ENGINEERING

Direct measurement of nanoscale filamentary hot spots in resistive memory devices

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Sanchit Deshmukh[1]†, Miguel Muñoz Rojo[1,2]†, Eilam Yalon[1,3], Sam Vaziri[1], Cagil Koroglu[1], Raisul Islam[1], Ricardo A. Iglesias[1], Krishna Saraswat[1,4,5], Eric Pop[1,4,5]*

Resistive random access memory (RRAM) is an important candidate for both digital, high-density data storage and for analog, neuromorphic computing. RRAM operation relies on the formation and rupture of nanoscale conductive filaments that carry enormous current densities and whose behavior lies at the heart of this technology. Here, we directly measure the temperature of these filaments in realistic RRAM with nanoscale resolution using scanning thermal microscopy. We use both conventional metal and ultrathin graphene electrodes, which enable the most thermally intimate measurement to date. Filaments can reach 1300°C during steady-state operation, but electrode temperatures seldom exceed 350°C because of thermal interface resistance. These results reveal the importance of thermal engineering for nanoscale RRAM toward ultradense data storage or neuromorphic operation.

INTRODUCTION

Future information technologies will need ultrahigh storage densities to process unprecedented amounts of data (*1*, *2*), beyond the capabilities of today's computing systems. Resistive random-access memories (RRAM) (*3*, *4*) promise such densities by being stackable in three dimensions (3D) (*5*) and by storing data in nanoscale (*6*–*8*) conductive filaments (CFs). RRAM can also be used at the heart of neuromorphic computing as a gradually programmable, synapse-like device (*9*–*11*). Typical RRAM cells have a compact crossbar structure ($4F^2$ footprint, where "*F*" is the minimum technology half-pitch) and benefit from low-temperature fabrication, compatible with standard complementary metal-oxide semiconductor (CMOS) technology (*12*). An RRAM cell includes a switching layer, e.g., a metal-oxide like HfO_2, Ta_2O_5 (*3*), or even emerging 2D materials like hexagonal boron nitride (*h*-BN) (*11*) or $MoTe_2$ (*13*), sandwiched between metallic top and bottom electrodes (TE and BE, respectively). RRAM operates through forming, partially breaking (reset), and reconnecting (set) nanoscale CFs with diameters as low as ~7 nm (*6*–*9*) in the switching layer. With enormous power densities (>10^{13} W/cm^3) in such nanoscale volumes, the corresponding temperature rise has been estimated to be as high as 1000 K across multiple studies (*14*–*16*) exclusively through electrothermal models of device behavior.

However, measurements of CF heating in RRAM devices have been very challenging, needing either destructive sample processing or indirect estimation. For instance, postmortem analysis of an RRAM device by transmission electron microscopy (*17*) suggested that the material could have heated to as much as 850 K. Other studies used a microthermal stage (*18*) or pulse-based electrical thermometry (*19*) to indirectly estimate RRAM thermal properties, without spatial resolution. Efforts to spatially resolve the localized heating in resistive memory have used optical techniques that are diffraction-limited and require nonstandard cells (*20*, *21*), or scanning probe techniques without quantifying the temperature (*22*, *23*).

In this work, we quantify individual hot spots in realistic metal-oxide RRAM devices and directly attribute them to Joule heating in sub–10-nm diameter CFs under electrical bias. We achieve nanoscale temperature maps using scanning thermal microscopy (SThM) with a novel calibration approach (*24*), while comparison with detailed simulations reveals that the electrode materials and their thermal coupling with the CF ultimately determine heat spreading in RRAM devices and, thus, thermal cross-talk in RRAM arrays.

RESULTS

Our crossbar RRAM devices use HfO_2 as the switching metal oxide and TEs that are either conventional metals (TiN), single-layer graphene (SLG), or two-layer graphene (2-LG). All devices are capped with a thin layer of Al_2O_3 (see Materials and Methods). The ultrathin graphene TEs are used because they allow the most intimate thermal coupling between the SThM tip and the buried CF, as further described below. As an example, Fig. 1A displays repeatable switching current versus voltage (*I*-*V*) for an RRAM device with SLG as TE, shown in the optical image inset. The CF in this RRAM cell is initially formed at ~4 V under 1-μA current compliance (see Materials and Methods), and Fig. 1B displays >200 switching cycles (also see fig. S1).

Figure 1C shows a schematic of the SThM scanning probe technique (*25*–*28*) that enables temperature measurements with nanoscale resolution, using a sharp V-shaped thermoresistor in direct contact with the sample surface. Our SThM setup can simultaneously map the topography and heating (in terms of the SThM voltage V_{SThM}; see Materials and Methods) at the top surface of the sample under steady-state bias conditions. We compare V_{SThM} scans and quantify top temperature rise (ΔT_S) above the ~20°C ambient for multiple bias conditions, to study self-heating in our RRAM crossbars. Figure 1D shows the topography of the 1.5 × 1.5 μm^2 cell area, while Fig. 1 (E and F) corresponds to V_{SThM} surface maps of the SLG device in the low-resistance state (LRS) with 0- and 90-μW dissipated electrical power (*P*) in the device, respectively.

We detect a single hot spot on the TE surface (Fig. 1F), a clear signature of highly localized Joule heating from the CF. This

[1]Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA. [2]Instituto de Micro y Nanotecnología, IMN-CNM, CSIC (CEI UAM+CSIC), Madrid, Spain. [3]Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa 32000, Israel. [4]Department of Materials Science & Engineering, Stanford University, Stanford, CA 94305, USA. [5]Precourt Institute for Energy, Stanford University, Stanford, CA 94305, USA.
*Corresponding author. Email: epop@stanford.edu
†These authors contributed equally to this work.

Downloaded from https://www.science.org on March 30, 2022

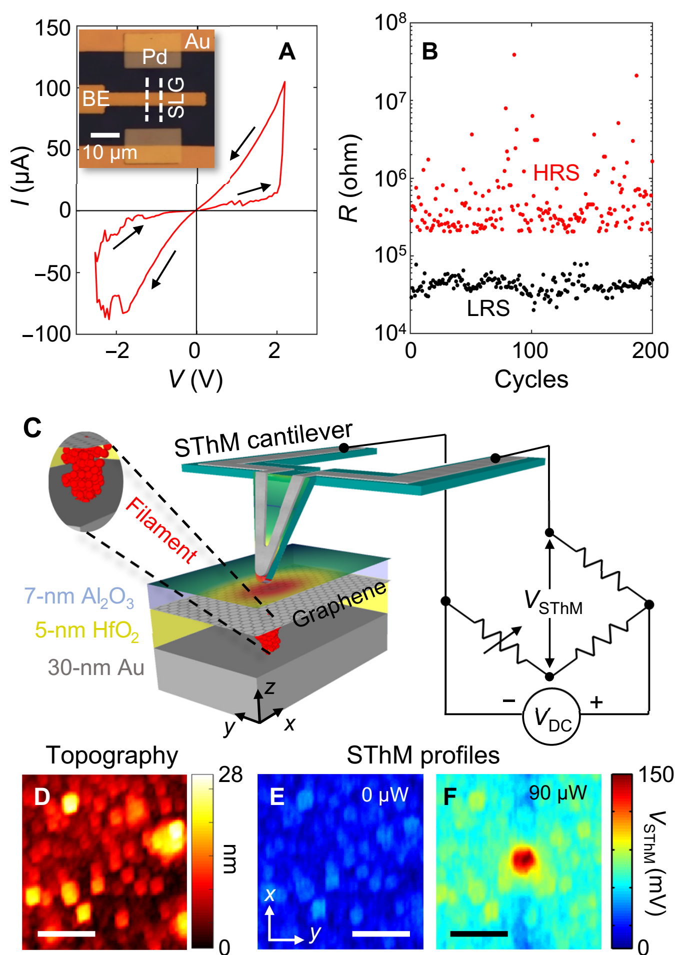

Fig. 1. Graphene-contacted RRAM and SThM. (**A**) Measured *I-V* of RRAM device with SLG as the TE using adaptive pulsed switching. Inset: Optical image of such a device. (**B**) Best devices switched >200 cycles with 1-μs adaptive pulsed switching (see section S1). (**C**) Schematic of SThM measurement on RRAM devices, showing a Wheatstone bridge connected to a typical SThM cantilever. The cantilever scans with a low V_{DC} in physical contact with the top surface of the device. For the axes displayed: *x* and *y* are in and perpendicular to the SThM scan direction, respectively, while *z* is in the vertical direction. (**D**) Topography scan and (**E** and **F**) steady-state SThM on SLG TE device from (A) in the LRS at 0 and 90 μW, respectively. Scale bar, 500 nm.

represents the first direct observation of a CF hot spot with sub–100-nm resolution in a crossbar RRAM device. We note that this surface hot spot is not the same as the highest temperature in the entire device, which is likely within the buried CF. We also note that hot spots are not detected in the high-resistance state (HRS) for devices with higher resistance ratio ($R_{HRS}/R_{LRS} > 10$), further underscoring the CF origin of self-heating. The steady-state power in HRS can be too small to enable any hot spot observation under SThM, and at higher electrical power, the device typically switches to the LRS.

