Abstract—High reset energy is an ongoing issue for phase-change memory (PCM) devices. Prior work demonstrates that smaller PCM switching volume and thermal isolation can reduce the reset energy. In this paper, we fabricate and measure a planar confined PCM device with a multilayer dual-layer stack (DLS) of SiO$_2$/Al$_2$O$_3$ insulator. Devices with contact area of 500 × 20 nm and lengths of 2 μm show exceptionally low reset energies of 18.25 ± 15.8 pJ and low reset current densities of 0.94 ± 0.51 MA/cm$^2$. Implementing the DLS enables a 60% reduction in reset energy compared with SiO$_2$-isolated devices.

Index Terms—Phase-change memory (PCM), reset energy, thermal conductivity, thermal design.

I. INTRODUCTION

Phase-change memory (PCM) has reached production as a nonvolatile storage-class memory (SCM), with access times and costs between NAND flash and DRAM [1]–[3]. In order to create high-density SCM, lower reset energy PCM is required. The high reset energy stems from the temperatures (between 800–1000 K) required to heat and melt the phase-change layer [1]. Two main techniques have been implemented toward lowering the reset energy: reducing the phase-change volume melted via device size scaling [4]–[8] and cell design [9]–[12] or achieving better heat confinement in the structure to create higher temperatures for an applied energy [13]–[15]. To reduce the volume melted, small contact areas have been implemented, down to two nm via carbon nanotube contacts, lowering the reset energy to ∼1 pJ [5]–[7]. In addition, unique cell designs, such as mushroom [8], edge [9], [10], and confined [11] structures have been implemented to limit the phase-change volume melted. Several works investigated the thermal confinement of the mushroom type structure by implementing thermally resistive bottom electrodes. TiN/TaN multilayers [13], fullerenes [14], and graphene bottom electrodes [15] were used and show 90% power reduction, 75% current reduction, and 40% current reduction compared with control devices with standard electrodes, respectively. For confined and edge structures, the relative heat loss through the surrounding insulators is significantly higher. Specifically, for confined cells as shown in Fig. 1(a), thermal simulations show that 89% of the heat is dissipated into the insulator [16]. A strong reduction of the reset energy is predicted by reducing the thermal conductivity of the insulator [16], [17].

In this paper, we implement an SiO$_2$/Al$_2$O$_3$ multilayer [16], [17] for the insulator of a planar confined PCM device. A false-colored TEM of the multilayer SiO$_2$/Al$_2$O$_3$ multilayer stack (DLS) is shown in Fig. 1(b). Devices using the DLS as the insulator require ∼60% less reset energy than devices with SiO$_2$. Scanning thermal microscopy (SThM) and Raman mapping are used to identify the highest temperature region and thus the phase-transformation region: the interface between the phase-change layer and the patterned electrode. Scaling the phase-change volume (via length scaling) reduces the energy, implying that the bulk region is where most of the energy is generated during switching. The fabrication process flow and a further description of the multilayer SiO$_2$/Al$_2$O$_3$ are described in Section II. Section III describes how the various device electrical measurements were conducted. Specifically, the following data will be presented: DC sweep response,
resistance transition due to applied current, endurance, and reset energy for different device dimensions. In Section IV, SThM measurements will be used to identify the location of the switching region and finite-element (FE) simulations and analytical models will be used to describe the heat generation through the PCM device.

II. FABRICATION PROCESS FLOW AND DIELECTRIC STACK

The planar confined PCM devices were fabricated using the process flow outlined in Fig. 2. First, the bottom insulator layer was deposited onto Si wafers. The samples were split between control films, which had 50 nm of plasma-enhanced chemical vapor deposition (PECVD) SiO₂, and DLS films, which had an atomic layer deposition (ALD) oxide comprising of [SiO₂ (2)/Al₂O₃ (1)]₁₇ multilayer (where [SiO₂ (A)/Al₂O₃ (B)]C means A-nm thick SiO₂ layers and B-nm thick Al₂O₃ layers repeated C times). The ALD SiO₂ and Al₂O₃ were deposited using O₂ plasma with Tris[dimethylamino]silane for SiO₂ and Trimethylaluminum for Al₂O₃ deposition. Next, a 20-nm W layer was deposited using e-beam evaporation. The W layer was then patterned and dry-etched with SF₆ [20] to form the electrode pads and metal line. For the W line, the width remaining was 500 nm and the length was varied to 4, 10, and 20 µm. Next, half of the metal region was patterned to form the phase-change region. The width for this pattern was 500 + 100 nm overshot on each edge (to account for misalignment to the W metal line) and the length was half of the W line length +1 µm into the pad. This patterned region was then etched to remove all of the W and filled with 20 nm of sputtered Ge₂Sb₂Te₅ (GST). The remaining GST on the unpatterned regions was then lifted off to create the self-aligned phase-change layer seen in Fig. 2(b). Following lift-off, the entire wafer was coated with a top insulator. The SiO₂ control samples were formed by depositing 10 nm of ALD SiO₂ + 40 nm of PECVD SiO₂. The DLS samples were formed by ALD depositing [SiO₂ (2)/Al₂O₃ (1)]₁₇, after the top insulator was deposited, via holes were patterned and wet-etched with 20:1 buffered oxide etch to contact the metal pads. Finally, 5-nm Ti + 85-nm Pt pads were patterned, deposited by e-beam evaporation, and lifted off. The final structure is shown via scanning electron microscope (SEM) in Fig. 2(f).

