Coupled Ionic and Electronic Transport Model of Thin-Film Semiconductor Memristive Behavior

Dmitri B. Strukov,* Julien L. Borghetti, and R. Stanley Williams

The memristor, the fourth passive circuit element, was predicted theoretically nearly 40 years ago, but we just recently demonstrated both an intentional material system and an analytical model that exhibited the properties of such a device. Here we provide a more physical model based on numerical solutions of coupled drift-diffusion equations for electrons and ions with appropriate boundary conditions. We simulate the dynamics of a two-terminal memristive device based on a semiconductor thin film with mobile dopants that are partially compensated by a small amount of immobile acceptors. We examine the mobile ion distributions, zero-bias potentials, and current–voltage characteristics of the model for both steady-state bias conditions and for dynamical switching to obtain physical insight into the transport processes responsible for memristive behavior in semiconductor films.

1. Introduction

In 1971, Leon Chua proposed the existence of a fourth fundamental electronic circuit element, the memristor, to complement and complete the set of previously known passive devices (capacitor, resistor, and inductor).[1] The primary characteristic of this element that made it remarkable was its “pinched hysteresis loop”, that is, when energized by a sinusoidal voltage, the resulting current–voltage (i–v) characteristic was a Lissajous curve that could not be duplicated with any network of the other fundamental passive devices. In 1976, Chua and Kang[2] generalized this concept to a specialized set of dynamical devices they called memristive systems, the primary characteristic of which was again the pinched hysteresis loop, or the fact that such devices cannot store energy and therefore, the current flowing through them must be zero when the applied voltage is zero. However, memristors and memristive devices “remember” the total charge that flows through them by changing their resistance. These properties arise from the defining equations for memristive systems,

\[ v = R(w, i) i \]  
\[ \frac{dw}{dt} = f(w, i) \]

in which \( v \) is the voltage, \( i \) the current, \( R \) the instantaneous resistance, and \( w \) is a state variable (or variables), and for “pure” memristors neither \( R \) nor \( f \) are explicit functions of \( i \).

These devices were primarily intellectual curiosities until this year, when we showed that both memristance and memristive behavior arise naturally in thin film semiconductor systems for which the equations of motion for current carriers (electrons and holes) and for mobile charged dopants are coupled in the presence of an applied field.[3] In that paper, we presented a simple analytical model to show how variable resistors in series, one representing a region of highly doped semiconductor and the other undoped material, could be constructed to satisfy Equations 1a and b. The state variable \( w \) corresponded to the position of a sharp dividing line between the doped and undoped semiconductor, and was bounded between two limits, 0 and \( L \), that corresponded to the positions

**Keywords:**
- ion transport
- memristive behavior
- modeling
- resistive switching
- thin films
of the metal contacts at either side of a semiconductor film. For the first time, we were able to write an equation expressing the memristance, \( M(q) \), of a device in terms of its physical properties and geometrical structure, which revealed that the magnitude of the charge \( q \)-dependent term, which is responsible for the hysteresis, is inversely proportional to the thickness of the device \( L \). Thus, memristance and memristive behavior are much stronger at the nanometer scale than at the micrometer scale and have emerged in importance with the shrinking of electronic device feature sizes. Another realization to come out of our work was the importance of the boundary conditions as \( w \) approached 0 or \( L \) on the \( i-v \) characteristics of a device; a semiconductor memristor driven hard or long enough will become a memristive system when \( w \) approaches a boundary, and different boundary conditions produce dramatically different \( i-v \) curves that replicate many anomalous results that have been reported in the past.\(^{[3]} \)

In this paper, we provide a more physical model of memristive behavior based on numerical solutions of coupled drift-diffusion equations for electrons, holes, and ions to simulate the dynamics of two-terminal devices based on a metal/semiconductor/metal structure. The model contains enough physics to enable us to see details of the dynamics but is still simple enough to allow us to obtain significant insight into the behavior. It can be applied to materials that act as solid-state electrolytes as well as nonstoichiometric compounds with mobile defects, for example, oxygen vacancies. The results of our present investigation validate some of the assumptions of our analytical model and provide further insight into such issues as the nature of the boundary conditions and other limitations on the motion of dopants that were not previously evident. Most importantly, the model provides a foundation from which we can construct more quantitative simulations of real devices.