Despite the direct surface measurement, extracting the CF temperature from the measured V_{SThM} remains challenging at first sight. The SThM tip temperature rise ΔT_{tip} is lower than the sample surface temperature rise ΔT_S (*29*), which itself is lower than the temperature rise of the CF, ΔT_{CF}. We first discuss our approach to quantify ΔT_S from V_{SThM}. Thermal coupling of the sample to the SThM tip occurs not only by direct heat conduction but also through convection (*30*) and a water meniscus (*30*, *31*) at the tip-sample interface. The result is a net thermal exchange radius (r_{th}) around the tip-sample contact as shown in Fig. 2A. These effects can be combined through an SThM calibration factor $F(w) = [V_{SThM,\Delta}(w) - V_{SThM,0}(w)]/\Delta T_S$, where the first term is the measured SThM voltage at the center of an isothermal calibration line of width w and temperature ΔT_S, and the second term is the SThM voltage at the same position with $\Delta T_S = 0$.

Because of the nanoscale hot spots in these RRAM devices, calibration of the SThM tip on micrometer-scale isothermal lines is insufficient. Thus, we calibrated the SThM tips on metal heater lines of nominal widths ranging from 50 to 750 nm, as detailed in Materials and Methods and figs. S2 and S3. The extracted calibration factors in Fig. 2B indicate $F(w) \approx 6.5 \pm 1$ mV/K for $w > 200$ nm, while $F(w)$ decreases for $w < 200$ nm, in good agreement with an expected $r_{th} \approx 100$ nm from previous work (*31*). Nearly 46% of heat transfer from sample to SThM tip is through the thermal exchange radius instead of direct conduction. Because the hot spot measured in Fig. 1F was < 200 nm (full width at half maximum) and not an isothermal feature, we implement two deconvolution approaches, simple and Wiener (*32*), to numerically extract arbitrary temperature profiles from V_{SThM} (see section S3 for details), as verified against measured values for $F(w)$ in Fig. 2B. We also validate our SThM calibration by comparison to Raman thermometry on the same device using a single MoS_2 layer transferred on top because of its higher Raman sensitivity (*33*), as shown in Fig. 2C. Good agreement between the Gaussian-averaged SThM temperature and the Raman thermometry data for the same device (Fig. 2D) validates our SThM calibration.

In Fig. 3 (A to D), we reveal how the hot spot generated by the CF changes with different TEs, including TiN/Ti/Pt, TiN, 2-LG, and SLG. A single, nanoscale hot spot imaged in all devices is shown in the corresponding figures. However, the TE vertically and laterally spreads the heat generated by the CF, thus reducing ΔT_S below ΔT_{CF}. With the ultrathin graphene TE, we minimize the temperature drop across the thickness of the TE, allowing the most intimate thermal coupling between the SThM tip and the buried CF. Imaging RRAM devices with TE thickness from 50 to sub–1 nm (Fig. 3, A to D) reveals that the hot spot width appears >120 nm, but the true diameter of the buried CF is expected to be smaller, even below ~10 nm as observed for HfO_2-based RRAM by Celano *et al.* (*6*). The normalized ΔT_S at similar power levels are shown in Fig. 3E for all four device types. The narrowest hot spot is imaged in the device with the thinner TiN electrode (Fig. 3B), confirming more lateral heat spreading in the thicker metal TE (Fig. 3A) and in devices with 2-LG and SLG TE (Fig. 3, C and D), due to the larger thermal healing length (L_H) in these configurations (see section S5).

DISCUSSION

Our observations point to heat spreading in the TE and not the CF diameter, being most important for thermal cross-talk in dense RRAM arrays. Figure 3F displays ΔT_S as a function of applied power. Because of the combined TE and Al_2O_3 capping thickness, the ΔT_S is evaluated ~55 nm above the CF for the device with 50 nm TE and ~7 nm above the CF for the devices with graphene TE (see Materials and Methods). Here, we observe three trends: First, among devices with

Downloaded from https://www.science.org on March 30, 2022

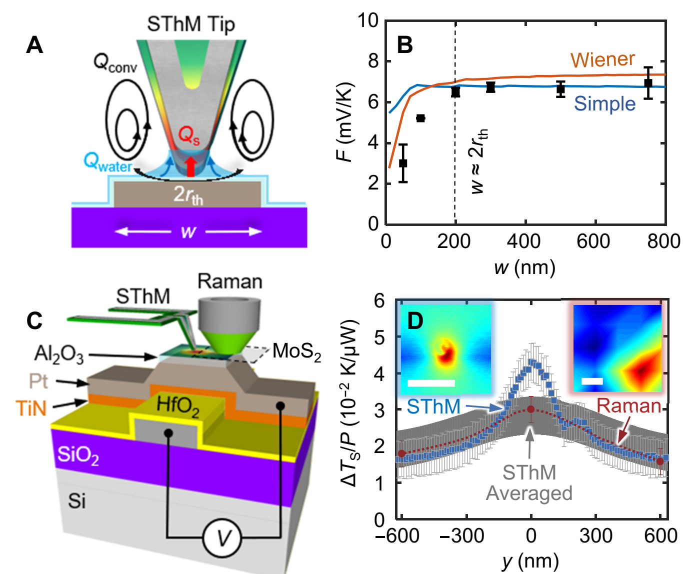

Fig. 2. SThM calibration and validation with Raman thermometry. (**A**) Tip-sample thermal exchange on metal line heaters of different widths (w), showing an effective thermal exchange radius r_{th} due to tip-sample conduction (Q_S) shown in red arrow, by convection (Q_{conv}) shown in black, and by water meniscus (Q_{water}) in blue. (**B**) Extracted calibration factors using SThM scans indicating an $r_{th} \approx 100$ nm by individual linear fits (black, error bars show 95% confidence intervals), simple deconvolution (blue), and Wiener deconvolution (orange). See section S3. (**C**) Combined schematic of SThM and Raman thermometry measurement on same RRAM crossbar with single-layer (1 liter) MoS_2 as Raman-active layer. (**D**) Extracted SThM (blue squares) and Raman (maroon circles) top temperature rise normalized by applied electrical power. The gray "band" replots the SThM data by averaging it across the Gaussian laser spot size (~600 nm), revealing good agreement with the Raman thermometry. Inset images show SThM (left) and Raman ΔT_S (right). Temperature color range for insets is from 0 to 38 K. Scale bars, 750 nm.

