Unipolar resistive switching in metal oxide/organic semiconductor non-volatile memories as a critical phenomenon

Citation for published version (APA):

DOI:
10.1063/1.4936349

Document status and date:
Published: 01/01/2015

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
opennaccess@tue.nl
providing details and we will investigate your claim.

Download date: 12. Mar. 2019
Unipolar resistive switching in metal oxide/organic semiconductor non-volatile memories as a critical phenomenon

Benjamin F. Bory,¹ Paulo R. F. Rocha,²,³ Henrique L. Gomes,² Dago M. de Leeuw,³ and Stefan C. J. Meskers¹,a)
¹Molecular Materials and Nanosystems and Institute for Complex Molecular Systems, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
²Instituto de Telecomunicações, Av. Rovisco, Pais, 1, 1049-001 Lisboa, Portugal
³Max-Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

(Received 9 August 2015; accepted 11 November 2015; published online 24 November 2015)

Diodes incorporating a bilayer of an organic semiconductor and a wide bandgap metal oxide can show unipolar, non-volatile memory behavior after electroforming. The prolonged bias voltage stress induces defects in the metal oxide with an areal density exceeding 10¹⁷ m⁻². We explain the electrical bistability by the coexistence of two thermodynamically stable phases at the interface between an organic semiconductor and metal oxide. One phase contains mainly ionized defects and has a low work function, while the other phase has mainly neutral defects and a high work function. In the diodes, domains of the phase with a low work function constitute current filaments. The phase composition and critical temperature are derived from a 2D Ising model as a function of chemical potential. The model predicts filamentary conduction exhibiting a negative differential resistance and nonvolatile memory behavior. The model is expected to be generally applicable to any bilayer system that shows unipolar resistive switching.

I. INTRODUCTION

Many metal-insulator-metal (MIM) systems show non-volatile, electrically induced resistive switching and have been proposed as replacements for standard NAND-flash non-volatile circuitry.¹ A large variety of materials can give rise to resistive switching²–⁴ including organic semiconductors.⁵–⁹ The operating mechanism of organic memory cells is under intense investigation. At present, no consensus has been reached.

A number of studies have appeared indicating the importance of oxide or alkali halide interface layers in the resistive switching process involving organic semiconductors.¹⁰–¹⁴ In organic memories comprising Al electrodes, switching has been attributed to the native oxide. Reproducible memories with a yield of unity were realized by deliberately adding a thin sputtered Al₂O₃ layer in series with an organic semiconducting layer.¹⁵ In their pristine state, these hybrid diodes are highly resistive. Before the diodes show memory properties they have to be electroformed by applying a high bias voltage.¹⁵,¹⁶ When the electric field strength in the Al₂O₃ layer approaches the limit for dielectric breakdown, so-called “soft-breakdown” is induced. The electroforming can be performed by applying a short high voltage pulse or by increasing the bias and setting the current compliance. Diodes can also be formed by constant current stress where the current compliance is set and the voltage is monitored over time.¹⁷ In this way electroforming can be investigated at low power dissipation. Electroforming is generally attributed to the formation of oxygen vacancies and the evolution of oxygen upon electroforming has been well documented in the past.¹⁸

The electroformed device can be switched between a high conductance on-state and a low conductance off-state. The switching can be induced by applying voltage pulses. When voltage pulses of the same polarity are used the switching is called unipolar. In contrast, reversible switching requiring voltage pulses of opposite polarity is referred to as bipolar. Switching in hybrid Al₂O₃ memories is unipolar. Electroforming and resistive switching was first reported in 1962 by Hickmott for thin anodic Al₂O₃ films.¹⁹ Early research on Al₂O₃ up to the 1980s has been thoroughly reviewed by Deamley et al.,²⁰ Oxley,²¹ and by Pagnia and Sotnik.²²

In this study we focus on electroforming and resistive switching in Al₂O₃/polymer diodes. Fabrication and characterization of the diodes is presented in Section II. Section III describes experimental results on the electroforming culminating in an estimate of the areal density of defect in the Al₂O₃ layer. The switching is characterized by electroluminescence (EL) and current noise measurements. Statistical thermodynamics of phase transitions in 2D systems is used to explain the unipolar switching in the fabricated memories. Details are presented in Section IV. Concluding remarks are listed in Section V.

II. EXPERIMENTAL

Diodes investigated typically consist of an Al bottom electrode, a thin sputtered Al₂O₃ layer, a spincoated layer of...
a spirofluorene semiconducting polymer\textsuperscript{23} and a Ba/Al top electrode. To optically access the diodes and detect electroluminescence an ITO electrode was used instead of Al. The layout of the hybrid diodes is shown in Fig. 1(a) and the schematic flat band diagram is shown in Fig. 1(b). The devices were fabricated in clean room conditions on 6 in.\textsuperscript{2} glass wafers. The 30 nm thick Al bottom electrode was thermally evaporated through a patterned shadow mask. The Al\textsubscript{2}O\textsubscript{3} was sputtered with a thickness varying from 20 nm to 50 nm. A typical thickness is 20 nm. The spirofluorene layer of about 80 nm was spincoated in air from toluene. The Ba/Al top contact was thermally evaporated through a shadow mask that allows for overlap with the bottom electrode of precise active areas. All the devices on a wafer were encapsulated with 9 glued stainless steel caps. The polymer was laser ablated where the glue was applied to avoid diffusion of water through the polymer. A getter was placed into the stainless steel cap to exclude H\textsubscript{2}O. The glass substrate was then cut into 9 encapsulated modules, each containing devices with areas between 1 and 9 mm\textsuperscript{2}. A picture of a module is shown in Fig. 1(c). Due to the encapsulation, both pristine devices as well as electroformed memories are stable for years.