Previous research on the modeling of mixed ionic/electronic conductors focused on their stationary or/and static \( i-v \) characteristics.\(^{[4–7]} \) The results of dynamic simulations of a system with Schottky-type contacts have been discussed,\(^{[8]} \) but details of the simulation approach have not yet been published. Detailed transient dynamics have been examined,\(^{[9]} \) but only the displacement electronic current was considered. More recently, the impact of oxygen vacancy drift on the drain current of a field effect transistor in the context of thin film that contains charged mobile electronic conductors focused on their stationary or/and static \( i-v \) characteristics.\(^{[10]} \) The model geometry was significantly different from the one considered in this paper, that is, it was a three-terminal device in which ion drift was mostly perpendicular to the electronic current.

2. Model

We consider a 1D model of the device with a semiconductor thin film that contains charged mobile \( n \)-type dopants with a concentration distribution \( N_D(x) \) confined by electrodes at \( x = 0 \) and \( x = L \) (Figure 1). The dimensions of the device along the \( y \)- and \( z \)-directions are assumed to be much larger than \( L \), which is a reasonable approximation for thin film structures even if the active region is the gap between a conducting channel in a much thicker layer of semiconductor and one of the electrodes. The active layer is partially compensated by a small amount of immobile and uniformly distributed \( p \)-type dopants with concentration \( N_A \leq N_D \), the average concentration of mobile dopants \( \langle N^\mu_D(x) = \int N_D(x)/L \, dx \rangle \). In the discussion that follows, all quantities that are explicitly denoted as functions of \( x \) are also functions of time and will be examined after certain time intervals have elapsed.

For simplicity, we assume that both donors and acceptors are shallow, and thus their energy levels coincide with the corresponding conduction \( E_C \) and valence \( E_V \) band edges. In this case, one can neglect thermal generation and recombination currents for relatively large band gap materials. Moreover, the goal of this paper is to study the dynamics on the ion-drift time scale. Since the electron-lattice relaxation is much
faster, electron \( n(x) \) and hole \( p(x) \) concentrations can be approximated with Fermi–Dirac statistics, that is, with the separate quasi-Fermi potentials \( \phi_n(x) \) and \( \phi_p(x) \), respectively.\(^{[11]} \) Thus, the steady-state currents from the drift-diffusion approximation for electrons and holes are determined from

\[
\nabla \cdot (-en(x)\mu_n \nabla \phi_n(x)) = 0
\]

(2)

and

\[
\nabla \cdot (ep(x)\mu_p \nabla \phi_p(x)) = 0
\]

(3)

while the Poisson equation for the active layer with uniform permittivity \( \varepsilon_0 \) is

\[
- \varepsilon \varepsilon_0 \Delta \phi(x) = e(p(x) - n(x)) + D_D(x)N_D(x) - f_A(x)N_A
\]

(4)

Here, \( e \) is the unit charge, \( \phi(x) \) the electrostatic potential, and \( \mu_n(\mu_p) \) is the electron (hole) mobility, which is assumed to be independent of the field. \( f_D(x) \) and \( f_A(x) \) are ionization factors for donors and acceptors, correspondingly, which are found from the population statistics\(^{[11]} \) and typically equal to unity for the considered simulation parameters.

The non-equilibrium mobile ion distribution and ion flux \( J_{ION}(x) \) can also be found from drift-diffusion theory (in this paper we are neglecting nonlinear drift in high electric fields\(^{[12]} \)) and the continuity equation, that is, for singly charged positive ions,

\[
J_{ION}(x) = -eD_i \nabla N_D(x) - eN_D(x)u_i \nabla \phi(x)
\]

(5)

and

\[
e \frac{\partial N_D(x)}{\partial t} = -\nabla \cdot J_{ION}(x)
\]

(6)

where \( D_i \) and \( u_i \) are the ion diffusion constant and mobility, respectively, which are related via the Einstein–Nernst equation.

In this study we consider only bulk-limited transport, and the interfaces are assumed to be purely ohmic for electrons. This, for example, might correspond to either metal electrodes with adjacent \( \Delta \)-doped semiconductor interfacial layers or heavily doped semiconductor electrodes with a band gap similar to that of the transport layer. The electron and hole concentrations are fixed to their equilibrium levels defined by the Fermi potential at the interface, such that for an applied bias \( v \) the boundary conditions are

\[
\phi_n(0) = \phi_p(0)
\]

(7)

and

\[
\phi_n(L) = \phi_p(L) = \phi_n(0) + v
\]

(8)

The interfaces are assumed to be trap free and completely blocking for mobile ions, so that the total number of mobile dopants, \( LN_D^+ \) is constant, that is,

\[
J_{ION}(0) = J_{ION}(L) = 0
\]

(9)

In this study, we neglect Joule heating and explicit temperature effects.