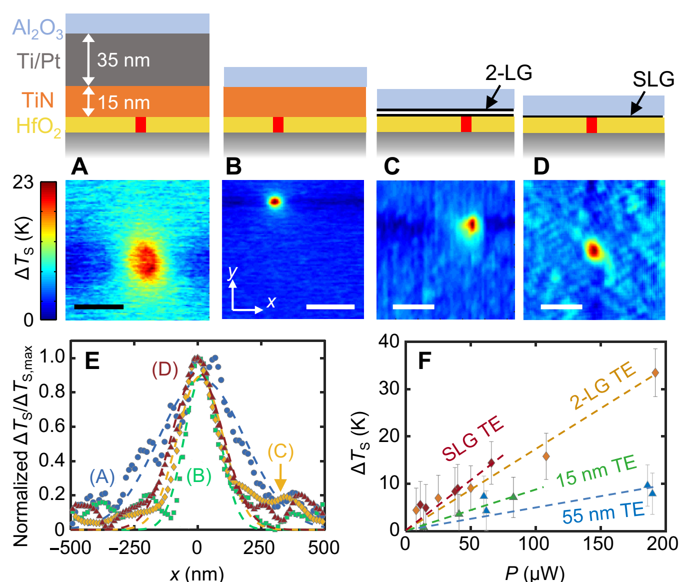

Fig. 3. Comparison between four different RRAM types measured. RRAM device stacks and typical SThM extracted temperature with the following: (**A**) 15-nm TiN/2-nm Ti/33-nm Pt and (**B**) 15-nm TiN TE. For graphene devices, we measured the following: (**C**) 2-LG and (**D**) SLG as TE. All devices are capped by ~7-nm Al_2O_3. Scale bars, 500 nm. (**E**) Normalized ΔT_S across hot spots at 100- to 200-μW electrical power (symbols), corresponding to the different TEs shown in (A) to (D). Each temperature profile is normalized to its peak value, with the baseline subtracted. Dashed lines show Gaussian fits to the corresponding data. Lateral heat spreading is least for the device type in (B), with 15-nm TiN TE. (**F**) Estimated top surface temperature rise versus electrical power for devices with different TE. Proximity of SThM tip to filament has a weak effect on ΔT_S, pointing to high thermal interface resistance in all four cases. Power shown on the x axis is after subtracting the power dissipated in the series resistance and electrodes (see section S10).

graphene TE, the one with 2-LG displays similar ΔT_S within variability as compared with the device with SLG. This is due to slightly better lateral heat spreading in the 2-LG countered by the slightly higher series resistance due to interlayer resistance. Second, among devices with metal TE, the one with 50-nm thickness displays lower ΔT_S than the device with 15-nm metal TE. This is consistent both with better lateral heat spreading and with wider initial CF diameter in the device with 50 nm TE, because of the lower series resistance of this TE (*6*). Third, between the graphene and metal TE devices, our simulations (described below) suggest that effects from different L_H and filament diameter are insufficient to explain their different ΔT_S values. Instead, we find that the thermal boundary conductance (TBC) of the CF-TE interface (*34*), denoted by $G_{CF\text{-}TE}$, and the combined effect of the shape and thermal conductivity of the CF (k_{CF}) itself play important roles in determining the temperature difference between the CF and top surface.

We also perform SThM measurements during device operation, as shown in Fig. 4A, simultaneously with the measured *I-V*, here for an SLG TE for the most intimate coupling between SThM tip and CF. An example of SThM measurement on relatively conductive, ohmic HRS is also seen in Fig. 4A. We clearly observe a transition from HRS to LRS at $\Delta T_S \approx 330$ K (i.e., a surface temperature of ~350°C), with the (buried) ΔT_{CF} expected to be much higher. This observation is not predicted from electrothermal RRAM models in the literature (*15*, *16*) and demands further investigation. This high ΔT_S can elevate neighboring RRAM device temperatures in an array through the electrodes every switching cycle, to result in a substantial array reliability challenge as thermal cross-talk. To gain insight into heat spreading from the CF and to estimate ΔT_{CF}, we perform electrothermal simulations, displayed in Fig. 4B at three different input electrical powers. This electrothermal model is agnostic to electronic transport mechanisms within the CF while attempting to derive a thermal understanding of the RRAM device and CF material stacks to match SThM measurements and measured voltages. We match our simulations to measured ΔT_S profiles at $P \sim 100$ μW (applied during SThM measurement) with CF diameter d_{CF} ~4 nm for the device with TiN TE and ~13 nm for the device with SLG TE, simultaneously fitting the measured electrical resistance of the device as well. Further details are in section S6.

We compare measurements during device operation with our simulations in the LRS, as shown in Fig. 4C. The best agreement is found using a double conical filament shape in our simulations (see inset), consistent with direct measurements on HfO_2 RRAM devices by Celano *et al.* (*6*). Further filament properties and possible shape-related effects are addressed in section S8. The calculated maximum ΔT_{CF} during operation is also plotted in Fig. 4C, reaching up to ~1100 K at the filament constriction. The thermal resistance of the filament-TE interface ($1/G_{CF\text{-}TE}$) contributes more than

Downloaded from https://www.science.org on March 30, 2022

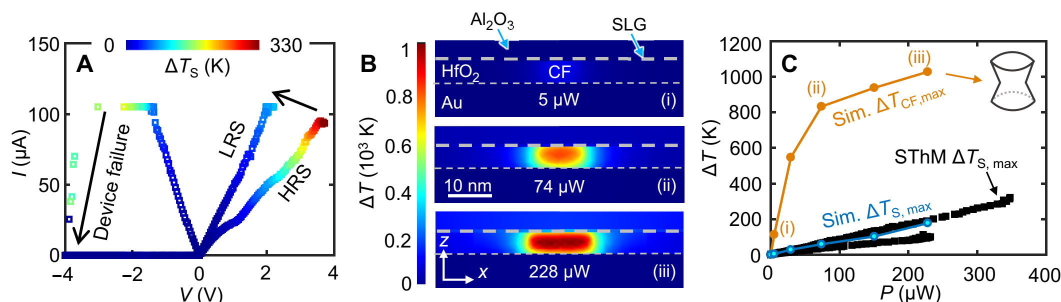

Fig. 4. Maximum filament temperature during operation. (**A**) Simultaneously measured *I-V* and ΔT_S during device operation, here with SLG TE. Symbol colors represent ΔT_S directly at the hot spot above the CF. The SThM is relatively slow compared with the CF time scale, which can cause device failure during this measurement. (**B**) Simulated device temperature cross sections, axisymmetric around the CF, at three different power levels. Temperatures at the top surface of the Al_2O_3 (ΔT_S) and in the filament center (ΔT_{CF}) differ depending on power and TBC of the CF-TE interfaces. (**C**) Maximum surface and filament temperature rise, as a function of input electrical power. Blue symbols are the simulated ΔT_S, in good agreement with the measured values corresponding to both LRS and HRS in (A) (black symbols). Orange symbols display the ΔT_{CF} (i.e., the maximum CF temperature rise). Simulated temperatures of $\Delta T_{CF,max}$ from individual panels in (B) are shown as (i), (ii), and (iii) in (C), respectively.

the thermal resistance of the nanoscale filament (see sections S7 and S8) at all power levels, while the thermal resistance of the TE itself is significantly lower in all cases. $G_{CF\text{-}TE}$ ~75 MW m^{-2} K^{-1} at 550 K for TiN as TE can be estimated by matching the simulations to measured ΔT_S and the electrical resistance and is found to be in the range of typical TiN interfaces (*35*). The same approach finds slightly higher TBC at the CF-SLG interface, and additional details about this thermal analysis are given in sections S8 and S9.

While the mechanism of resistance change in an RRAM device could be caused by changes of defect concentration (*36*) or changes in CF diameter (*37*), both effects require self-heating to cause chemical and structural changes in the CF region. Our work points to the most effective knob to control self-heating: the thermal interfaces to the CF. Lower $G_{CF\text{-}TE}$ increases heat confinement in the CF, creating a larger initial d_{CF} during forming or a higher initial defect density within a similar d_{CF}. RRAM cells with wider or more defect-rich initial CF are more easily programmed into analog memory states (*10*) with varying CF diameter or CF defect density, respectively. On the other hand, more heat confinement during switching provides more power per CF volume, thus making switching more abrupt. Our work shows a first proof of concept that the thermal properties of the CF-TE interface primarily determine heat confinement within a CF and, thus, RRAM device behavior for analog or digital memory operation.

In summary, these represent the first direct measurements of nanoscale hot spots caused by individual filaments in functioning metal-oxide RRAM devices. Using TEs ranging from conventional TiN (~50 nm thick) to SLG (sub–1 nm thick) enabled the most intimate thermal coupling of the SThM with the CF, while elucidating the heat spreading role of the TE. We also uncovered that the thermal resistance of the filament-TE interface is a more important limiter of heat confinement within a CF than the thermal resistance of the TE itself. From simulations, our study reveals a CF diameter of 4 nm with TiN TE and 13 nm with SLG TE at P ~ 100 μW with temperature rise as high as ~1100 K above ambient (i.e., over 1300°C). These results suggest that dense, future RRAM arrays can be made more tolerant to thermal cross-talk by using electrodes with low thermal conductivity. Individual devices could also be made more energy efficient by choosing electrodes with a TBC that is low at the filament-electrode interfaces to confine the CF heating, and high at the surrounding oxide interfaces to minimize lateral heat spreading. Nanoscale thermal engineering at the filament-electrode interfaces could also control heat confinement toward analog versus digital switching in RRAM devices.