The \textit{I-V} characteristics were measured with an Agilent semiconductor parameter analyzer 4156C. Throughout this paper positive bias is defined as the Ba/Al top electrode being charged negative. Typical \textit{I-V} characteristics in the on-state and in the off-state after electroforming are presented in Fig. 1(d). The current is Ohmic at low bias, below \textasciitilde 2 V, and superlinear at higher bias. The on-state shows a clear voltage controlled negative differential resistance (NDR) region.

### III. ELECTROFORMING, ELECTROLUMINESCENCE AND NOISE MEASUREMENTS

#### A. Trapping of electrons at the Al\textsubscript{2}O\textsubscript{3}/polymer interface

Initially, pristine metal/Al\textsubscript{2}O\textsubscript{3}/polymer/metal diodes have a very high electrical resistance. A sequence of events follows upon application of a bias voltage stress. Electroforming occurs. The first step is described as trapping of charges inside the diode. The quasi-static capacitance-voltage (QSCV) method\textsuperscript{24,25} is ideally suited to study traps in semiconductors that fill quickly but empty slowly. In Fig. 2, we compare QSCV measurements for diodes with Ba/Al and Pd top electrodes. As can be seen in Fig. 2(a), the QSCV measurements on diodes with a Pd anode, yields a capacitance value that is independent of the bias voltage applied and equal to the total geometric capacitance for the organic-Al\textsubscript{2}O\textsubscript{3} bilayer, $C_{\text{tot}} = (C_{\text{ox}}^{-1} + C_{\text{pol}}^{-1})^{-1}$\textsuperscript{26}. The work function of Pd is too high for injection of electrons. Furthermore, a significant barrier for injection of holes is expected.\textsuperscript{27} As a result no charge carriers are injected and no trapping of charge occurs in this type of diode.

For a diode with Ba/Al instead of a Pd top electrode (Fig. 2(d)) we find for reverse bias conditions also a constant capacitance equal to $C_{\text{tot}}$ indicating again that no charge carriers are injected. In contrast, the cyclic QSCV scans on the diodes with electron injection Ba/Al contact under forward bias show much larger values for the capacitance and history dependent behavior (Fig. 2(d)). Because the Ba/Al electrode is a quasi ohmic contact for injection of electrons into the polymer, the large QSCV in the Ba/Al diodes provides clear evidence of the existence of deep trap sites for electrons in the diode. As will be explained in

![FIG. 1. Diode layout. (a) Typical nominal e-only metal-insulator-semiconductor-metal diode layout where the Al\textsubscript{2}O\textsubscript{3} thickness is varied. (b) Flatband energy diagram where numbers are in eV. (c) Photograph of a module containing several diodes. The devices with an active area between 1 mm\textsuperscript{2} and 9 mm\textsuperscript{2} were encapsulated to exclude H\textsubscript{2}O. (d) \textit{I-V} characteristics after electroforming showing a low conductance off-state and a high conductance on-state with a pronounced negative differential resistance.](image-url)
more detail below the data indicate trapping of electrons at the Al$_2$O$_3$/polymer interface.

When sweeping the bias applied to a diode with a Ba/Al contact, from low to high, over a range that has not yet previously been applied to the diode, electrons are injected into the polymer. Subsequently these electrons are transported towards the oxide layer and get trapped at the oxide polymer interface. During the voltage sweeping the trapping current across the polymer layer effectively short-circuits the polymer capacitance and allows the observation of the oxide layer capacitance. The oxide geometric capacitance, $C_{ox}$, corresponds to the experimental QSCV value as illustrated in Fig. 2(d). Comparing Figs. 2(d)–2(f), the magnitude of the high capacitance $C_{ox}$ under the forward sweep is inversely proportional to the thickness of the oxide. This is consistent with our previous QSCV measurements. When the scan direction is reversed, the electrons cannot leave the deep trap sites at the interface. In the absence of a current through the polymer layer, the diode behaves as a two-layer capacitor, with a capacitance equal to the series sum of polymer and oxide capacitance. This is confirmed by a much lower capacitance value for the capacitance equal to the oxide capacitance. This is consistent with quasi-irreversible filling of electron trap sites at the Al$_2$O$_3$/polymer interface.

By following the maximum voltage in which this charging behavior can be observed, a lower limit of the areal density of trap sites for electrons at the Al$_2$O$_3$ polymer interface can be determined. We find a density of trap sites exceeding $3 \times 10^{17}$ m$^{-2}$. The deeply trapped electrons can be removed upon illumination. Optical detrapping experiments confirm the density of trap sites, $8 \times 10^{17}$ m$^{-2}$.

**B. Defect formation in the oxide**

Continuing the application of bias voltage stress to metal/Al$_2$O$_3$/polymer/Ba/Al diodes, a process following up on the filling of the electron traps at the Al$_2$O$_3$ interface, is the creation of positively charged defects in the oxide. To prove this defect formation and to estimate their density, we performed time dependent bias voltage stress experiments. Fig. 3 shows the current density upon application of a bias voltage step as a function of time. The height of the bias voltage is increased in steps of 1 V. The measurements are carried out in the dark and at room temperature. However, in between the scans, the diode is illuminated in order to remove the trapped electrons. Immediately after the application of the bias voltage the current density decreases over time. The initial current density arises from the filling of the electron trap sites as discussed above. As the trap sites fill up, the current density decreases over time. Eventually the current saturates to a finite value, which we ascribe to leakage current and tunneling of electrons through Al$_2$O$_3$. After applying a particular voltage step, the trap sites for electrons can be emptied by light illumination for about 10 min. The bias voltage is then incremented, and a next current decay trace recorded.