3. Simulation Results

Equations 2–6 with boundary conditions 7–9 are solved numerically using an iterative procedure to achieve self-consistency. More specifically, starting from initially defined mobile \( N_D(x) \) and fixed \( N_A \) ion distributions, the potentials \( \phi(x) \), \( \phi_n(x) \), and \( \phi_p(x) \) are determined from Equations 2–4 using a modified Newton–Raphson method.\(^{[13]} \) After that, a new \( N_D(x) \) is calculated using a backward difference approximation for the time derivative with the time step limited by the Courant condition.\(^{[14]} \) These two steps are repeated throughout the simulation until some specified goal is reached. Fast \( i-v \) measurements to characterize the resistance state of the system at a particular time are simulated by freezing \( N_D(x) \) and calculating the current as a function of voltage for that configuration of mobile dopants. All stationary profiles are calculated by starting with a uniform concentration, that is, \( N_D(x) = N_D^0 \), and then relaxing the system by simulating sufficiently large times so that no significant changes occur between iterations. As a sanity check, we compare the relaxed profile to one calculated from a stationary model with \( J_{ION} = 0 \) in Equation 5 to make sure that they are consistent.

We first simulate the ground state of the system, with zero voltage applied, in order to examine the equilibrium distribution \( N_D(x) \). An analytical solution of the Poisson–Boltzmann equation predicts that a 1D system with charged mobile ions confined between electrodes with compensating charge will settle into a U-shaped distribution.\(^{[15]} \) This intuitive result comes from the condition that the electrostatic attraction of ions to the image charge on both electrodes is the strongest near the interfaces, and hence a large ionic concentration gradient is required to compensate it. The two decreasing dopant concentrations from each interface meet in the center of the film with a slope of zero, which produces the U-shape.

When the semiconductor band gap and Fermi level pinning at the electrodes are included in the model, the simple analytical treatment is no longer appropriate. The equilibrium profile from the numerical calculation still has an approximate U-shape, as shown in Figure 1a, but depending on the specific conditions the \( U \) may be flattened or even inverted. For example, when the depletion width is smaller than the device length \( L \), only ions near the interfaces will be affected and thus the profile is flat in the middle of the device. The orientation of the \( U \), that is, either normal or inverted, depends on the alignment of the Fermi potentials between the electrodes and the semiconductor film. For the simulations reported here, the electrode Fermi potential was chosen such that it corresponds to the potential of appropriate semiconductor layers at each
electrode doped with $N_{D0}$ donors and $N_A$ acceptors (Figure 1a). There are three qualitatively different equilibrium profiles, given by the conditions $N_D > N_{D0}, N_D = N_{D0}$, and $N_D < N_{D0}$. The simulation results are reported for the latter, the most representative case with $N_D = 0.1N_{D0} = 10N_A$. In this case the equilibrium ion distribution has an upside-down U-shape with a flattened center (Figure 1a).

For these concentrations of mobile and fixed ions, we also show (with colored lines) the quasi-equilibrium characteristics for the ion distribution (Figure 1a) and the zero-bias potential (Figure 1b) resulting from a series of positive voltages applied to the right-hand side of the device ($x = L$) after holding each bias for a long enough time to reach steady state, as well as the $i$–$v$ curves used to characterize each state. These steady-state characteristics are attained when $J_{ION}(x) = 0$, or from Equation 5, $N_D(x)\mu_i\nabla \varphi(x) = -D_i\nabla N_D(x)$, that is, when the field and the concentration gradient are oppositely directed and the drift is exactly compensated by the diffusion. The initial application of an external voltage causes positive mobile ions to drift and results in effectively complete mobile ion depletion from one side of the device (the right-hand side for the applied bias at $x = L$, represented in Figure 1). Once the concentration of mobile donors in this region falls below that of the fixed acceptors, the device doping profile goes from $n^–n–n^+$ to $n^–n–p–n^+$. The $p$-type region produces a potential barrier for electrons (Figure 1b) with the maximum height mostly depending on the particular choice of $N_A$ and the built-in voltage of the induced $p–n^+$ junction, which for the parameters chosen is close to the bandgap value $E_G/e$. To induce a significant barrier height modulation, we chose $N_A$ such that the depletion width associated with it was half of the device length $L$.

