Physical Insights into the Transport Properties of RRAMs Based on Transition Metal Oxides

Toufik Sadi[1], Oves Badami[2], Vihar Georgiev[2], Jie Ding[3] and Asen Asenov[2]
[1]*Engineered Nanosystems Group, School of Science, Aalto University, PO Box 12200, 00076 AALTO, Finland*
[2]*School of Engineering, Electronic and Nanoscale Engineering, University of Glasgow, Glasgow G12 8LT, Scotland, UK*
[3] *College of Electrical and Power Engineering, Taiyuan University of Technology, 030024 China.*
toufik.sadi@aalto.fi

***Abstract*—Nowadays, resistive random-access memories (RRAMs) are widely considered as the next generation of non-volatile memory devices. Here, we employ a physics-based multi-scale kinetic Monte Carlo simulator to study the microscopic transport properties and characteristics of promising RRAM devices based on transition metal oxides, specifically hafnium oxide (HfO_x) based structures. The simulator handles self-consistently electronic charge and thermal transport in the three-dimensional (3D) space, allowing the realistic study of the dynamics of conductive filaments responsible for switching. By presenting insightful results, we argue that using a simulator of a 3D nature, accounting for self-consistent fields and self-heating, is necessary for understanding switching in RRAMs. As an example, we look into the unipolar operation mode, by showing how only the correct inclusion of self-heating allows the proper reconstruction of the switching behaviour. The simulation framework is well-suited for exploring the operation and reliability of RRAMs, providing a reliable computational tool for the optimization of existing device technologies and the path finding and development of new RRAM options.**

***Index Terms*—Kinetic Monte Carlo (KMC), resistive random-access memories (RRAMs), multi-scale models, transport phenomena.**

I. Introduction

For several decades, the semiconductor industry experienced a strong growth, thanks to device downscaling, leading to increased functionality and performance. However, as this miniaturization trend is maintained and Moore's law is approaching its limits, undesirable effects, such as excessive power dissipation and self-heating, hinder the performance of microchips. This has forced the industry to re-evaluate the von-Neumann architecture by moving towards in-memory computing. In this paradigm shift, devices based on resistive random access memories (RRAMs) are expected to play an important role, which necessitates the development of advanced physics-based simulators to understand better RRAM operation and provide optimal device designs.

The idea of memristor devices, such as RRAMs, was put forward theoretically almost 50 years ago [1]. Since their experimental demonstration 11 years ago [2], the interest in RRAMs has been increasing exponentially [3]–[5], being considered as the next generation of non-volatile memories. Indeed, the 'International Technology Roadmap for Semiconductors' (ITRS) cites a multitude of incentives for developing RRAMs, such as low cost and power dissipation, high endurance and three-dimensional (3D) crossbars integration [6]. The applications of RRAMs are also innumerable, ranging from high-density memories and novel processor architectures to neuromorphic computing and artificial intelligence [4].

In this work, we analyze the switching behaviour and certain interesting features of RRAM structures based on hafnium oxide (HfO_x), using a kinetic Monte Carlo (KMC) simulation framework. In Sec. II, we discuss the main attributes of the simulator and describe its original aspects. In Sec. III, we discuss the basic switching behaviour of the simulated devices, and highlight the importance of including coupled electro-thermal transport to capture correctly switching.

The research is funded by the EPSRC (UK), under grants no. EP/S000224/1 and no. EP/S001131/1.

II. Simulation Methodology

Most previous work on the simulation of RRAMs relied mostly on phenomenological models, such as the resistor breaker network [5], [7], which do not account accurately for self-heating and self-consistent fields. In addition, most existing models use two-dimensional (2D) approximations [9], [10] which may produce less reliable and insightful results [11]. The 3D KMC simulator used in this work is capable of providing a complete picture of particle dynamics in oxide based RRAMs. It incorporates several features that distinguish it from established phenomenological models [5], [9], [10], as discussed in Ref. [3].

We employ an in-house 3D device simulator, which has been previously used for gaining insight into the operation of SiO_x structures [3], [8], to study HfO_x-based RRAMs, a widely used transition metal oxide (TMO) in memristor technology. Hafnia is highly suitable for high-density CMOS integration due to their high dielectric constants. Figure 1(a) illustrates the simulation framework. Unlike previously used 2D and phenomenological models [5], [9], [10], our simulator uses a powerful combination of tools, describing accurately electron-ion interactions and reconstructing realistically the electroforming and rupture of conductive filaments in the 3D real space. It couples, in a self-consistent manner, electron and oxygen ion KMC trajectory simulations to the electric field and temperature distributions determined from the solution of Poisson's and the time-dependent heat diffusion equations.