MATERIALS AND METHODS

Device fabrication

We use commercially purchased Si wafers coated with thermally grown 30-nm SiO_2 as our starting substrates. Using a lift-off layer (Shipley LOL 1000) and standard Shipley 3612 photoresist-based optical lithography, we first define 5-μm-wide BE strips and contact pad patterns. Then, for fabricating the standard metal-insulator-metal devices, we deposit 30-nm-thick Pt in the defined BE pattern using electron-beam (e-beam) evaporation. For fabricating the devices with graphene as TE (or graphene-insulator-metal devices), we deposit 30-nm-thick Au as the BE material instead, using e-beam evaporation. We dissolve the unexposed photoresist, in each case, using *N*-methyl pyrrolidone (NMP) treatment at 70°C for 25 min, followed by successively rinsing our samples in acetone and isopropanol (IPA) for 2 min each, and drying with an N_2 blow gun. The metal (Pt or Au) coating the unexposed photoresist is lifted-off because of this process, leaving behind the patterned BE metal features. As Au is more ductile compared with Pt, the Au patterns show negligible lift-off edge features after lithography. We then deposit 5-nm-thick HfO_2 using atomic layer deposition (ALD) at 200°C on both sets of samples using tetrakis(dimethylamido)hafnium and deionized (DI) water as precursors.

To fabricate the standard metal-insulator-metal devices, after the HfO_2 ALD step, we use physical vapor deposition to sputter ~15-nm-thick TiN TE directly on the HfO_2 in a blanket manner. The TiN sputter process is done from a Ti target in a 3:1 Ar:N_2 ambient under 65-W dc bias. The TE is deposited in a different tool, within the quickest time allowed by the laboratory layout, i.e., a few minutes.

To fabricate the graphene-insulator-metal devices, after the HfO_2 ALD step, we wet transfer the SLG on top of the blanket HfO_2 by using DI water and chemical vapor deposited (CVD) graphene monolayers. For a subset of these devices, we wet transfer a second layer of graphene in a subsequent step to get devices contacted by 2-LG. The negligible lift-off features for the Au BE edges in this case ensure smooth graphene coverage after the wet transfer.

Downloaded from https://www.science.org on March 30, 2022

For the metal-insulator-metal devices, we pattern 5-μm-wide TE patterns and contact pads on top of the blanket TiN with optical lithography, on one subset as lift-off patterns, and on the remainder as etch patterns. The first subset of the TiN-sputtered samples is top contacted by ~2-nm-thick Ti followed by 33-nm-thick Pt, both deposited sequentially by e-beam evaporation and subsequently lifted off using NMP at 70°C for 25 min, followed by rinsing in acetone and IPA. These samples and the remainder of the TiN-sputtered samples are subjected to a SF_6 reactive ion etch plasma (to etch the remainder of the exposed TiN film), resulting in metal-insulator-metal crossbars. For the graphene-contacted devices, we pattern the SLG and 2-LG into 5-μm-wide TE stripes with optical lithography and a gentle 15-W O_2 plasma dry etch for 30 s. We contact the SLG and 2-LG TE strips with 30-nm-thick Pd on both ends, followed by 2-nm Ti/30-nm Au contact pad lithography. The Pd contacts to graphene are defined away from the crossbar area, so SLG or 2-LG serves as the TE to our RRAM devices, resulting in graphene-insulator-metal crossbars. We thus have four different RRAM crossbars: two as metal-insulator-metal devices—those with 50-nm-thick (Pt/Ti/TiN/HfO_2/Pt) and 15-nm-thick (TiN/HfO_2/Pt) metal TEs, and two as graphene-insulator-metal devices—those with transferred two-layer (2-LG/HfO_2/Au) and single-layer (SLG/HfO_2/Au) graphene TEs.

Last, we cap all our devices with 5-nm Al_2O_3 deposited by ALD at 200°C with trimethylaluminum and DI water as precursors. To ensure good Al_2O_3 coverage on the graphene-insulator-metal devices, we deposit a thin (~1.5 to 1.7 nm thick) Al "seeding" layer using e-beam evaporation, before the ALD step. This seeding layer partially oxidizes post-air exposure and pre-ALD deposition, adding ~2 nm to the ALD Al_2O_3 thickness, thus leading to the eventual ~7-nm-thick AlO_x only on top of the graphene TE devices. The metal-insulator-metal devices have a blanket 5-nm-thick Al_2O_3 layer on top.

For fabricating the metal line heaters in the SThM calibration sample (details of calibration are in section S2), we used 90-nm SiO_2/Si substrates. We first patterned four probe contact pads using optical lithography and lift-off of e-beam evaporated 2-nm Ti/50-nm Pd. We then fabricated heater patterns ranging from 50 to 750 nm with e-beam lithography using poly-methyl methacrylate with molecular weight 950 k in anisole (2% weight/volume) as resist layer. After patterning the metal lines, we deposited 2-nm Ti/30-nm Pd using e-beam evaporation and performed lift-off with NMP, acetone, and isopropanol. The final sample was coated with 10-nm Al_2O_3 deposited via ALD at 200°C.

Sample handling and preparation before SThM

The entire measurement setup and the experimenter follow careful electrical grounding protocols to prevent any electrostatic discharge (ESD) problems before mounting and manipulating the sample and the memory device. For biasing the devices, we used a probe station from Asylum Research and a Keithley 4200 parameter analyzer. The probe station is retrofitted on top of the vibration dampening stage of a MFP-3D atomic force microscopy (AFM) setup from Asylum Research before mounting samples. As-fabricated samples are mounted on this probe station on a glass slide to prevent any electrical leakage into the conductive stage. The sample is secured onto the glass slide with Kapton tape, while the glass slide is secured onto the metallic base of the probe station with permanent magnets to minimize spatial drift during SThM measurements.

Using micromanipulators for the probe station, the measurement probes are positioned within the measurement area of interest. The SThM tip (details further below) is brought into physical proximity of the area of interest by moving the entire stage but still >100 μm above the sample surface. The probes are then brought into physical contact with the device under test (DUT). The ground probe (connected to source-measure unit SMU2) is lowered on the BE pad for the DUT, while the bias probe (connected to SMU1) is lowered on the TE pad.

First, a low-voltage test is performed on the DUT for two reasons: one, to check for probing issues and, two, to ensure that the device is not in a preformed state because of unforeseen fabrication or ESD issues. For this test, a slow, dual-voltage sweep (from 0 to 2.5 V) is applied to the bias probe, and current measured at the ground probe is observed. If subpicoampere current is observed with negligible dependence on bias voltage, the micromanipulators are used to lower probes by a few micrometers to overcome any probing issue. If high current (>100 pA) is observed, the DUT is considered damaged, and micromanipulators are used to move to a different DUT. As-fabricated DUTs that show current in the range 10 to 20 pA with an increasing dependence on voltage within this voltage range are considered for the second step.

Second, the SThM tip is lowered for a topography measurement on the top surface of the DUT. This step ensures a relatively flat topography before electrical measurements, indicating a suitable device for SThM measurements with minimal topography artifacts and the absence of any surface features from ESD damage during preparation and handling.

Third, a forming and preliminary switching step is performed on the DUT that has passed the first two steps. For this purpose, we use a 10-kilohm off-chip series resistor for metal (TiN/Ti/Pt) TE devices to prevent current compliance overshoot and subsequent device breakdown during forming. For the TiN, SLG, and 2-LG TE devices, the on-chip series resistance from the respective TEs prevents current overshoot. Further forming details for devices shown in this study are in section S1. An intermediate topography check is performed after the forming step to confirm no topography damage due to current overshoot.

Fourth, the SThM tip is configured for SThM measurements and scanned over the device surface to ensure steady-state thermal signal from the setup before device heating measurements. RRAM device power across all measurements is calculated as shown in section S10.

