacancy is \(-0.11\), \(-0.15\) and \(-0.18\) eV for the di-, tri- and tetra-vacancy aggregates.
respectively. This implies that the attachment of vacancies to an existing aggregate may become more favourable as the aggregate grows. However, this is strongly dependent on the overall structure of the aggregate. It is also possible that a less favourable aggregate may re-organize into a lower energy structure allowing this energetic gain to continue.

4. Charged O di-vacancy

We calculated the neutral di-vacancy and its electronic structure using HSE06. The results are summarized in table 2. The binding energies of the neutral di-vacancies are comparable to those calculated using PW91 and PBE, varying by around 0.1 eV.

There are two occupied di-vacancy states in the middle of the band gap. The lower energy state is at 2.8 eV above the valence band whereas the second vacancy state appears at 3.4 eV above the valence band. Figure 5 shows the isosurface of the band-decomposed partial charge density associated with these vacancy states. Both states are a combination of Hf d-orbitals located inside both vacancies and can be described as two di-vacancy states. The higher and lower energy states (figures 5(a) and (b)) are antibonding and bonding, respectively, implying the di-vacancy has an overall formal bond order of zero, and may explain why the binding energy between these di-vacancies is so small.

In the case of \( D_{33}^- \) and \( D_{44}^- \), electrons localize on the Hf atoms between the two vacancy sites, and are distributed evenly between the vacancies. In the case of \( D_{34}^- \) the electrons localize onto the 4C vacancy site.

The highest occupied states which form when the di-vacancies are negatively charged lie between the di-vacancy states and the conduction band (CB). In \( D_{33}^- \) this state is singly occupied and is located 0.4 eV below the CB. \( D_{44}^- \) is the triplet state with two singly occupied states located 1 eV and 1.2 eV below the CB. In the lowest energy \( D_{34}^- \) the deeper state is 1.5 eV below the CB. This lower state is due to the deeper well created by the surrounding positive Hf cations, this creates a region of slightly positive charge, lowering the electronic levels and the overall energy of the structure.

This effect is also seen in the case of the lowest energy tetra-vacancy structure. In the PW91 calculations the CB is artificially lowered and covers up the negatively charged vacancy states, so that any extra electrons added to these systems delocalize. However, in the case of the lowest energy tetra-vacancy, the highest occupied states of the negatively charged vacancy are in the band gap. For the 1- and 3 cases a doublet state forms, for the 2- and 4 cases a singlet state forms. The lower energy level descends from 0.5 eV below the CB in the 1 case to 0.7 eV below the CB in the 4 case, while a higher energy state appears for the 3- and 4 cases at 0.2 eV below the CBM. These states do not appear in higher energy tetra-vacancy configurations. This is because in these low energy tetra-vacancies a very deep well is created by the surrounding Hf cations which can trap electrons. This deep well lowers the electronic states and lowers the overall energy of the aggregate.

To calculate the binding energies for negatively charged di-vacancies we assume that neutral O vacancies can trap one or two extra electrons and neutral and negative vacancies can diffuse in the vicinity of each other. The individual

### Table 2. Formation energies of oxygen vacancies and di-vacancies (eV) calculated using equation (2) and the HSE06 functional.

<table>
<thead>
<tr>
<th>Defect</th>
<th>( 2V_0^0 )</th>
<th>( V_0^0 + V_1^0 )</th>
<th>( V_1^0 + V_2^0 )</th>
<th>( V_2^0 + V_4^0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_{33}^- )</td>
<td>-0.16</td>
<td>0.13</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>( D_{44}^- )</td>
<td>0.13</td>
<td>0.04</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>( D_{34}^- )</td>
<td>0.06</td>
<td>0.38</td>
<td>-0.37</td>
<td></td>
</tr>
</tbody>
</table>

Note: \( D \) denotes a divacancy, subscripts denote the coordination of the constituent vacancies.
vacancies are calculated in separate cells in different charge states. When they are brought together, the electron density is re-distributed, forming occupied bonding and anti-bonding states described above. The binding energies of negatively charged di-vacancies are then calculated as the difference in total energies of an aggregate and individual vacancies and are shown in table 2. The binding energy varies depending on the constituent vacancies in the aggregate and the initial state at infinite separation. In the D\textsubscript{1h} case, the positive E\textsubscript{bind} indicates a slightly repulsive interaction. The only attractive vacancies are doubly negatively charged di-vacancies containing a 4C vacancy. These binding energies are generally small indicating charged vacancies do not interact strongly with one another.