978-1-7281-0940-4/19/$31.00 ©2019 IEEE

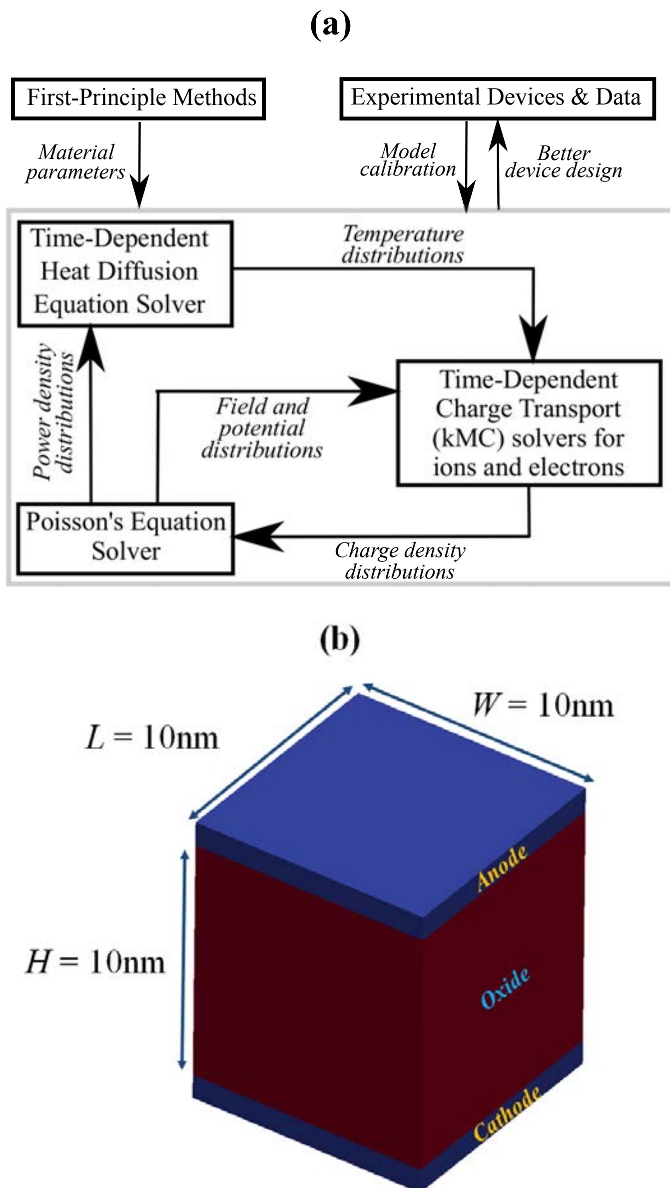

Fig. 1. (a) The simulation framework, coupling the KMC description of charge transport to the local temperature and electric field distributions in the oxide. Relevant material parameters, e.g. the activation energy, are obtained using first-principle methods. In general, the simulator is calibrated with experiments for enhanced predictivity power. (b) The simulated two-terminal RRAM structure, consisting of an oxide (HfO_x in this case) volume (thickness $T = 10$nm) sandwiched between the cathode and the anode. Realistic experimental structures may have electrode areas as large as $100\mu m \times 100\mu m$ [14], but it is sufficient to limit our study to a small contact area ($L \times W$=10nm×10nm here), to minimize computational cost.

The dynamic nature of the vacancy formation and annihilation, and electron trapping is considered accurately, as discussed rigorously in Refs. [3], [8], [12]. The ion and vacancy time-dependent dynamics (drift, diffusion, generation and recombination) are also modeled carefully, as discussed in Refs. [3], [8]. The effect of all the dominant electron transport mechanisms are carefully considered, including trap-assisted tunneling, trap-to-trap tunneling, Fowler-Nordheim tunneling, Poole-Frenkel emission, and direct tunneling mechanisms [3]. Electron and oxygen ion movements as well as ion-vacancy generation and recombination events are tracked down in time via the stochastic KMC algorithm, providing a realistic picture of the interplay between electrons, ions and vacancies as influenced by the evolving local electrostatic and temperature effects. More details about the simulation methodology and the included physical processes are given in Ref. [3].