Upon increasing the bias voltage in the voltage step experiments, the initial current density rises in subsequent measurements. As explained below, the increase in initial current density results from a rise in the capacitance $C_{ox}$ to that of the oxide layer. This indicates that positive charge accumulates in the oxide layer due to formation of positively charged defects. By looking at the end of each time trace, we noticed that the leakage current increases. This is consistent with the increasing defect density in the oxide.

From earlier studies on the kinetic of charging interface states, the current density as a function of time in the voltage step experiment is

![Fig. 2. Cyclic QSCV scans. Quasi-static capacitance-voltage characteristics for Al/Al$_2$O$_3$/polymer (80 nm)/metal diodes with 3 x 3 mm$^2$ area. (a), (b), and (c) diodes with Pd top electrode and oxide thickness of, respectively, 20 nm, 30 nm, and 40 nm. (d), (e), and (f) diodes with Ba/Al top electrode and, respectively, 30 nm, 40 nm, and 50 nm oxide thickness.](image-url)
where $V_o$ is the initial voltage drop over the polymer and $C_{ox}$ is the capacitance associated with the oxide layer. $V_o$ is determined by the bias applied and relative magnitude of the dielectric constants of the polymer, $\varepsilon_r=3$ (Ref. 29) and the oxide, $\varepsilon_r=9$. $n$ and $r$ are parameters describing the voltage dependence of the electron transport in electron-only diodes of the semiconducting polymer in the phenomenological relation26,30

$$J_{pol}^r(V_{pol}) = \frac{1}{r} (V_{pol})^{n+1}. \quad (2)$$

The parameters $r$ and $n$ can be obtained by fitting Eq. (2) to the experimental current density for the electron-only diode, Al/polymer/Ba/Al. We find $r = 100 \text{ m}^2/\text{A}$ and $n = 5.5$, consistent with an earlier analysis of electron transport in poly(spirofluorene). For a voltage step experiment of an oxide/polymer diode, Eq. (1) predicts for the charging current density at early times after the application of the bias step voltage: $J_{\text{Charge}} \propto C_{ox} V_{pol}^{\frac{n}{r}} \approx C_{ox}$. The predictions of Eq. (1) for the current density are illustrated for two values of the applied bias in the right part of Fig. 3. For voltages exceeding 2 V, the current densities practically overlap.

In summary, the experimental observation of an increase in the initial current density with increasing bias indicates a carrier trapping current. Positive charges must be accumulating in the oxide. This results in an effective “thinning” of the oxide layer. Earlier studies have shown that irreversible electroforming takes place under conditions where high electrical power is available when the applied potential induces field strengths in the metal oxide comparable to the critical field strength for dielectric breakdown.15 For diodes with an Al$_2$O$_3$ layer of 20 nm thickness, this requires bias voltages around 12 V.

### C. Electroluminescence in Al$_2$O$_3$/polymer resistive switching diodes

After electroforming, the Al$_2$O$_3$ diodes exhibit electroluminescence. This is demonstrated in Fig. 4. The current densities (upper trace in Fig. 4) are high and show a maximum at around 4.5 V and a second maximum at a higher voltage of about 15 V. In between these two voltages, the diodes exhibit a NDR region. In a voltage range close to 4.5 V, the maximum of the current density shows highly erratic electroluminescence with large bursts in intensity. This is consistent with the NDR region. The light emitted in this voltage range covers a broad part of the visible and near infrared spectrum. In contrast, the electroluminescence at high bias shows much less temporal variability and is of a distinct blue color with spectral characteristics matching those of the photoluminescence emitted by the semiconducting polymer.

Electroluminescence in electroformed metal/Al$_2$O$_3$/metal diodes has been previously reported.22,31,32 In these Al$_2$O$_3$ diodes, the photon energy of the EL is lower for bias voltages near the onset of the NDR region in comparison with EL under high bias voltage.33 Al$_2$O$_3$ diodes can also emit electrons.33 Research on electron emission in low temperature Al$_2$O$_3$ cathodes has indicated the importance of ionized defects in the oxide.34 Therefore both the electroluminescence and electron emission from the Al$_2$O$_3$ memory diodes during operation clearly indicate the importance of ionization and charge recombination at defects in the oxide. The oxygen vacancy defects that have been proposed to play an active role in the resistive switching have levels that lie within the gap of the pristine insulating material.35 Considering the voltages applied, ionization and neutralization of the defects via recombination processes should occur during the switching between high and low conduction states.36

The observation of electroluminescence with a wavelength centered at 475 nm from the diodes during the resistive switching argues against creation and destruction of
metal filaments between the electrodes as the source of the resistive switching.\textsuperscript{37,38} The electroluminescence also indicates that migration of charged defects may not be necessary to explain the resistive switching. Ionization and recombination of charges at defects resulting in electroluminescence provides a more likely explanation for the fast redistribution of charges in the diodes.\textsuperscript{39,40}

D. Filamentary conduction in formed diodes

The homogeneity of the electrical conduction in the electroformed diodes has been investigated with an IR enhanced CCD camera.\textsuperscript{41} The spatially resolved thermal images show hot spots in the on-state due to highly conductive paths. In the off-state, the spots disappear. However, the spots are not created and destroyed upon switching. Upon repeated switching between the on- and off-states, the same original hot spots were detected in the thermal image. From these observations it has been concluded that upon switching, filaments are neither generated nor destroyed, but that individual filaments are turned on and off, like switches. This conclusion has been corroborated by the anomalous temperature dependence of the on-state\textsuperscript{42} and by electrical noise measurements. We note that scanning probe measurements confirm the existence of conducting filaments in Al$_2$O$_3$.\textsuperscript{43}