SThM measurement setup

The scanning thermal microscope is an add-on from Anasys Instruments that was added to the MFP-3D AFM from Asylum Research. The probes were purchased from Anasys Instruments and consist of a thin Pd line on Si_3N_4. Each probe is connected to one of the arms of a Wheatstone bridge. This bridge is initially matched; an increase in temperature at the SThM probe tip causes a change in the resistance of the Pd line and subsequently a voltage mismatch in the bridge. This voltage change across the Wheatstone bridge is read out as the SThM voltage signal (V_{SThM}). A higher V_{SThM} indicates a higher SThM probe resistance and, qualitatively, a higher temperature at the SThM tip. To electrically isolate our device from the Pd line of the SThM tip, we cap our devices with Al_2O_3 as shown in the schematic in Fig. 1C. During all our measurements, we apply the same force to the surface of the devices and account for the thermal drift (see section S3). We used a gain of 1000× while extracting tip

Downloaded from https://www.science.org on March 30, 2022

voltage change and a relative deflection set point of −0.5 V. Images were analyzed with MATLAB and Asylum Research software.

For steady-state measurements, the SThM tip moved with a scan rate of ≈0.7 to 0.8 Hz over a ≈5-μm scan line (corresponding to width of the scan image), while the device biasing voltage is held constant. Each scan line has 256 pixels. A total of 256 such scan lines taken in succession led to ≈5-μm length of the scan image.

For measurements during device operation, the SThM tip is stationary, in contact with the sample surface, while a slow (0.1 V step every 0.5 s) voltage sweep is applied to the device, to allow ample time for the sample + tip system to thermally equilibrate. The SThM tip is positioned at the observed location of the hot spot from prior steady-state scans of the same device to consequently observe ΔT_S directly above the CF.

Raman measurement setup

Raman spectroscopy was carried out using a Horiba LabRam instrument with a 532-nm laser and 100× long working distance objective with numerical aperture of 0.6 and gratings of 1800 gr/mm. Step sizes in the Raman maps were 0.25 μm, and the acquisition time of device thermal map was ~20 min. The laser spot radius is ~0.3 μm, and with 4-s acquisition time at each point with 5% laser power, the incident laser power is 150 μW (<10-μW absorbed power) to avoid laser heating in excess of the electrical heating. Temperature calibration was done with a Linkam THMS600 stage, following the procedure outlined elsewhere (*33*).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abk1514

REFERENCES AND NOTES

1. M. M. S. Aly, M. Gao, G. Hills, C. S. Lee, G. Pitner, M. M. Shulaker, T. F. Wu, M. Asheghi, J. Bokor, F. Franchetti, K. E. Goodson, C. Kozyrakis, I. Markov, K. Olukotun, L. Pileggi, E. Pop, J. Rabaey, C. Ré, H.-S. P. Wong, S. Mitra, Energy-efficient abundant-data computing: The N3XT 1,000x. *Computer* **48**, 24–33 (2015).
2. H. S. P. Wong, S. Salahuddin, Memory leads the way to better computing. *Nat. Nanotechnol.* **10**, 191–194 (2015).
3. H. S. P. Wong, H. Y. Lee, S. Yu, Y. S. Chen, Y. Wu, P. S. Chen, B. Lee, F. T. Chen, M. J. Tsai, Metal-oxide RRAM. *Proc. IEEE* **100**, 1951–1970 (2012).
4. S. Yu, Resistive random access memory (RRAM). *Synth. Lect. Hum. Lang. Technol.* **2**, 1–79 (2016).
5. M. Yu, Y. Cai, Z. Wang, Y. Fang, Y. Liu, Z. Yu, Y. Pan, Z. Zhang, J. Tan, X. Yang, M. Li, R. Huang, Novel vertical 3D structure of TaO_x-based RRAM with self-localized switching region by sidewall electrode oxidation. *Sci. Rep.* **6**, 21020 (2016).
6. U. Celano, L. Goux, R. Degraeve, A. Fantini, O. Richard, H. Bender, M. Jurczak, W. Vandervorst, Imaging the three-dimensional conductive channel in filamentary-based oxide resistive switching memory. *Nano Lett.* **15**, 7970–7975 (2015).
7. D.-H. Kwon, K. M. Kim, J. H. Jang, J. M. Jeon, M. H. Lee, G. H. Kim, X.-S. Li, G.-S. Park, B. Lee, S. Han, M. Kim, C. S. Hwang, Atomic structure of conducting nanofilaments in TiO_2 resistive switching memory. *Nat. Nanotechnol.* **5**, 148–153 (2010).
8. C. Pan, Y. Ji, N. Xiao, F. Hui, K. Tang, Y. Guo, X. Xie, F. M. Puglisi, L. Larcher, E. Miranda, L. Jiang, Y. Shi, I. Valov, P. C. McIntyre, R. Waser, M. Lanza, Coexistence of grain-boundaries-assisted bipolar and threshold resistive switching in multilayer hexagonal boron nitride. *Adv. Funct. Mater.* **27**, 1604811 (2017).
9. S. Choi, S. H. Tan, Z. Li, Y. Kim, C. Choi, P.-Y. Chen, H. Yeon, S. Yu, J. Kim, SiGe epitaxial memory for neuromorphic computing with reproducible high performance based on engineered dislocations. *Nat. Mater.* **17**, 335–340 (2018).
10. W. Wu, H. Wu, B. Gao, N. Deng, S. Yu, H. Qian, Improving analog switching in HfO_x-based resistive memory with a thermal enhanced layer. *IEEE Electron Dev. Lett.* **38**, 1019–1022 (2017).
11. Y. Shi, X. Liang, B. Yuan, V. Chen, H. Li, F. Hui, Z. Yu, F. Yuan, E. Pop, H.-S. P. Wong, M. Lanza, Electronic synapses made of layered two-dimensional materials. *Nat. Electron.* **1**, 458–465 (2018).
12. H. Chen, S. Yu, B. Gao, P. Huang, J. Kang, H. S. P. Wong, HfO_x based vertical resistive random access memory for cost-effective 3D cross-point architecture without cell selector, in *IEEE International Electron Devices Meeting (IEDM)* (2012), pp. 20.27.21–20.27.24.
13. I. M. Datye, M. M. Rojo, E. Yalon, S. Deshmukh, M. J. Mleczko, E. Pop, Localized heating and switching in $MoTe_2$-based resistive memory devices. *Nano Lett.* **20**, 1461–1467 (2020).
14. D. Ielmini, Resistive switching memories based on metal oxides: Mechanisms, reliability and scaling. *Semicond. Sci. Technol.* **31**, 063002 (2016).
15. A. Padovani, L. Larcher, O. Pirrotta, L. Vandelli, G. Bersuker, Microscopic modeling of HfO_x RRAM operations: From forming to switching. *IEEE Trans. Electron Dev.* **62**, 1998–2006 (2015).
16. S. Menzel, M. Waters, A. Marchewka, U. Böttger, R. Dittmann, R. Waser, Origin of the ultra-nonlinear switching kinetics in oxide-based resistive switches. *Adv. Funct. Mater.* **21**, 4487–4492 (2011).
17. J. Kwon, A. A. Sharma, C.-Y. Chen, A. Fantini, M. Jurczak, A. A. Herzing, J. A. Bain, Y. N. Picard, M. Skowronski, Transient thermometry and high-resolution transmission electron microscopy analysis of filamentary resistive switches. *ACS Appl. Mater. Interfaces* **8**, 20176–20184 (2016).
18. Z. Jiang, Z. Wang, X. Zheng, S. Fong, S. Qin, H. Chen, C. Ahn, J. Cao, Y. Nishi, H. S. P. Wong, Microsecond transient thermal behavior of HfOx-based resistive random access memory using a micro thermal stage (MTS), in *IEEE International Electron Devices Meeting (IEDM)*. (2016), pp. 21.23.21–21.23.24.
19. A. A. Sharma, M. Noman, M. Skowronski, J. A. Bain, High-speed in-situ pulsed thermometry in oxide RRAMs, in *Proceedings of Technical Program - 2014 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)*. (2014), pp. 1–2.
20. S. Kumar, Z. Wang, N. Davila, N. Kumari, K. J. Norris, X. Huang, J. P. Strachan, D. Vine, A. L. D. Kilcoyne, Y. Nishi, R. S. Williams, Physical origins of current and temperature controlled negative differential resistances in NbO_2. *Nat. Commun.* **8**, 658 (2017).
21. E. Yalon, S. Deshmukh, M. M. Rojo, F. Lian, C. M. Neumann, F. Xiong, E. Pop, Spatially resolved thermometry of resistive memory devices. *Sci. Rep.* **7**, 15360 (2017).
22. J. M. Goodwill, G. Ramer, D. Li, B. D. Hoskins, G. Pavlidis, J. J. McClelland, A. Centrone, J. A. Bain, M. Skowronski, Spontaneous current constriction in threshold switching devices. *Nat. Commun.* **10**, 1628 (2019).
23. E. Puyoo, D. Albertini, Conductive filament localization within crossbar resistive memories by dcanning joule expansion microscopy. *IEEE Electron Dev. Lett.* **41**, 848–851 (2020).
24. S. Deshmukh, M. M. Rojo, E. Yalon, S. Vaziri, E. Pop, Probing Self-Heating in RRAM Devices by Sub-100 nm Spatially Resolved Thermometry, in *76th Device Research Conference (DRC)*. (2018), pp. 1–2.
25. Y. Zhang, W. Zhu, F. Hui, M. Lanza, T. Borca-Tasciuc, M. M. Rojo, A review on principles and applications of scanning thermal microscopy (SThM). *Adv. Funct. Mater.* **30**, 1900892 (2019).
26. F. Menges, P. Mensch, H. Schmid, H. Riel, A. Stemmer, B. Gotsmann, Temperature mapping of operating nanoscale devices by scanning probe thermometry. *Nat. Commun.* **7**, 10874 (2016).
27. A. A. Wilson, M. M. Rojo, B. Abad, J. A. Perez, J. Maiz, J. Schomacker, M. Martín-Gonzalez, D.-A. Borca-Tasciuc, T. Borca-Tasciuc, Thermal conductivity measurements of high and low thermal conductivity films using a scanning hot probe method in the 3ω mode and novel calibration strategies. *Nanoscale* **7**, 15404–15412 (2015).
28. M. M. Rojo, S. Grauby, J. M. Rampnoux, O. Caballero-Calero, M. Martin-Gonzalez, S. Dilhaire, Fabrication of Bi_2Te_3 nanowire arrays and thermal conductivity measurement by 3ω-scanning thermal microscopy. *J. Appl. Phys.* **113**, 054308 (2013).
29. L. Shi, S. Plyasunov, A. Bachtold, P. L. McEuen, A. Majumdar, Scanning thermal microscopy of carbon nanotubes using batch-fabricated probes. *Appl. Phys. Lett.* **77**, 4295–4297 (2000).
30. K. Kim, J. Chung, G. Hwang, O. Kwon, J. S. Lee, Quantitative measurement with scanning thermal microscope by preventing the distortion due to the heat transfer through the air. *ACS Nano* **5**, 8700–8709 (2011).
31. E. Puyoo, S. Grauby, J.-M. Rampnoux, E. Rouvière, S. Dilhaire, Thermal exchange radius measurement: Application to nanowire thermal imaging. *Rev. Sci. Instrum.* **81**, 073701 (2010).
32. N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications* (MIT press, Cambridge, MA, 1964), vol. 8.
33. E. Yalon, C. J. McClellan, K. K. H. Smithe, M. M. Rojo, R. L. Xu, S. V. Suryavanshi, A. J. Gabourie, C. M. Neumann, F. Xiong, A. B. Farimani, E. Pop, Energy dissipation in monolayer MoS_2 electronics. *Nano Lett.* **17**, 3429–3433 (2017).
34. E. Yalon, A. A. Sharma, M. Skowronski, J. A. Bain, D. Ritter, I. V. Karpov, Thermometry of filamentary RRAM devices. *IEEE Trans. Electron Devices* **62**, 2972–2977 (2015).
35. J. P. Reifenberg, K. Chang, M. A. Panzer, S. Kim, J. A. Rowlette, M. Asheghi, H. P. Wong, K. E. Goodson, Thermal boundary resistance measurements for phase-change memory devices. *IEEE Electron Dev. Lett.* **31**, 56–58 (2010).