III. Results and Discussion

A. *Simulated Structure and Practical Considerations*

Here, we illustrate how 3D electrothermal modelling, accounting for self-heating effects and self-consistent fields, as neglected in other KMC simulation models (see for example Ref. [9]), provides a deep physical insight into RRAM switching. The studied devices simply consist of the oxide (HfO_x in this case) layer (thickness$\sim$ 10nm) located between two electrodes (the cathode and the anode), as illustrated in Fig. 1(b). As discussed extensively in literature, the memristive behaviour of oxide-based RRAMs is a direct consequence of the forming and destruction of conductive filaments, which are formed by direct electrical conductive paths between the cathode and the anode [3], [8], [13]. These filaments are created by the generation of oxygen ion-vacancy pairs, whose rates and transport are in general governed by the electric field (potential) and temperature distribution within the oxide. While the experimental structures used to validate the simulator may have electrode areas as large as $100\mu m \times 100\mu m$ [14], the simulations can be limited to a small contact area ($L \times W$=10nm×10nm here), which can represent a region incorporating e.g. a grain boundary. This is common practice in Monte Carlo modeling methods, aiming to reduce the computational cost while allowing reasonable numerical simulation accuracy [3].

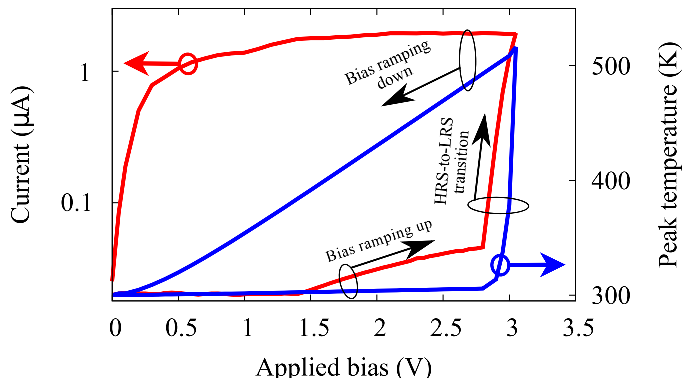

Fig. 2. The $I-V$ and peak temperature curves, as bias is ramped up towards the CF forming and then lowered down to 0V.

B. *Basic Characteristics*

Figure 2 shows the $I - V$ and peak temperature curves obtained during the electroforming of the conductive filament, using an electric current compliance limit of 2μA. Figure 3 shows the distribution of the oxygen vacancies created, and

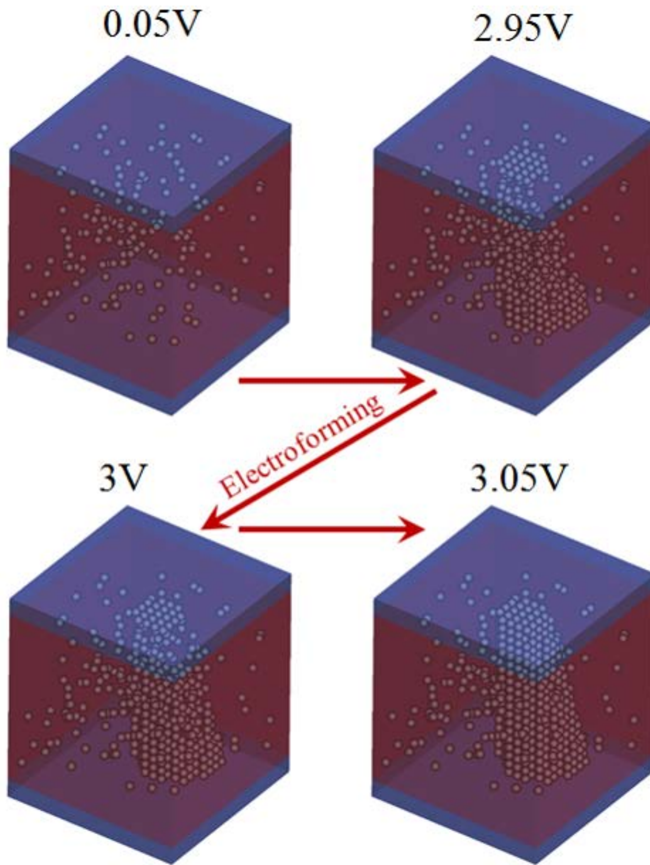

Fig. 3. The generated vacancy distributions as bias is ramped up towards forming.