At low bias, in the Ohmic regime, the transport is almost activation-less. In the off-state at high bias the conduction is thermally activated. In a temperature range of 200 K–300 K, an activation energy of about 66 meV has been reported.\textsuperscript{42} Interestingly, the on-state behaves completely different. Switching diodes were programmed in the on-state and the current was measured upon lowering the temperature. The current doubles in a temperature range of 150 K, which corresponds to a positive temperature coefficient of the electrical resistivity of about 0.01 K\textsuperscript{−1}. This is an anomalously large value when compared to typical values of metals, such as 0.0039 K\textsuperscript{−1} for Cu. Furthermore, in contrast to the smooth decrease in the resistivity of metals, the current over temperature of our switching diodes decreases in a step-like fashion. This strongly suggests that additional conducting filaments become active upon lowering the temperature. This behavior contradicts the widely held view that switching is due to filaments that are formed reversibly by the diffusion of metal atoms. Instead, the anomalous temperature dependence together with small-signal impedance measurements have indicated that creation of filaments is controlled by filling of shallow traps localized at the oxide/polymer interface.\textsuperscript{42}

Filamentary conduction and trap assisted switching unambiguously follow from electrical noise measurements.\textsuperscript{44} The current noise, $S_I$ ($\text{A}^2/\text{Hz}$), is presented in Fig. 5 as a function of frequency on a double logarithmic scale. The memory diode is programmed in the on-state. At low bias, where the current is Ohmic, the noise follows the omnipresent 1/f dependence.\textsuperscript{45} A comprehensive theoretical analysis of 1/f noise in phase change memories has been presented by Nardone et al.\textsuperscript{46} We have analyzed the 1/f noise using the empirical Hooge relation.\textsuperscript{44,47} The figure of merit extracted of $10^{-21}$ cm$^3$/Ω is typical for conducting materials such as metals and doped semiconductors. However, averaged over the whole active area of the memory diode, the current density is low. This indicates that the regions responsible for current transport and noise are narrow but highly conducting. Comparable noise measurements were obtained for the off-state. The extracted figure of merit scales with the conductivity.

At high bias, close to the NDR region, the average noise level increases. Typical time records of 250 ms from a continuous measurement during 6 h are presented in Fig. 6. The time records show random telegraph noise (RTN); large discrete current fluctuations of about 45 nA, corresponding to $\Delta R/R \sim 5\%$. RTN has been observed in a variety of systems such as $p$-$n$ junctions,\textsuperscript{48} MIM junctions,\textsuperscript{49,50} MOS transistors\textsuperscript{51} and has been used to characterize resistance fluctuations in chalcogenide resistive random access memories.\textsuperscript{52} RTN is described as discrete fluctuations in current-voltage or current-time characteristics when charge transport is controlled by the statistical capture/emission of electrons at electron trap sites. Especially when transport occurs through current-carrying filaments, large current fluctuations can occur. RTN noise is due to individual events involving trapping of charge carriers that cause localized and discrete modulations of the current. RTN is therefore an effective tool for probing the dynamic behavior of an electrically inhomogeneous system such as our memory diode.

IV. UNIPOLAR MICROSCOPIC SWITCHING MECHANISM

A. Introduction

We explain the electrical bistability of the non-volatile memories by the phase coexistence of two thermodynamically stable phases. The two phases occur in the two

![FIG. 6. RTN Noise at high bias. Electrical noise of a memory diode programmed in the on-state at a temperature of 220 K. The time records show the current RTN fluctuations under an applied bias of 1.2 V at different intervals during 6 h of measurement. The discrete current levels are presented on the left side.](image-url)
Phase coexistence occurs in many physical systems, such as ferromagnets. The phase coexistence depends on temperature. At a critical high temperature, the phase coexistence vanishes and the domains disappear. At the same time the phase composition can be controlled by external physical quantities such as a magnetic field in the case of ferromagnets. For the non-volatile memories discussed here, the external driving force is the electrical potential, automatically leading to negative differential resistance. Crucial ingredients for phase coexistence and a critical point are a cooperative interaction between quasi particles in the thermodynamic system leading to a negative exchange parameter interaction \( J \) and an interaction of the quasi-particles with an external potential.

This section is organized as follows. First, in Section IV B, we introduce the quasi-particle that consists of a trapped electron and a trapped hole together with their image charges in the metallic electrode and then discuss its cooperative electrostatic interaction with other quasi-particles at zero temperature and zero electric potential. In Section IV C, we adopt the 2D Ising model to describe the coexistence of a mainly neutral and a mainly ionized phase. Strictly speaking, the Ising model applies to systems in thermodynamic equilibrium. For a complete description of the switching kinetics in the memory diodes, we adapt the model to also include effects of carrier injection and associated power dissipation. We calculate the phase composition and the critical temperature. Subsequently the composition is obtained as a function of chemical potential. This yields domains with different work functions which automatically leads to a nonvolatile memory exhibiting filamentary current filaments. We note that the formation of charged double layers near metal electrodes is well known in wet electrochemistry.53

Phase coexistence occurs in many physical systems, such as ferromagnets. The phase coexistence depends on temperature. At a critical high temperature, the phase coexistence vanishes and the domains disappear. At the same time the phase composition can be controlled by external physical quantities such as a magnetic field in the case of ferromagnets. For the non-volatile memories discussed here, the external driving force is the electrical potential, automatically leading to negative differential resistance. Crucial ingredients for phase coexistence and a critical point are a cooperative interaction between quasi particles in the thermodynamic system leading to a negative exchange parameter interaction \( J \) and an interaction of the quasi-particles with an external potential.