The multilayer DLS film comprising of [SiO₂ (2)/Al₂O₃ (1)] repeated shows low thermal conductivity and was developed in previous work [16]. The through-plane thermal conductivity of this film was measured using a steady-state electrothermal measurement of a metal heater on the film [16]. By measuring the change in electrical resistance of the metal heater due to applied power and at different temperatures, the thermal resistance of the multilayer film was extracted [16], [21]. Then, assuming 1-D heat conduction through the film, an effective thermal conductivity was calculated. Samples with various number of layers with the same total thickness (60 nm) were deposited, and the resulting thermal conductivities are shown in Fig. 3. The thermal conductivity of the DLS film is significantly lower than bulk ALD SiO₂ [18] and Al₂O₃ [19] films. Increasing the number of layers reduces the thermal conductivity, ultimately reaching a 64% reduction when compared with the bulk SiO₂ thermal conductivity. The decrease in thermal conductivity is attributed to two main reasons: a change in film quality resulting in lower density and thus lower intrinsic thermal conductivity of each of the layers, and a finite thermal boundary resistance (TBR) between the two layers of ~1.5 m²K/GW, which contributes significantly when the number of layers is large [16]. This multilayer film was leveraged to create a thermally confined PCM device [17] and is further explored in this paper.

III. ELECTRICAL MEASUREMENTS

Electrical resistance measurements of the devices were performed using an HP 4156C Parameter Analyzer, programming pulses were applied via an Agilent 811110A pulse generator, and a 707B Keithley switch matrix was used.

Fig. 2. (a)–(d) Process flow to fabricate planar confined PCM devices (top view). (a) First, an insulator blanket layer of either SiO₂ or DLS was deposited. Then, W was e-beam evaporated. The W was then patterned to form pads and lines, while the remaining area was dry-etched. (b) Phase-change region was patterned, then W was dry-etched, and then a GST layer was deposited and lifted-off. This created a self-aligned GST region. (c) Top insulator blanket layer was deposited with the same film as the bottom layer. The insulator layer was then patterned and wet-etched to form vias to contact the pads. (d) Large pads were then patterned. Ti + Pt was deposited to form the pads and contact the PCM device. (e) Side profile view of the final fabricated structure. (f) SEM of the final fabricated device (top view).

Fig. 3. Through-plane thermal conductivity of SiO₂/Al₂O₃ multilayers with 60-nm total thickness and varying individual layer thicknesses ([SiO₂ (A)/Al₂O₃ (B)]C; means A-nm thick SiO₂ layers and B-nm thick Al₂O₃ layers repeated C times). The thermal conductivity decreases as the layer thickness decreases [16]. The multilayer films exhibit lower thermal conductivity than their bulk ALD SiO₂ [18] and Al₂O₃ [19] constituents.
to switch between the two tools. A 200-mV DC voltage was used to read the resistance of the device. Reset and Set were performed by applying a 5/50/5 and 100/100/100 ns rise/width/fall time pulse, respectively. During programming, the input impedance of a Tektronix DPO4045 Oscilloscope was used as a 50-Ω resistor to ground to measure the current. For all of the following measurements, the devices were first cycled ~1000 times before the actual measurements to prevent any anomalous effects associated with early cycle switching, such as phase-segregation [22].