Downloaded from https://www.science.org on March 30, 2022

36. A. Marchewka, B. Roesgen, K. Skaja, H. Du, C.-L. Jia, J. Mayer, V. Rana, R. Waser, S. Menzel, Nanoionic resistive switching memories: On the physical nature of the dynamic reset process. *Adv. Electronic Mater.* **2**, 1500233 (2016).
37. R. Degraeve, A. Fantini, N. Raghavan, L. Goux, S. Clima, Y. Y. Chen, A. Belmonte, S. Cosemans, B. Govoreanu, D. J. Wouters, P. Roussel, G. S. Kar, G. Groeseneken, M. Jurczak, Hourglass concept for RRAM: A dynamic and statistical device model, in *Proceedings of the 21th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*. (2014), pp. 245–249.
38. A. Kalantarian, G. Bersuker, D. C. Gilmer, D. Veksler, B. Butcher, A. Padovani, O. Pirrotta, L. Larcher, R. Geer, Y. Nishi, P. Kirsch, Controlling uniformity of RRAM characteristics through the forming process, in *2012 IEEE International Reliability Physics Symposium (IRPS)*. (2012), pp. 6C.4.1-6C.4.5.
39. A. Banerjee, H. Satoh, A. Tiwari, C. Apriono, E. T. Rahardjo, N. Hiromoto, H. Inokawa, Width dependence of platinum and titanium thermistor characteristics for application in room-temperature antenna-coupled terahertz microbolometer. *Jpn. J. Appl. Phys.* **56**, 04CC07 (2017).
40. E. Pop, Energy dissipation and transport in nanoscale devices. *Nano Res.* **3**, 147–169 (2010).
41. M. K. Samani, X. Z. Ding, N. Khosravian, B. Amin-Ahmadi, Y. Yi, G. Chen, E. C. Neyts, A. Bogaerts, B. K. Tay, Thermal conductivity of titanium nitride/titanium aluminum nitride multilayer coatings deposited by lateral rotating cathode arc. *Thin Solid Films* **578**, 133–138 (2015).
42. W. Jang, Z. Chen, W. Bao, C. N. Lau, C. Dames, Thickness-dependent thermal conductivity of encased graphene and ultrathin graphite. *Nano Lett.* **10**, 3909–3913 (2010).
43. R. E. Taylor, J. Morreale, Thermal conductivity of titanium carbide, zirconium carbide, and titanium nitride at high temperatures. *J. Am. Ceram. Soc.* **47**, 69–73 (1964).
44. S. Vaziri, E. Yalon, M. Muñoz Rojo, S. V. Suryavanshi, H. Zhang, C. J. McClellan, C. S. Bailey, K. K. H. Smithe, A. J. Gabourie, V. Chen, S. Deshmukh, L. Bendersky, A. V. Davydov, E. Pop, Ultrahigh thermal isolation across heterogeneously layered two-dimensional materials. *Sci. Adv.* **5**, eaax1325 (2019).
45. P. Sun, N. Lu, L. Li, Y. Li, H. Wang, H. Lv, Q. Liu, S. Long, S. Liu, M. Liu, Thermal crosstalk in 3-dimensional RRAM crossbar array. *Sci. Rep.* **5**, 13504 (2015).
46. V. E. Dorgan, M.-H. Bae, E. Pop, Mobility and saturation velocity in graphene on SiO_2. *Appl. Phys. Lett.* **97**, 082112 (2010).
47. M. A. Panzer, M. Shandalov, J. A. Rowlette, Y. Oshima, Y. W. Chen, P. C. McIntyre, K. E. Goodson, Thermal properties of ultrathin hafnium oxide gate dielectric films. *IEEE Electron Dev. Lett.* **30**, 1269–1271 (2009).
48. G. Chen, P. Hui, Thermal conductivities of evaporated gold films on silicon and glass. *Appl. Phys. Lett.* **74**, 2942–2944 (1999).
49. R. M. Costescu, M. A. Wall, D. G. Cahill, Thermal conductance of epitaxial interfaces. *Phys. Rev. B* **67**, 054302 (2003).
50. J. C. Duda, C. Y. P. Yang, B. M. Foley, R. Cheaito, D. L. Medlin, R. E. Jones, P. E. Hopkins, Influence of interfacial properties on thermal transport at gold:silicon contacts. *Appl. Phys. Lett.* **102**, 081902 (2013).
51. D. H. Hurley, M. Khafizov, S. L. Shinde, Measurement of the Kapitza resistance across a bicrystal interface. *J. Appl. Phys.* **109**, 083504 (2011).
52. D. Villaroman, X. Wang, W. Dai, L. Gan, R. Wu, Z. Luo, B. Huang, Interfacial thermal resistance across graphene/Al_2O_3 and graphene/metal interfaces and post-annealing effects. *Carbon* **123**, 18–25 (2017).
53. Z.-Y. Ong, M. V. Fischetti, A. Y. Serov, E. Pop, Signatures of dynamic screening in interfacial thermal transport of graphene. *Phys. Rev. B* **87**, 195404 (2013).
54. S. Deshmukh, E. Yalon, F. Lian, K. E. Schauble, F. Xiong, I. V. Karpov, E. Pop, Temperature-dependent contact resistance to nonvolatile memory materials. *IEEE Trans. Electron Dev.* **66**, 3816–3821 (2019).
55. S. Deshmukh, R. Islam, C. Chen, E. Yalon, K. C. Saraswat, E. Pop, Thermal modeling of metal oxides for highly scaled nanoscale RRAM, in *2015 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. (2015), pp. 281–284.
56. D. Niraula, V. G. Karpov, Heat transfer in filamentary RRAM devices. *IEEE Trans. Electron Dev.* **64**, 4106–4113 (2017).
57. K. T. Regner, J. A. Malen, Nondiffusive thermal transport increases temperature rise in RRAM filaments. *IEEE Electron Dev. Lett.* **37**, 572–575 (2016).
58. S. Larentis, F. Nardi, S. Balatti, D. C. Gilmer, D. Ielmini, Resistive switching by voltage-driven ion migration in bipolar RRAM—Part II: Modeling. *IEEE Trans. Electron Dev.* **59**, 2468–2475 (2012).
59. N. Stojanovic, D. H. S. Maithripala, J. M. Berg, M. Holtz, Thermal conductivity in metallic nanostructures at high temperature: Electrons, phonons, and the Wiedemann-Franz law. *Phys. Rev. B* **82**, 075418 (2010).
60. W. Wu, H. Wu, B. Gao, P. Yao, X. Zhang, X. Peng, S. Yu, H. Qian, A Methodology to Improve Linearity of Analog RRAM for Neuromorphic Computing, in *2018 IEEE Symposium on VLSI Technology*. (2018), pp. 103–104.