This section is organized as follows. First, in Section IV B, we introduce the quasi-particle that consists of a trapped electron and a trapped hole together with their image charges in the metallic electrode and then discuss its cooperative electrostatic interaction with other quasi-particles at zero temperature and zero electric potential. In Section IV C, we adopt the 2D Ising model to describe the coexistence of a mainly neutral and a mainly ionized phase. Strictly speaking, the Ising model applies to systems in thermodynamic equilibrium. For a complete description of the switching kinetics in the memory diodes, we adapt the model to also include effects of carrier injection and associated power dissipation. We calculate the phase composition and the critical temperature. Subsequently the composition is obtained as a function of chemical potential. This yields domains with different work functions which automatically leads to a nonvolatile memory exhibiting filamentary conduction exhibiting a negative differential resistance. Finally, in Section IV D, the predictions of the Ising model are related to the experimental measurements as described in Section III.

B. Quasi particles and their interaction energy, \( J \)

In Fig. 7, we have summarized graphically the experimental evidence on charge accumulation at the oxide-organic interface. Electrons injected into the organic semiconductor accumulate at the interface of the polymeric organic semiconductor with the metal oxide occupying deep trap sites with areal density estimated as \( 8 \times 10^{17} \text{cm}^{-2} \). Prolonged bias voltage stress induces the formation of defects in the oxide. Presumably, these defects are anion vacancies that can exist in various charge states. Once the density of trapped charges exceeds a certain threshold, soft breakdown of the oxide occurs and electroforming is complete. The electroformed diodes show electrical bistability, and experimentally it was found that diodes with a thin oxide layer (10–20 nm) can be switched most reliably.15 Therefore, we focus our discussion on thin oxide layers. In diodes with thin oxide, the trapped carriers are close to the metal electrode. The image charges in the metal should be taken into consideration when investigating the electrostatic potential energy of arrangements of trapped charges. In particular, one of the simplest, low energy charge configurations that one can think of is the quartet of charges illustrated in the left side of Fig. 7. This quasi-particle consists of a trapped electron in the organic semiconductor, an ionized defect in the metal oxide, and their two image charges in the metal.

We now calculate the interaction energy, \( J \), between the quasi-particles. When the interaction energy is negative, then at absolute temperature, the quasi-particles will condense into an array. In this array, the charges mutually stabilize each other, with for instance a trapped electron binding two charged vacancies together. The stabilization can be supported with the following intuitive argument. Looking at the two uppermost layers of charges, the trapped electrons, and the defects, we notice that a similar layered arrangement of charges also occurs naturally in ionic crystal, e.g., NaCl. The crystallographic (1,1,1) and (2,2,2) planes in a NaCl crystal each contain ions of the same charge that are arranged in a triangular 2D lattice. The triangular arrangement is further illustrated in Fig. 8.

The array of condensed charges contains a bilayer of positive defects in the oxide and negatively charged trapped electrons in the organic semiconductor. In Fig. 7, we have illustrated an idealized structure of the charge condensate assuming that the defects in the oxide and the trap sites in the organic semiconductor are positioned in a triangular lattice with lattice constant \( a \). From the density of trap sites

![Fig. 7. (Upper) Schematic representation of charges near the oxide/organic semiconductor interface in electroformed, metal oxide organic semiconductor diodes. Arrays of mainly neutral defects have a high work function and constitute the off-state. Arrays of mainly ionized defects have a low work function and constitute the on-state. (Lower) Schematic representation of a current filament.](image)
determined experimentally, $8 \times 10^{17} \text{m}^{-2}$, we estimate the lattice constant $a$ to be in the nanometer range, $a \approx 1/\sqrt{(8 \times 10^{17} \text{m}^{-2})} = 1.1 \text{nm}$. The distance of the closest approach between trapped electrons and defects in the vertical direction is denoted by $z$, and $b$ indicates the distance between the defects in the oxide and in the metal. $b$ and $z$ are also expected to be in the nanometer range. The triangular lattice of trap sites for electrons can be displaced relative to the triangular lattice of the defects in the oxide by a variable offset $h$, expressed in units of the lattice constant $a$. If $h = 0$ corresponds to a perfectly linear quartet. If $h \neq 0$, the individual vertical quartets of charges do not have a linear geometry in the vertical direction, but are actually bent (see Fig. 8(b)). A bent quartet has a higher electrostatic energy than a linear one, and so an energy penalty, $\Delta E_{\text{bend}}$, should be associated with bending. The case $h = 1$ corresponds to a perfect hexagonal packing of the positive and negative charges.

The electrostatic stabilization energy of each quartet by its neighbors, $E_{\text{stab}}$, in the array can be calculated by considering the pairwise interaction $w_{ij}$ between different quartets $i$ and $j$

$$w_{ij} = \begin{cases} 0 & \text{if } i = j \\ \frac{1}{2} \sum_{n=1}^{4} \sum_{m=1}^{4} \frac{q_{n,i}q_{m,j}}{|r_{n,j} - r_{m,j}|} & \text{otherwise} \end{cases}$$

where $q_{n,i}$ indicates the individual elementary charges in the quartets. We illustrate the numerical calculation of the interaction energy $J$ for a finite sized, triangular array of $10 \times 10$ quartets (Figs. 8(d)–8(f)). Energies are in units of $E_{\text{ref}} = q_e^2/4\pi\varepsilon_0 e_a a$ with $a$ the lattice constant of the triangular lattice of trapped charges. The average stabilization of a quartet in the $N \times N$ array through interactions with other quartets $\langle \Delta E_{\text{array}} \rangle$ is now

$$\langle \Delta E_{\text{array}} \rangle = \frac{1}{2N^2} \sum_{i,j} w_{ij}.$$  \hspace{1cm} (4)