The device has only one equilibrium state, for which the $i$–$v$ curve is essentially linear and has the lowest resistance possible for the system. This is defined as the ON state (Figure 1c) and represents one of the boundaries for the resistance variation of the device. On the other hand, device states formed by increasing the bias voltage have increasingly nonlinear (and slightly asymmetric) $i$–$v$ curves with correspondingly increasing values of the zero-bias resistance (Figure 1c). There exists one limiting OFF state corresponding to the $i$–$v$ with the highest resistance possible for a device with given parameters, which is the other boundary for the resistance. This state can in principle be reached by applying a large positive external bias voltage so that all the mobile ions are pushed into a narrow region near one interface of the device and the potential barrier height reaches its maximum value $E_G/e$. Applying a larger bias voltage will not increase the resistance of the device. In practice, we do not attempt to simulate a system for which the applied voltage might yield an unphysically high concentration of mobile dopants within the present model.

Figure 2 illustrates the dynamical switching properties of our model. Initially, we apply a fixed positive voltage $v_+$ (we choose $v_+ = 120v_0$, such that from Figure 1a the maximum concentration $N_D(x)$ is less than $N_{D0}$ but the resistance modulation is still significant) that was only capable of turning
the device partially OFF to the right hand side of the device \((x = L)\), and we then determined the mobile dopant distribution, the zero-bias potential, and the electric field, along with corresponding \(i-v\) sweeps to characterize each state at various time intervals, as shown in Figure 2a–c, respectively. We see that the dopant distribution was pushed to the left (toward \(x = 0\)) with an extremely abrupt front between low and high ion concentrations. We define an effective vertical boundary for a particular time by the point \(x = w\), such that \(N_D(w) = N_A\) and plot \(w/L\) versus time in the inset of Figure 2a; the boundary initially moves rapidly, but decelerates significantly as it approaches the steady-state limit at \(w/L \approx 0.58\) for \(v = v_+\). This behavior can be understood from Figure 2c by analyzing the evolution of the electric field at \(w\); the field is essentially constant for \(x < w\) and equal to \(E(w)\), and is nearly linear for \(x > w\) with a slope \(-eN_A/(\epsilon \epsilon_0)\) since \(N_D(x)\) is very small, with the result that \(E(w) \approx v_+/L - e(L-w)^2N_A/(2\epsilon \epsilon_0)\). The rate of change in \(w\) is related to \(E(w)\) via the simple drift law \(dw/dt = -\mu_i E(w)\), which can be integrated by \(w(t) \propto \tanh[\lambda t]\), that is, exponential deceleration. The movement of the boundary causes the growth of the potential barrier in the gap (which is approximately \(eN_A(L-w)^2/(2\epsilon \epsilon_0)\)) by exposing more of the fixed acceptors. Consequently, the electrical resistance increases until the steady-state condition is reached.

Next, the bias voltage polarity is reversed (Figure 2f–j), with a negative voltage \((v = -v_+)\) applied at \(x = L\). As can be seen from Figure 2f for different time intervals, the ion concentration moves back toward its initial equilibrium distribution. However, this evolution of the distribution for switching back to the ON state is not the time-reversed image of the OFF-switching motion. The speed of the distribution front (defined as before and shown in the inset of Figure 2f) between high and low concentrations is much faster for the ON-switching case, and a new front appears at the left-hand side of the device, near \(x = 0\), as the negative bias at \(x = L\) attracts ions that had been at the opposite metal contact toward the center of the device. By applying the negative bias voltage for a long enough time, the device could be switched OFF again by attracting the dopants away from the left-hand contact and thus forming a potential barrier at that side; the beginning of this process can be observed in Figure 2f–h. It is unclear if the boundary speed decelerates as it approaches the ON state since the dynamics are much more complex than for the transition toward the OFF state. Although the internal state of the device is significantly different from the equilibrium state, as shown by the final plots of both \(N_D(x)\) and \(\varphi(x)\) in Figure 2f and g, respectively, the resistance of the device in the final state is only slightly larger than that of the equilibrium state.

When a sinusoidal voltage is applied to the device, as shown in Figure 3, it displays the pinched hysteresis loop of a memristor or memristive device. However, if we plot the integral of the voltage as a function of the integral of the current density, as in the inset of Figure 3, we see that the result is not a single curve, as it must be for a memristor,[1] and thus this model more properly represents a memristive device. This is because the rate of the ON and OFF transitions is different, and thus the hysteresis loops in the positive and negative quadrants of the \(i-v\) plots are not symmetric. Figure 3 also shows the results of the analytical model introduced in Reference [3], that is, using the equivalent circuit shown below panels a and f in Figure 2:

\[
v(t) = \left(\frac{R_{\text{ON}} w(t)}{L} + R_{\text{OFF}} \left(1 - \frac{w(t)}{L}\right)\right) i(t) \tag{10a}\]

and

\[
dw(t)/dt = -\mu_i R_{\text{ON}} i(t)/L \tag{10b}\]