Acknowledgments: Work was performed in part at the Stanford Nanofabrication Facility and the Stanford Nano Shared Facilities, which receive funding from the National Science Foundation as part of the National Nanotechnology Coordinated Infrastructure Award ECCS-1542152. S.D. acknowledges support from Stanford Nonvolatile Memory Technology Research Initiative (NMTRI). S.D. acknowledges R.B. Lockwood for feedback on the manuscript, A. Sood for discussions on thermal interface insights, and G. Wetzstein for feedback on the deconvolution approach. **Funding:** This work was supported in part by ASCENT, one of six centers in JUMP, a Semiconductor Research Corporation (SRC) program sponsored by DARPA. **Author contributions:** S.D. and E.P. wrote the manuscript. S.D., M.M.R., E.Y., and E.P. designed the experiments. S.D. and M.M.R. performed SThM measurements on RRAM. M.M.R. performed SThM measurements on the calibration sample. R.A.I. contributed to calibration measurements. M.M.R. and E.Y. performed electrical thermometry of the calibration sample. S.D. and S.V. fabricated the devices and the calibration sample. S.D., M.M.R., and E.Y. performed the initial analysis to extract calibration factors. S.D. conceptualized and implemented the deconvolution approach for calibration, with contribution from C.K. E.Y. and S.D. built the COMSOL model and performed simulations. E.Y., S.D., and S.V. performed Raman thermometry measurements. S.D., R.I. and K.S. worked on the thermal conductivity estimation. **Competing Interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 24 June 2021
Accepted 4 February 2022
Published 30 March 2022
10.1126/sciadv.abk1514

Downloaded from https://www.science.org on March 30, 2022

Science Advances

Direct measurement of nanoscale filamentary hot spots in resistive memory devices

Sanchit DeshmukhMiguel Muñoz RojoEilam YalonSam VaziriCagil KorogluRaisul IslamRicardo A. IglesiasKrishna SaraswatEric Pop

Sci. Adv., 8 (13), eabk1514. • DOI: 10.1126/sciadv.abk1514

View the article online
https://www.science.org/doi/10.1126/sciadv.abk1514
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org on March 30, 2022

Use of this article is subject to the Terms of service

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.
Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Science Advances
AAAS

Supplementary Materials for

Direct measurement of nanoscale filamentary hot spots in resistive memory devices

Sanchit Deshmukh, Miguel Muñoz Rojo, Eilam Yalon, Sam Vaziri, Cagil Koroglu, Raisul Islam, Ricardo A. Iglesias, Krishna Saraswat, Eric Pop*

*Corresponding author. Email: epop@stanford.edu

Published 30 March 2022, *Sci. Adv.* **8**, eabk1514 (2022)
DOI: 10.1126/sciadv.abk1514

This PDF file includes:

Sections S1 to S10
Figs. S1 to S9
Tables S1 to S3
References

S1. RRAM Switching

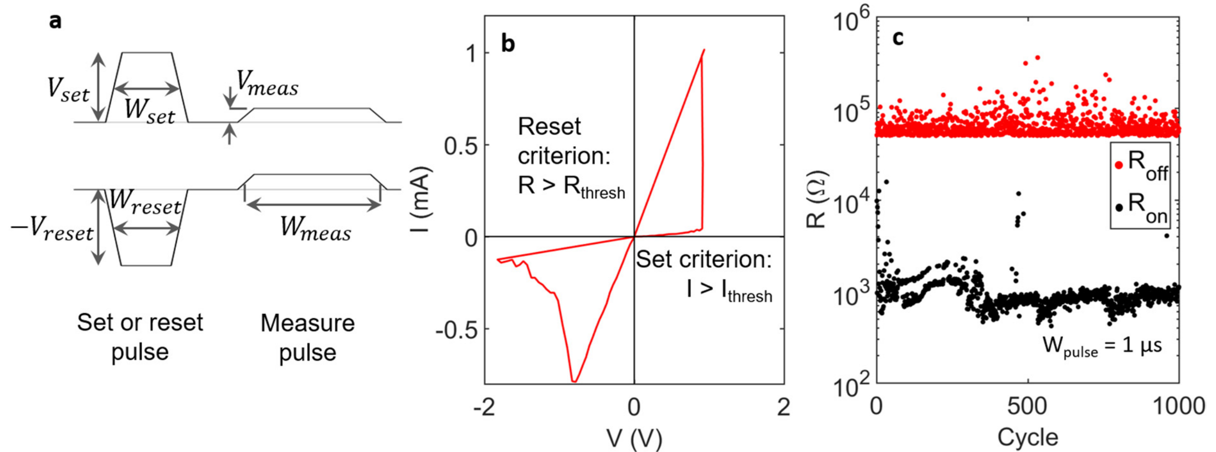

Fig. S1. Pulsed switching of RRAM devices. (a) Switching pulse routine consisting of a set or reset pulse of programmable voltage (V_{set} or -V_{reset}), and pulse widths (W_{set} or W_{reset}). This is followed by a measure pulse of low voltage (V_{meas} ~ 0.25 V) and longer width (W_{meas} ~ few 100 µs). **(b)** Switching in typical metal-HfO_2-metal crossbar (as discussed in Figs. 1(A-B) for SLG TE) using pulse routine from panel (a). Every data point is a pulsed measurement. The reset criterion ($R > R_{thresh}$) or set criterion ($I > I_{thresh}$) is checked in the measure pulse, depending on the current state of the device. **(c)** Switching in the same device as panel (b), showing an on/off ratio ~ 50 for 1000 cycles, with $W_{set} = W_{reset}$ = 1 µs.