To calculate the overall stabilization energy per quartet, $E_{\text{stab}}$, we include the penalty for bending the quartet

$$E_{\text{stab}} = \langle \Delta E_{\text{cluster}} \rangle + \Delta E_{\text{bend}}.$$  \hspace{1cm} (5)

Values for $E_{\text{stab}}$ computed numerically are illustrated in Figs. 8(d)–8(f). The negative energies evidence a net electrostatic stabilization of the quartets by clustering into an array. The stabilization energy for infinitely large arrays is somewhat more difficult to calculate because the infinite sums of interaction energies are only conditionally convergent. Nevertheless using, e.g., the Ewald summation technique, \cite{54,55} for rapid convergence, a net stabilization similar in magnitude to the energy for the $10 \times 10$ condensate relative to the isolated linear quartets as a function of the distances $h$, $z$, and $b$ (see (a)), in units of the lattice constant. Energies are in units of $E_{\text{ref}} = q_e^2/4\pi\varepsilon_0 e_a a$, where $a$ is the lattice constant of the triangular lattice of trapped charges.

FIG. 8. (a) Triangular lattice of trapped charges at the metal oxide/organic semiconductor interface. (b) Isolated quartets of trapped charges in their lowest energy linear arrangement, $h = 0$, and in a bent arrangement, $h \neq 0$, with higher electrostatic potential energy. (d)–(f) stabilization energies from Eq. (5) for an array of $10 \times 10$ charge quartets consisting of a positively charged defect in the oxide a trapped electron in the organic semiconductor plus image charges (see (b)). The negative energies show the average stabilization per (bent) quartet in the $10 \times 10$ condensate relative to the isolated linear quartets as a function of the distances $h$, $z$, and $b$ (see (a)).
estimate for the defect density ($\rho_{\text{defect}} = 1/\mathfrak{a}^2 = 8 \times 10^{17} \text{ m}^2$, see Section III), the typical energy scale of the condensation of charges is then $E_{\text{ef}} = qe^2/4\pi \mathfrak{a} \rho_{\text{defect}} = 0.14 \text{ eV}$. From Fig. 8, we see that $E_{\text{stab}} \approx 3 E_{\text{ef}}$. Thus we expect the electrostatic stabilization energy per quasi-particle to be on the order of half an eV.

C. 2D Ising model and mean field solution

So far we have considered quasi-particles consisting of a quartet of charges and their mutual interactions. The quasi-particles are not necessarily thermodynamically stable. The trapped electron can recombine with the trapped hole reestablishing a neutral defect. The excess energy can be released, e.g., as heat or by emission of a photon. Conversely, a neutral defect can ionize, forming the quasi-particle of a quartet of charges. To calculate the average of the areal density of charged quartets, or degree of ionization, we calculate the total energy of a 2D lattice of sites that can be either empty or occupied by a quasi-particle. This total energy has two main contributions. First, the creation of an individual quasi-particle via ionization of defects requires a certain amount of energy. This energy of an individual defect can be described by a chemical potential $\mu$, which depends on the electrical potential at the interface. Next, the interactions between the charged quasi-particles require a second term involving their interaction energy $4J$. Here, a factor of 4 is introduced for easy conversion to the standard formulation of the 2D Ising model. For simplicity we consider a simple 2D-lattice model with only nearest neighbor interactions. The Hamiltonian energy function for the system of quasi-particles can then be given by

$$\mathcal{H} = \mu \sum_i c_i + 4J \sum_{\langle i,j \rangle} c_i c_j, \quad (6)$$

where $c_i$ denotes now the ionization state of the defect on site $i$ in the oxide layer and can take the values of either 0 (neutral) or 1 (+, ionized state). The first summation in Eq. (6) runs over all lattice sites $i$ that can potentially be occupied by a quasi-particle while the second summation is over pairs $i,j$ of neighboring lattice sites. $4J$ indicates the electrostatic interaction energy between a pair of quasi-particles located on nearest neighbor sites. We assume a square lattice in which each site has 4 nearest neighbors. Energy is in units of $k_B T$

In order to obtain the average degree ionization from Eq. (6), we make use of the well-known analogy between the lattice model expressed by Eq. (6) and the Ising model for magnetism.57,58 By making the substitutions $c_i = (S_i + 1)/2$, $\mu = 2H - 8J$, to the lattice model Hamiltonian Eq. (6) one obtains the Hamiltonian function Eq. (7) pertaining to the 2D-Ising model for magnetic phase transitions

$$\mathcal{H} = H \sum_i S_i + J \sum_{\langle i,j \rangle} S_i S_j, \quad (7)$$

where $S_i$ denotes the state of the spin at site $i$ with possible values +1 and -1. $J$ stands for the nearest neighbor exchange interactions with $J < 0$ corresponding to ferromagnetic coupling. $H$ represents the magnetic field.

In the mean field approximation and including only nearest neighbor interactions, solutions for the 2D Ising problem are well known.57,58 The mean field solutions are illustrated in Fig. 9. Fig. 9(a) shows the temperature at which a particular average degree of ionization of the defect can