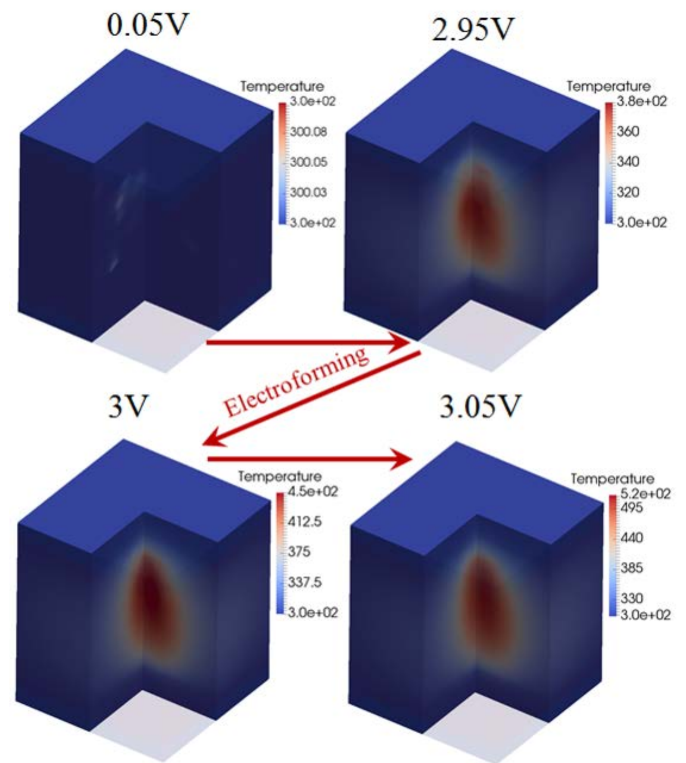

Fig. 4. The temperature distributions as bias is ramped up towards forming.

Fig. 4 shows the corresponding local temperature distributions, as bias is increased and the CF is gradually created. Figure 2 illustrates the expected memristive characteristics of the RRAM device. Figure 3 highlights the three-dimensional nature of conductive filaments. At biases below 2.5V, very few vacancies are generated. As bias is increased, more vacancies are generated and filament seeds start to appear (e.g. at 2.9V). Such seeds start to grow as bias is further increased. At around 3V, an accelerated generation of vacancies occurs, leading to the creation of a full conductive filament, linking both electrodes; at this condition, percolation paths are created, as an abrupt jump in the device current is observed.

Figure 4 shows how the oxide temperature reaches values beyond 500K during the electroforming process. The peak temperature tends to occur within the filament volume, where the combination of the elevated current densities and the low oxide thermal conductivity can lead to such very high values. In general, the elevated local temperatures, resulting mainly from Joule heating, affect significantly the device behavior, as they can boost the probability of vacancy generation and ion hopping, but also electron transport via trap-assisted tunneling and other relevant mechanisms [3], [8], [12], [15].

C. Switching and Self-Heating

The critical role of device self-heating can be illustrated by looking into the reset process considering the unipolar RRAM switching mode. Unlike the bipolar switching mode, where the bias is further reduced, from 0V to negative values, to realize the reset process (after the CF is formed) [3], this regime is achieved in the unipolar mode by increasing the bias from 0V to positive values. Figure 5 shows the oxygen vacancy and temperature distributions just before and after the filament is ruptured, at a bias voltage of around 2.4V, during the reset process for a unipolar mode. It can be seen that the fully-conductive filament is broken near the top surface (near the anode), resulting in the device switching from a high-current (ON) low-resistance state (LRS) to a low current (OFF) high-resistance state (HRS). As expected, the ON-to-OFF transition also results in the peak temperature dropping considerably, as current densities are reduced. The RRAM device experiences such transition because the oxygen ions in the anode contact move back to the oxide volume, recombining with the nearby vacancies and breaking the CF. This phenomenon occurs thanks to ion diffusion, which is facilitated by the elevated temperatures in the oxide just before the transition. It has been verified that such transition cannot be easily reconstructed, using simulations, in the unipolar RRAM operation mode without the inclusion of thermal effects. Self-heating in the bipolar switching mode is not of critical importance, as ions easily drift back to the oxide as a certain reset bias (field) is

reached.