For this calculation, we use the same sinusoidal external bias, \(\mu_i\), and \(R_{\text{ON}}\) as used in the numerical simulations, but the value of \(R_{\text{OFF}} = 5R_{\text{ON}}\) to avoid hard boundary conditions. The simple analytical model is a memristor, with the charge a single-valued function of the flux as shown in the inset of Figure 3. This comparison clearly reveals the two major differences between the numerical model for a semiconductor thin film memristive device and a pure memristor: the asymmetry in the hysteresis loops caused by the different rates of ON and OFF switching and the highly nonlinear \(i-v\) characteristic of the OFF state of the device. (Note that in the pure memristor model the ionic diffusion component was neglected[3]).

4. Discussion and Summary

The model presented here with drift-diffusion equations for current carriers (Equations 2 and 3) and charged mobile
dopants (Equations 5 and 6), coupled by the Poisson equation (Equation 4), both justifies and significantly extends the earlier analytical model for memristive behavior in a metal-semiconductor-film/metal device with mobile dopants.\textsuperscript{[3]} The observation of the abrupt front between the high and low concentration regions of the mobile dopant distribution is a partial confirmation of the central assumption of the earlier model. However, we see that in fact there is not a single state variable that describes the dynamical evolution of the resistance of the system, but rather a state function that evolves in time, which can be taken to be either the mobile dopant concentration distribution \(N_D(x)\) or the potential \(v(x)\) for \(0 \leq x \leq L\) in the semiconductor. We also see that the boundary conditions for the potential and the ion flux (Equation 7–9) lead to a set of soft boundary conditions for the resistance states of the system, that is, the front retreats from the limiting ON with a constant speed but decelerates significantly when approaching the OFF state.

The present model also makes several predictions about the behavior of a thin film semiconductor memristive device. First, the ON and nearby low-resistance states have ohmic \(i-v\) characteristics, but as the zero-bias resistance of the device increases toward the OFF state, the \(i-v\) curves become increasingly nonlinear because of the \(n-p-n\) potential barrier that develops in the semiconductor. Second, for a voltage bias polarity that increases the resistance state of the device, the ultimate OFF state may be approached in a series of steps. The resistance will increase with time for a particular constant applied voltage magnitude until it saturates when the system achieves a steady state corresponding to the condition \(J_{\text{ION}}(x) = 0\) in Equation 5. After that, increasing the magnitude of the voltage will increase the resistance further, until finally the maximum resistance OFF state of the system is attained. A corollary to this prediction is that for the same magnitude but opposite polarity of applied voltage, the speed of switching ON will be significantly faster than switching OFF. This is again related to Equation 5: the electric field and the dopant concentration gradients point in the same direction for switching ON but in opposite directions for switching OFF. A final prediction of this model is that it is possible to invert the switching polarity of a device; for example, applying an ON-switching voltage for a long enough time will eventually cause ions from the grounded side of the device to drift toward the center of the film, thus turning the device OFF. In fact, this effect has long been the basis for determining the mobility of mobile ions by measuring the peak position of the transient electric current that is associated with the transition through the ON state.\textsuperscript{[8,16]}

The model is quite general, and our choice of \(n\)-type mobile species is arbitrary. If the mobile species have \(p\)-type character,\textsuperscript{[17]} the same qualitative results will be obtained from the model. Moreover, similar dynamics will occur if one of the interfaces is transparent to the ions,\textsuperscript{[17,18]} such as when mobile species can completely escape from or be injected into the transport layer.

The model presented here is a foundation for performing more complete simulations. It can be extended to include the generation and recombination currents for both ions and electrons (which was necessary to accurately simulate electron transport for phase change material systems\textsuperscript{[10]}); interface ionic traps;\textsuperscript{[20]} nonlinear drift of electrons and ions, including both the effects of Joule heating and high electric fields;\textsuperscript{[12]} and the presence of both neutral and ionized mobile dopants.\textsuperscript{[21]} By extending the model to higher dimensions, it can capture conductance channel formation\textsuperscript{[18,22]} and other localized effects.

Acknowledgements

This research was supported in part by DARPA. The authors thank D. R. Stewart and A. M. Bratkovsky for insightful comments.


Received: September 9, 2008
Revised: November 11, 2008
Published online: February 18, 2009