\[\text{FIG. 9. Mean field 2D Ising lattice model. (a) Phase composition at various reduced temperatures } \theta \text{ for reduced chemical potential } \mu^* = 0 \text{ (black) and 0.2 (red). The critical point occurs at } \theta_c = 1, \mu^* = 0 \text{ and degree of ionization equal to one half. (b) Degree of ionization versus reduced chemical potential at } \theta = 0.6, \text{ illustrating the coexistence of a mainly neutral “off” state and a highly ionized “on” state for potentials } \mu^* \text{ below 0.4. At a higher chemical potential only the “off” state is thermodynamically stable. (c) Current in the on and off state calculated as the degree of ionization times the reduced potential. At } \mu^* \text{ about 0.4 the electroformed diode shows a negative differential resistance as at high potential only the off state is stable. The NDR is indicated by the dashed red line. (d)Maximal on/off ratio for the current in the electroformed diode as a function of the reduced chemical potential for various values of reduced temperature.}\]
occur. Here, a reduced temperature $\theta = k_b T/|J|$ and a reduced chemical potential $\mu^* = (\mu/|J|) + 2$ are noted. Note that $\mu^*$ is equal to two times the reduced magnetic field $h = H/|J|$ in the spin Ising model. In Fig. 9, the black lines with a pitchfork shape illustrate the case for $\mu^* = 0$, where there is effectively no preference for a defect to be ionized or neutral. At high temperature and $\mu^* = 0$, there is always enough thermal energy available to ionize/neuterize any defect, and we see that the average degree of ionization of a defect is $1/2$, as demanded by a maximum configurational entropy. At $\mu^* = 0$ and $\theta_c = 1$ the system reaches its critical point. Upon lowering the temperature below the critical point, two phases can coexist, one with mainly neutral defects and the other with mainly ionized defects. Below the critical point, the creation of an isolated quasi-particle on a neutral lattice site is energetically unfavorable. However, the cooperative interactions between the quasi-particles help to stabilize a cluster of particles. Hence, the surface area subdivides into domains that are either almost completely neutral or fully ionized albeit that they have the same free energy per site. At low temperature, the small thermal energy available to quasi-particles does not allow them to escape from their mutually attractive interactions in the cluster. Quasi-particles have to cluster together to gain enough stabilization to escape recombination.

Upon raising the chemical potential $\mu^*$ we introduce a preference for the defects to become neutral. This is illustrated in Fig. 9(a) by the red curve calculated for $\mu^* = 0.2$. Below a reduced temperature $\theta = 0.7$ the two phases still coexist. However, only the mainly neutral phase is thermodynamically stable above $\theta = 0.7$.

**D. Relation of the 2D Ising predictions with experimental measurements**

The total electrostatic stabilization of the in nearest neighbor approximation on a square lattice with four nearest neighbors is $16|J|$. Making use of the estimate $E_{\text{stat}} \approx 0.5$ eV, as discussed in Section IV B, we find $J \approx 0.03$ eV. Using this estimate of $J$, one then calculates a critical temperature $T_c = 0.4/|k_b|$ of about a 1000 K. Such a high critical temperature is consistent with the experimental non-volatile nature of the conduction states of the memory diode. We note that the critical temperature depends on the value of the interaction energy $J$ and thus implicitly on the density of defects. Room temperature, $T = 300$ K, corresponds to $\theta = 0.2$ for which the calculations predict $\mu^*$ to be close $0.7$. This then translates into critical potential for switching on the order of a volt, in reasonable agreement with the experiment.

For a large separation between the defects, i.e., a large lattice constant $a$, the interactions between neighboring lattice cells become dipolar in character and so the nearest neighbor interaction $J$ is expected to scale with the defect density as $1/a^4 \propto \rho_{\text{defect}}$. This means that at fixed temperature $T$, the reduced temperature $\theta$ is inversely proportional to defect density. With this scaling in mind, Fig. 9(a) now provides a qualitative explanation for the soft breakdown of the oxide in the electroforming process. Starting at low defect density and $\mu^* = 0$, the reduced temperature $\theta$, is high, and the ionization state of the defects is uniform and equal to $1/2$. With increasing bias voltage stress, defects are formed and the defect density increases. At a certain defect density, the reduced temperature $\theta$ will go below the critical point $\theta_c = 1$. Now phase segregation between ionized and non-ionized domains is possible and the diode starts to show bistability.

Fig. 9(b) gives a different representation of the solution for the mean field Ising model that is more appropriate for explaining the resistive switching. The average degree of ionization is plotted on the vertical axis, with the applied external potential on the horizontal scale. At reduced temperature $\theta = 0.6$ and at zero potential, two phases can coexist. Under these conditions the phases are not “pure” and have an average degree of ionization 0.95 and 0.05. Upon raising the potential, the phase that has mainly ionized defects is destabilized and vanishes at $\mu^* = 0.4$.

Associated with the ionized phase is a step in the electrostatic potential profile in the perpendicular direction to the surface with respect to the neutral phase. This potential step modifies the effective work function of the metal electrode. This may be compared to the shift in metal work function that can be induced by self-assembled monolayers on the surface of noble metals.

The formation of charged double layers in the ionized phase can explain earlier experiments showing that the resistive switching is associated with a change of the effective work function of a metal electrode covered with a thin layer of alkali halide. Assuming that in the double layer the charge density is the same as the trap density for the electrons at the polymer interface, $8 \times 10^{17}$ m$^{-2}$, we can easily calculate the work function shift $\Delta \psi = \alpha z e / |q|$, with $z$ the distance between the layers of opposite charge and $\alpha$ the (average) 2D charge density. Taking a distance $z$ of about one nanometer, the presence of the double layer will induce a shift of the effective work function of the metal electrode of typically several volts. The shift of the work function alters the charge injection and thus modulates the electrical resistance. The part of the metal electrode covered by the ionized phase will have a lower barrier for electron injection from the metal into the polymer and can be associated with the high conduction on-state of the diode. We assume that electrical current will be limited by injection. The neutral phase hardly carries any current. We simplistically equate electrical conduction with the product of degree of ionization with applied potential.

Fig. 9(c) illustrates the dependence of the conduction of the two phases on applied potential. At low potential and $\theta < 1$, two conduction states are possible, “on” and “off,” depending on the history of the diode. Upon raising the potential, the ionized phase loses its stability and diodes switch from the high to the low conduction state, corresponding to a negative differential resistance.