Several electrical measurements were performed, including a DC voltage sweep to set the devices, an $R–I$ curve showing the resistance and currents measured after applying different voltages, from which the average reset energy was derived. First, the DC sweep was measured to identify the threshold switching voltage. Fig. 4(a) shows an example of the DC sweep performed on the $L_{\text{heater}} = 10 \, \mu \text{m}$ device. The devices showed a threshold voltage between 1 and 2 V. Next, pulsed measurements were performed via gradually increasing voltages, with 0.5 or 0.2 V steps for reset and set. As the voltage pulses were applied, the current during each pulse, $I_{\text{reset}}$, was determined using the following equation:

$$ I_{\text{reset}} = \frac{1}{t_{\text{pulse}}} \int I_{\text{osc}}(t) \, dt $$

(1)

where $I_{\text{osc}}(t)$ was the measured voltage across the 50-Ω oscilloscope resistor as a function of time and $t_{\text{pulse}}$ is the total time the pulse was applied at its maximum value. For the 5/50/5 ns rise/width/fall time pulse, 55 ns was used since $1/2 \times [5-\text{ns rise} + 5-\text{ns fall}] + 50 \, \text{ns pulselength} = 55 \, \text{ns}. Fig. 4(b) shows an example of the resulting currents and DC resistance after pulses were applied on the $L_{\text{heater}} = 10 \, \mu \text{m}$ devices. The devices were initially in the set state ($R_{\text{set}} \leq 200 \, \text{kΩ}$) and switch to the reset state ($R_{\text{reset}} \geq 2 \, \text{MΩ}$) after undergoing the voltage pulses. The $L_{\text{heater}} = 2, 5, 10 \, \mu \text{m}$ DLS devices switch at 94 ± 51, 129 ± 62, and 190 ± 13 µA, while the SiO2 devices switch at 240 ± 83, 329 ± 70, 341 ± 27 µA, respectively (25 measurements on a single devices for each length). The $L_{\text{heater}} = 2 \, \mu \text{m}$ DLS devices were thus able to show current densities of 0.94 ± 0.51 MA/cm².

A write-verify process was used to cycle the devices and identify the reset energy. In order to perform the process, a 2-V reset pulse is first applied. After the reset pulse, the device resistance is measured. If the resistance measured is below the $R_{\text{reset}}$ threshold, then another reset pulse is applied, increasing the previous voltage by 0.5-V increments. The reset pulses are increase until the resistance is above the $R_{\text{reset}}$ threshold. Then, set pulses are applied in a similar manner with 0.8-V starting set voltage and 0.2-V steps until the resistance measured is below the $R_{\text{set}}$ threshold. The process was repeated to cycle the device. Fig. 5(a) and (b) shows the required voltage measured to reset $L_{\text{heater}} = 5 \, \mu \text{m}$ SiO2 and DLS devices. The current driven at reset was measured 25 times randomly selected from 1000 cycles for each of the device dimensions. An example plot of the measured currents and voltages during reset for the $L_{\text{heater}} = 5 \, \mu \text{m}$ DLS and SiO2 devices is shown in Fig. 5(c). There is a strong linear correlation between the reset voltage, $V_{\text{reset}}$, and current, $I_{\text{reset}}$. Thus, the reset current can be written as a function of the reset voltage, $I_{\text{reset}}(V_{\text{reset}})$.

In addition, for a given reset voltage, the current driven does not change significantly between cycles. The variation in reset current is mostly due to the variation in the required voltage to reset. The energy required to reset PCM devices can thus be estimating using

$$ E_{\text{reset}} = V_{\text{reset}} \times I_{\text{reset}}(V_{\text{reset}}) \times t_{\text{pulse}} $$

(2)

The reset energy was measured for DLS and SiO2 samples with $L_{\text{heater}} = 2, 5$, and $10 \, \mu \text{m}$ and is shown in Fig. 6. The DLS devices require ~60% lower energies to switch compared with the SiO2-isolated devices at all dimensions. In addition, the reset energy decreases with the overall length of the device with the $L_{\text{heater}} = 2 \, \mu \text{m}$ devices showing 18.25 ± 15.8 pJ switching energies. The devices were also cycled for endurance testing using the write-verify process.
Fig. 7. Measured endurance of $L_{\text{heater}} = 10\,\mu\text{m}$ (a) SiO$_2$ and (b) DLS PCM devices. A write-verify process was used. Devices with all different $L_{\text{heater}}$ dimensions were able to maintain $10^5$ resistance window for $>10^5$ cycles.

Fig. 8 shows the spatial distribution of the temperature rise for a DLS PCM device ($W_{\text{heater}} = 1\,\mu\text{m}$ and $L_{\text{heater}} = 5\,\mu\text{m}$). The AFM topography map acquired simultaneously determines the exact position of the GST channel and the electrodes (marked with white dashed lines). Evidently, the highest temperature region is locally at the W/GST contact with some temperature change also generated throughout the bulk GST channel. This implies that the phase-change switching region is near the interface, since that is the highest temperature region. In addition, the finite temperature change along the bulk GST implies that significant power is consumed along the bulk GST to heat up the entire region. Since the bulk volume is much larger than the interfacial switching region, significant power is required to reach even a small temperature change of the bulk GST.