We also attempted to calculate positively charged and doubly positively charged di-vacancies. However, every configuration where the vacancy–vacancy distance was within 4 Å was found to be unstable. In each case the vacancies would repel each other and spontaneously separate until they were sufficiently screened by at least one Hf cation.

5. Discussion and conclusions

Our results demonstrate that the aggregation of neutral oxygen vacancies in m-HfO\textsubscript{2} is feasible and that vacancy aggregates, if created, will be stable at room temperature. The binding energy per vacancy becomes larger as the aggregate grows, implying the addition of a vacancy to an aggregate larger than a tetra-vacancy may be even more favourable. However, this depends on the structure of the aggregate as a whole, so aggregate reorganizations may need to occur for this process to continue. The negative di-vacancy was also shown to be reasonably stable.

There is little data in the literature to compare the binding energies of the HfO\textsubscript{2} vacancy aggregates and those calculated in other oxides. It is interesting to note, however, that the binding energies of the di-vacancies are comparable to those of alkaline earth diatomic molecules. These diatomics have similar bonding and antibonding ordering of highest occupied molecular orbitals and a formal bond order of zero and, as a consequence, have small binding energies. In particular, the properties of Be\textsubscript{2} were measured using laser-induced fluorescence [39]. It was found that Be\textsubscript{2} has an E\textsubscript{bind} of \(-0.10\) eV and bond length of 2.45 Å which is very similar binding energy and interatomic distance to the calculated HfO\textsubscript{2} di-vacancies.

At low vacancy concentrations, the energy gain due to vacancy interaction does not necessarily result in the formation of stable vacancy aggregates because an entropic contribution to the free energy due to the number of ways in which isolated vacancies are accommodated can outperform the binding energy [40]. However, the concentration of oxygen vacancies in HfO\textsubscript{2} films in CMOS devices usually exceeds \(10^{18}\) cm\textsuperscript{-3} [41] and needs to be greater than 4% for conductive filaments to form in RRAM devices, [42] which are in the focus of this work. High oxygen deficiency in these systems is often induced by the oxygen gettering action of the Ti or Hf layer, which forms a highly sub-stoichiometric region near the top interface with a metal electrode [43–45].

The kinetics of neutral oxygen vacancy aggregation has been studied only in MgO [46]. This study has demonstrated that the defect interaction energy is the key factor in aggregate formation and that the vacancy binding energies as small as 0.03 eV are sufficient for aggregates to form once the vacancy diffusion activation energy (about 3.5 eV) can be overcome. In MgO stable aggregates are formed within several minutes when the vacancy concentration exceeds \(10^{18}\) cm\textsuperscript{-3} and the temperature is above 1500 K [46]. Based on the calculated neutral vacancy diffusion barriers of 2.4 eV in HfO\textsubscript{2} [28] film temperatures of the order of 1000 K are required to accelerate the vacancy aggregation. Such high temperatures, however, are likely to lead to breaking of the loosely bound aggregates and stimulate aggregation at lower energy sinks, such as grain boundaries and dislocations.

Within the context of a RRAM conductive filament forming process, it is expected that the oxide will be under bias and it is therefore likely that some vacancies will be neutral or positively charged. However, high vacancy diffusion barriers in the neutral charge state and small binding energies suggest that the process of aggregation of these vacancies via thermally activated diffusion is slow and inefficient under the voltage/temperature conditions of practical interest. Positive vacancies can diffuse more efficiently and may get converted into a neutral state after merging with other already aggregated vacancies—such a cluster can capture injected electrons very effectively, and it shares trapped electrons between all linked nearby vacancies. An alternative mechanism may involve generation of new vacancies in the vicinity of pre-existing vacancies. The mechanism of this process is considered in a separate study [47].

Acknowledgments

This work was supported by EPSRC and the COST Action CM1104. Via our membership of the UK’s HPC Materials Chemistry Consortium, which is funded by EPSRC (EP/L000202), this work made use of the facilities of HECToR and ARCHER, the UK’s national high-performance computing service, which is funded by the Office of Science and Technology through the EPSRC’s High End Computing Programme. This work also gratefully acknowledges access to the Chinook supercomputer at EMSL, a national scientific user facility sponsored by the Department of Energy’s Office of Biological and Environmental Research and Located at Pacific Northwest National Laboratory. The authors gratefully acknowledge Thomas Durrant and Samuel Murphy for valuable discussions.

References

[40] Bersuker G et al 2006 J. Appl. Phys. 100 094108