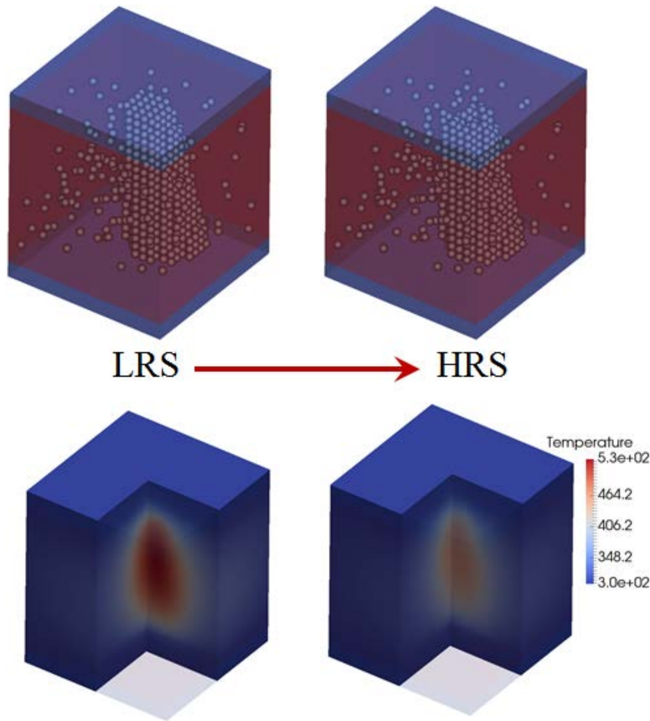

Fig. 5. The vacancy and temperature distributions, in the reset process for a unipolar mode, just before and after the CF is ruptured (bias $\sim$ 2.4V).

IV. Outlook

We applied a kinetic Monte Carlo simulation framework to study the operation and switching physics of hafnia-based RRAM devices. We discussed the need for using 3D models accounting for self-heating and self-consistent fields to capture carefully the expected switching behaviour. In addition to exploring RRAM physics and operation, the model can be applied to investigate reliability issues and bottlenecks of RRAM technology development, such as the design of reliable oxygen storage or supply systems to increase the RRAM reliability and endurance. Physics-based modelling is only the initial step in designing high-performance RRAMs. Data from the physical simulation of experimental devices will contribute to the development of analytical (compact) models, which will be integrated into circuit simulators to design cross-bar circuits for interesting applications.

References

[1] L. Chua, "Memristor-the missing circuit element," IEEE Trans. Circuit Theory, vol. 18, pp. 507–519, 1971.

[2] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," Nature, 453, pp. 80–83, 2008.

[3] T. Sadi *et al.*, "Investigation of resistance switching in SiOx RRAM cells using a 3D multiscale kinetic Monte Carlo simulator," Jour. Phys.: Condens. Matter, 30, pp. 084005, 2018.

[4] A. Mehonic and A. J. Kenyon, "Emulating the electrical activity of the neuron using a silicon oxide RRAM cell," Front. Neurosci., 10, pp. 1–10, 2016.

[5] S. C. Chae *et al.*, "Random circuit breaker network model for unipolar resistance switching," Adv. Mater., 20, pp. 1154–1159, 2008.

[6] The ITRS Report 2013, http://www.itrs2.net/2013-itrs.html. Accessed 5 Jun 2018.

[7] S. Brivio and S. Spiga, "Stochastic circuit breaker network model for bipolar resistance switching memories," J. Comput. Electron., vol. 16, pp. 1154–1166, 2017.

[8] T. Sadi *et al.*, "Advanced Physical Modeling of SiO_x Resistive Random Access Memories," In Proc. SISPAD, pp. 149–152, 2016.

[9] S. Yu, X. Guan, and H.-S. P. Wong, "On the stochastic nature of resistive switching in metal oxide RRAM: physical modeling, Monte Carlo simulation, and experimental characterization," In: 2011 IEEE Int. Electron Devices Meeting (IEDM), p. 17.3.1., 2011.

[10] S. Kim *et al.*, "Physical electro-thermal model of resistive switching in bi-layered resistance-change memory," Sci. Rep., 3, p. 1680, 2013.

[11] T. Sadi and A. Asenov, "Microscopic KMC Modeling of Oxide RRAMs," In: Nikolov G., Kolkovska N., Georgiev K. (eds) Numerical Methods and Applications. NMA 2018. Lecture Notes in Computer Science, vol. 11189. Springer, Cham, pp. 290–297, 2019.

[12] L. Vandelli *et al.*, "A physical model of the temperature dependence of the current through SiO_2/HfO_2 stacks," IEEE Trans. Electron Devices, 58, pp. 2878–2887, 2011.

[13] M. Buckwell, L. Montesi, S. Hudziak, A. Mehonic, A. J. Kenyon, "Conductance tomography of conductive filaments in intrinsic silicon-rich silica RRAM," Nanoscale, vol. 7, pp. 18030–18035, 2015.

[14] A. Mehonic *et al.*, "Resistive switching in silicon sub-oxide films," J. Appl. Phys., vol. 111, p. 074507, 2012.

[15] G. C. Jegert, "Modeling of leakage currents in high-k dielectrics," PhD Dissertation, Tech. Univ. Munich, Germany, 2011.