2-D FE simulations were performed to better understand the heating and improvement from using the DLS film. The following Fourier heat diffusion equations with TBRs were used to simulate the steady-state heat flow:

$$\nabla \cdot (k \nabla T) + Q_{\text{heat}} = 0 \quad (3)$$

$$Q'' = \frac{\Delta T}{TBR} \quad (4)$$

where (3) was solved at each node with the following node parameters: $k$ is the thermal conductivity, $T$ is the temperature, and $Q_{\text{heat}}$ is the heat generated. For the W/GST interface, $Q''$ was the heat flux across the interface and had a TBR of $24\,\text{m}^2\text{K}/\text{GW}$ [26]. Fig. 9(a) outlines all boundary conditions used for the simulation. The W and GST contacts with the pad region and Si wafer were assumed to be isothermal, since they act as a heat sink/reservoir, with volumes much larger than the regions simulated. The thermal conductivities used were identical to previous work [16], except for the insulator. The insulator thermal conductivities were selected to match their high temperature values of 1.6 and 0.48 W/(m × K) for SiO$_2$ and DLS films, respectively [16].

Joule heating was assumed to be the primary heat source for this simulation. Thus, the heat generated could be broken up into two different heating types based on the most electrically resistive locations: 1) the heat generated at the interface to
Fig. 9. 2-D FE thermal simulation. (a) Simulation boundary conditions and temperature legend. (b)–(e) Resulting simulation temperature profiles. Simulations compared (b) and (c) SiO2 devices and (d) and (e) DLS devices. Devices with different Lheater lengths were simulated, and the normalized temperature profile near the GST/W interface appeared the same, since the heat transfer length is much shorter than the bulk GST and W length. (b) and (d) Simulations assuming all heat generated in the bulk GST. For the same power, that DLS devices generate 2.4× the temperature change of SiO2 devices. (c) and (e) Simulations assuming all heat generated at the GST/W interface. For this case, heating of the interface showed 1.6× the temperature change in DLS devices compared with SiO2 devices.

simulate the effect of contact resistance and 2) the heat generated in the bulk of the phase-change region. In addition, any potential heating from the contact thermoelectric effect could be lumped with the contact resistance heating. Heating may have also occurred in the bulk W, but the measured resistance of W only lines (identical to devices except without the W etch/GST deposition step) was significantly lower than the PCM devices. Also, SThM measurements did not show a temperature change in the bulk W region, thus the energy consumed by the W was likely not significant.

Fig. 9(b) and (d) shows the simulated temperature change due to bulk heating for SiO2 and DLS films, respectively. In the case of bulk heating, most of the heat is dissipated through the insulator stack and thus, the thermal resistance is inversely related to the insulator thermal conductivity. Since the thermal resistance represents the temperature change generated per applied power, this means that the insulator thermal conductivity linearly decreases the energy required to switch. The simulations are in agreement with the analytical estimation, showing a 59% reduction in power to reach the melting temperature for DLS devices compared with the SiO2 devices. This is in good agreement with the measured energy reduction attained between the SiO2 and DLS devices.

Another important detail to note is the decrease in total power required to reach the melting temperature with Lheater length. Fig. 10(a) shows the relative power required to reach the melting temperature in the bulk for various dimensions. Since the volume heated decreases linearly with Lheater length, the required power linearly decreases as well. For interface heating, Fig. 9(c) and (e) show the simulated temperature change due to interface heating for SiO2 and DLS films, respectively. For this case, the heat travels along the GST and W layers while gradually dissipating into the insulator, similar to a thermal fin. A thermal healing length Lthermal can be calculated which is the transfer length of the temperature profile and describes the relationship between the thermal conductivity of the insulator, k_{insulator}, and the metal, k_W, as follows:

\[ L_{\text{thermal}} \propto \sqrt{\frac{k_W}{k_{\text{insulator}}}} \] (5)

Thus, the effective volume heated is larger if the insulator thermal conductivity decreases. As a result, the simulated reduction in power is 36% for DLS devices compared with SiO2 devices. The temperature profile near the GST/W interface was identical for L_{heater} = 2, 5, and 10 µm devices, since the thermal length was \( \sim 100 \text{ nm} \), thus \( L_{\text{thermal}} \ll L_{\text{heater}} \).

In addition, since the volume heated remains constant with device length, the relative power did not depend on L_{heater} as shown in Fig. 10(b). Simulating only interface heating does not agree with the measured results for both reduction in power from changing the insulator and changing L_{heater}. Thus, the required energy to switch is consumed mostly with heating up the bulk, rather than heating the switching region at the interface.