In a large area diode the surface area is subdivided into domains that are either almost completely neutral or fully ionized. The domains with ionized defects have a low work function that allows for injection of electrons. The neutral domains have a high work function that limits the injection of electrons. As a consequence the domains with ionized
defects constitute current filaments. The filament can be visualized through, e.g., thermal imaging and current noise measurements. We note that a filament involves electrical conduction through the oxide. However, the state of the filament, on- or off, is determined by the ionization state at the interface between the oxide and polymer. As a consequence the switching can be extremely fast and the dominant noise in the current is random telegraph noise.

In Fig. 9(d), we plot the maximum difference in conductivity of the diodes assuming that the electrode is completely covered by either ionized or neutral phase. At low reduced temperature (implying low $T$ and/or high $\rho_{\text{defect}}$) the on/off ratio is high and the bistability extends to high values of $\mu^0$. At higher $\theta$, both the contrast and the bistability limit are reduced.

Now we can comment on the process of switching itself. Due to the interactions between neighboring sites, switching is a collective process of many defects. We assume that when an ionized defect becomes neutral when recombinig with an electron an excited state can be produced. The collective switching of the ionized defects to the neutral state will then reveal itself by bursts of light. We note that the collective switching provides an alternative approach to stimulated emission that is not controlled by Bose-Einstein statistics of the photon and may be of practical importance. The collective switching can also account for the large quasi-discrete fluctuations in the electrical current (see Fig. 6). Earlier studies on random telegraph noise in SiO$_2$ by Farmer et al. have already provided evidence for collective ionization of defects as the cause for anomalous random telegraph noise.$^{62}$

Finally, we discuss the electrical bistability of the diode and its unipolar switching. Application of a bias voltage will result in an electric field that destabilizes the dipolar charged double layer. Thus with increasing bias, the electrochemical potential $\mu$ of the quasi-particles in the Ising type model increases. At a certain value of the applied bias, the electrochemical potential will reach its critical value and the charged double layer constituting the on-state is no longer stable. Thus application of bias voltage to a memory diode in the on-state exceeding the stability limit will induce a switch to the off-state, which is observed as a negative differential resistance in the $IV$ characteristics.

We now discuss switching from the off-state to the on-state. In the off-state defects are mainly neutral while in the on-state they are charged. Hence switching from off- to on-requires injection and/or redistribution of charge. If we start with a memory diode in its off-state at a high bias voltage, then upon lowering the bias below the critical value the on-state becomes stable. However, the formation of the on-state requires redistribution of charge and, therefore, is kinetically hampered. When the bias is lowered sufficiently quickly the diode remains in the off-state. If however the bias is lowered sufficiently slowly for movement of charge to occur, the diode may switch to the high conduction on-state.

We note that the memory diode connected to an external electrical power supply constitutes an open thermodynamic system. Under application of bias, continuous power dissipation and entropy generation will occur even if the diode remains in the same conduction state. Under application of bias voltage below the critical limit, the diodes may be in several possible stationary states. Ultimately, one expects the diodes to adopt the stationary state with the highest kinetic stability. Yet, the question “which of the possible stationary states is the most stable?” cannot be answered on the basis of equilibrium thermodynamic considerations alone but requires a detailed analysis of all kinetic processes involved.$^{63}$ Such an evaluation of kinetic stability is beyond the scope of the present investigation. Experiments indicate that for voltages below the critical bias for switching, the on state has the highest stability. It was found that the time interval $t_d$ needed to switch a diode from the off- to the on-state depends exponentially on the magnitude of the bias applied, $t_d = t_0 \exp(-\gamma V_d).$ This experimental finding is consistent with the transition from the off-state to the on-state driven by injection and redistribution of charge.

V. SUMMARY AND CONCLUSION

Unipolar resistive switching in Al$_2$O$_3$/organic semiconductor diodes involves defects in the oxide that are created during electroforming. We extracted a defect density on the order of $10^{17}$ m$^{-2}$. Reproducible electroforming is possible by including a layer of a semiconducting polymer. First, this polymer layer acts as a current limiting series resistance that prevents thermal runaway during electroforming and allows for memory operation without external current limitation. Second, the presence of the polymer introduces an internal polymer/oxide interface, where charges can condense into a 2D charged double layer.

We have introduced quasi-particles consisting of an ionized defect in the oxide, a trapped electron in the polymer, and the corresponding image charges in the electrode. By considering Coulombic interactions between quasi-particles, their cooperative interaction energy, $J$, has been estimated as about 0.1 eV. Quasi-particles can recombine to a neutral defect. A 2D lattice model with sites that can be neutral or occupied by a quasi-particle, analogous to the Ising model, has been introduced to calculate the total free energy. The model predicts the coexistence of a mainly neutral and a mainly ionized phase for temperatures below a critical limit. The phase composition is obtained as a function of chemical potential. The two phases have a different work function which automatically leads to a nonvolatile memory exhibiting filamentary conduction with a negative differential resistance. The Ising model can account for the electrical bistability. Thus unipolar resistive switching in Al$_2$O$_3$/organic semiconductor diodes can be regarded as an emergent behavior that arises when the density of ionizable defects in the metal oxide exceeds a critical limit. Predictions of the Ising type model are related to the experimental measurements. The model is expected to be generally applicable to any bilayer system that shows unipolar resistive switching.

ACKNOWLEDGMENTS

The work forms part of the research programme of the Dutch Polymer Institute (DPI), Project DPI No. 704, BISTABLE. We gratefully acknowledge the financial support received from Fundação para Ciência e Tecnologia (FCT).