The relative energy generated in the two locations can be approximated by assuming that the current is the same for both the bulk GST and the interface region. Thus

\[ E_{\text{interface}} / E_{\text{bulk}} = R_{\text{interface}}(T_{\text{interface}}) / R_{\text{bulk}}(T_{\text{bulk}}) \] (6)

where \( E_x \) is the energy at location x and \( R_x(T_x) \) is the electrical resistance at temperature T and location x. For the bulk GST region, the resistance decreases over five orders of magnitude as the temperature increases [1]. Thus, the bulk GST region likely consumes a large part of the energy, despite not being the switching region. The interface resistance, however, is less dependent on temperature [27]. Thus, as the bulk heats up and becomes more conductive, gradually more energy is generated at the interface. Therefore, the highest temperature and switching occur at the interface.

The energy required to switch the PCM device can be used as an upper bound for the volume switched. If there is no energy lost to the surroundings, the energy to heat the phase-change layer to the melting temperature and the latent heat of
melting is \( \sim 1.2 \text{ aJ/nm}^3 \) [28]. For 18 pJ, the energy required to switch the DLS \( L_{\text{heater}} = 2 \mu m \) device, the volume melted must be below \( 1.5 \times 10^4 \text{ nm}^3 \) or melted length \( < 1.5 \mu m \) for the devices used in this paper. Thus, it is impossible for the entire bulk GST region to be melted and the GST must be switching at the interface. The devices in this paper show much lower energy and reset current density than other reported devices of similar dimensions (typically \( \sim 10 \) nJ and \( \sim 10 \) MA/cm\(^2\)) [4], [17]. Further thermal confinement in the insulators can reduce the reset energy and will continue to show significant benefit for all device structures, including confined cells where the primary heat loss path is into the insulator [16].

V. Conclusion

This paper utilized the planar confined PCM device to reduce the volume melted, and a thermally resistive multilayer insulator SiO\(_2\)/Al\(_2\)O\(_3\) to achieve low reset energy in PCM devices. Reset energies as low as 18.25 \( \pm 15.8 \) pJ and current densities of 0.94 \( \pm 0.51 \) MA/cm\(^2\) were measured for a contact area of \( A = 10^4 \) nm\(^2\) and heater/GST length of \( L_{\text{heater}} = 2 \mu m \). SThM and Raman mapping suggest that the phase-change region near the heater/GST interface exhibits the highest temperature change and thus represents the switching region. Simulations and analytical models show that most of the energy is used to heat the bulk GST region until the heater/GST contact resistance begins to dominate, at which point the heating and thus switching occurs near the interface. The thermally resistive DLS devices show a strong 60% reset energy reduction and a path to further reduce reset energy via thermal confinement in the insulator.

REFERENCES


Scott W. Fong received the B.S. degree in engineering physics from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2013, where he is currently pursuing the Ph.D. degree, with a focus on phase-change memory device simulation and fabrication to improve performance via thermal design.

Christopher M. Neumann received the B.S. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 2012 and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2015, where he is currently pursuing the Ph.D. degree. His research interests include 2-D materials and resistive switching and phase change memory devices.
Eilam Yalon received the B.S. degree in materials engineering and physics and the Ph.D. degree in electrical engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 2009 and 2015, respectively.

He is currently a Post-Doctoral Researcher with Stanford University, Stanford, CA, USA with a focus on energy efficient memory and computing devices via 2-D, resistive switching, and phase-change materials.

Miguel Muñoz Rojo received the Ph.D. degree in condensed matter physics & nanotechnology and the M.S./B.S. degree in physics from the Autonomous University of Madrid.

He is currently a Post-Doctoral Researcher with Stanford University, Stanford, CA, USA. His current research focus is the transport properties of novel 2-D materials and their applications in electrical and thermal devices.

Eric Pop (M’99–SM’11) received the B.S. and M.S. degrees from the Massachusetts Institute of Technology, Cambridge, MA, USA, and the Ph.D. degree from Stanford University, Stanford, CA, USA.

He is currently an Associate Professor of electrical engineering with Stanford University. His current research interests include energy efficient electronics and data storage, novel 2-D and 1-D devices & materials, and energy conversion and harvesting.

H.-S. Philip Wong (F’01) was with the IBM T. J. Watson Research Center, Yorktown Heights, NY, USA, from 1988 to 2004. He joined Stanford University, Stanford, CA, USA, as a Professor of electrical engineering, in 2004, where he is currently the Willard R. and Inez Kerr Bell Professor. He is the founding Faculty Co-Director of the Stanford SystemX Alliance, an industrial affiliate program focused on building systems.