In order to switch our fabricated resistive random access memory (RRAM) devices repeatedly, we use an adaptive write-verify pulsing approach. After the initial forming step of our RRAM device-under-test (DUT) using a DC bias, the DUT is in the low resistance state (LRS). The DC bias formed our device under the effect of the current compliance as determined partially by the series resistance from the top electrode (TE). Time-to-form was determined to not be crucial (*38*) in controlling subsequent device switching. In order to reset these devices to the high resistance state (HRS), we apply a voltage pulse with a fixed width (W_{reset}) and a tunable amplitude (-V_{reset}). After successfully resetting the DUT, we apply a set pulse with a fixed width (W_{set}) and varying amplitude (V_{set}) until the DUT is back in the LRS. After every set or reset pulse, we apply a wider (pulse width ~ W_{meas}) measure pulse of amplitude (V_{meas}) ~ 0.25 V to probe the current state of the DUT (whether in LRS or HRS). The generic set, reset and corresponding measure pulses are shown in fig. S1(a). In order to determine the state of the device – HRS or LRS, we define a reset and set criterion in our pulsing control script, respectively. In the specific case shown in fig. S1(b) (and in Figs. 1(A-B)), we use a set criterion $I > I_{thresh}$ (= 100 µA for Fig. 1B and 1 mA for fig. S1(c)) and a reset criterion $R > R_{thresh}$ (= 200 kΩ for Fig. 1B and 50 kΩ for fig. S1(c)), implemented in the Keithley 4200 instrument (see Methods) with custom-written C code.

A current lower limit as the set criterion helps immediately identify sharp current rise in the device during the measure pulse, while the resistance threshold in the HRS case lets us define a lower limit on the reset resistance. With this approach, we can achieve > 200 cycles of switching in the single layer graphene (SLG) top electrode (TE) RRAM device in Fig. 1B, and > 1000 cycles of switching in our TiN/Pt TE device in fig. S1(c). It is important to note that the series resistance from the graphene prevents current overshoot during the forming process for the SLG TE device. We can switch similar metal-contacted devices repeatably using adaptive pulsed switching measurements.

S2. Scanning Thermal Microscopy (SThM) Calibration

To calibrate the conversion factor between the SThM tip voltage signal (V_{SThM}) and the temperature rise of our sample surface (ΔT_S), we used a calibration sample with a set of Ti/Pd heater lines of varying widths from w = 50 nm to 750 nm. The calibration procedure included a measurement of the temperature coefficient of resistance (TCR) of the Ti/Pd heaters capped with thin Al_2O_3, just like our RRAM devices. This Al_2O_3 capping is needed to block electrical leakage between the SThM tip and the sample.

We perform 4-probe electrical measurement on every metal line of different w. These measurements are performed at 5 background temperatures on a heated electrical stage (T_{stage}). The electrical data from a typical measurement for a line with w = 500 nm is shown in fig. S2(a). The 4-probe electrical resistance R_0 of the metal heater (at a given w) at zero power (i.e. no Joule heating, so $T = T_{stage}$,) is extracted for each T_{stage} and allows us to estimate the w-dependent TCR (denoted below as α) as:

$$\alpha = \frac{R_0(T = T_{stage} + \Delta T) - R_0(T = T_{stage})}{\Delta T \times R_0(T = T_{stage})} = \frac{1}{R_0} \times \frac{\partial R_0}{\partial T_{stage}} \tag{1}$$

The $R_0 w$ vs. T_{stage} plot is shown in fig. S2(b) for all widths. The TCR values are plotted in fig. S2(c), with 95% confidence intervals. The TCR is lower for lower w, due to line edge roughness scattering (*39*).

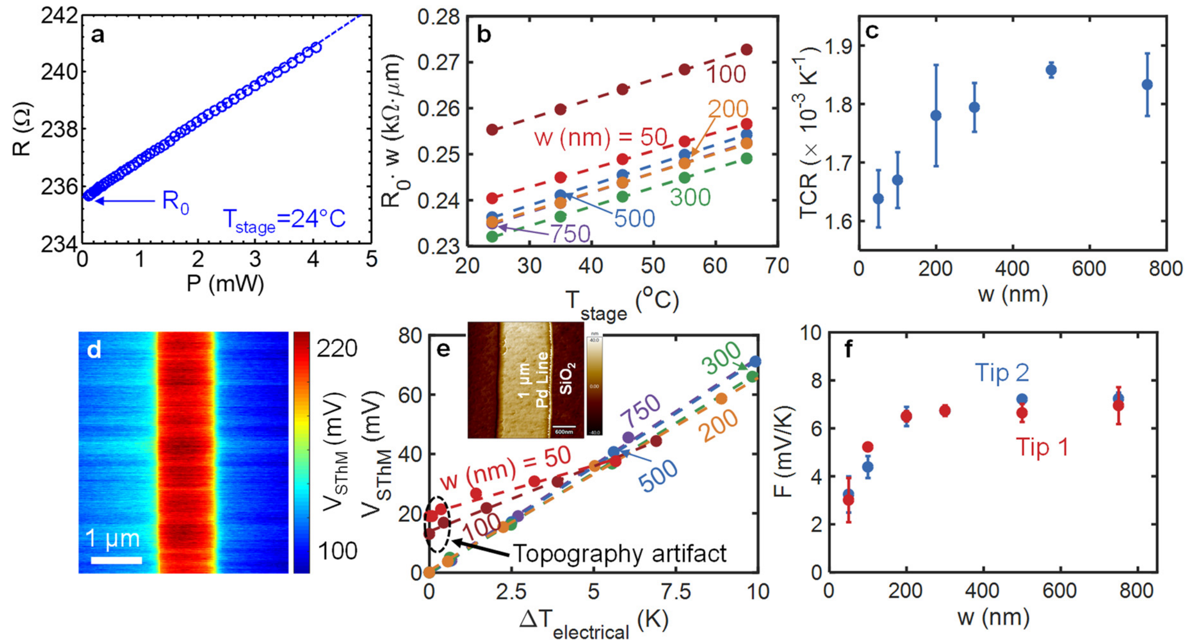

Fig. S2. SThM probe calibration with metal heaters. (a) Four-probe resistance measurement of 500 nm wide Ti/Pd heater line at stage temperature T_{stage} = 24 °C. The y-intercept of the linear fit to the data at P = 0 mW is R_0. **(b)** R_0 times heater width (w) measured at five different T_{stage} for all w fabricated. **(c)** Extracted temperature coefficient of resistance (TCR) for all heaters with different w. Uncertainty in w is 5 nm. **(d)** Typical SThM scan of metal heater at non-zero power. **(e)** Measured V_{SThM} at center of every line of different w plotted vs. estimated ΔT from electrical measurement. At w < 200 nm, the slope of the linear fit decreases (due to tip-sample heat exchange) and the y-intercept increases (due to topography artifacts from the line edge). **(f)** Extracted calibration factor F (= slope of linear fit in panel e) for different w for two different SThM tips from different batches using the same calibration sample.

Next, using the TCR for every line, we extract the metal line temperature ($\Delta T_{\text{electrical}}$) which depends on the Joule heating input power of the line. We extract $\Delta T_{\text{electrical}}$ values for > 4 electrical bias conditions for all widths using eq. 2:

$$\Delta T_{electrical}(P) = \frac{1}{\alpha}\left(\frac{R_P}{R_0} - 1\right) \quad (2)$$

Here, $\Delta T_{\text{electrical}}(P)$ is the estimated temperature rise for input electrical power P and R_P is the measured 4-probe resistance of the heater line at power P. We then measure $V_{\text{SThM}}(w)$ for the same input power P on the metal lines (fig. S2(d)) and extract $F(w)$ as the slope of the linear fit to the V_{SThM} *vs*. $\Delta T_{\text{electrical}}$ plot for every w, shown in fig. S2(e). $F(w)$ is expected to remain constant for the SThM tip across different samples with the same thin Al_2O_3 capping layer as the calibration sample. For narrow line widths in fig. S2(e), two separate trends are seen: first the slope is lower due to reduced heat transport from the sample surface to the tip (as a result of the thermal exchange radius r_{th}) and second the y-intercept of the plot increases to a non-zero V_{SThM} due to the topography effect from the line edges in every thermal map. The obtained factors $F(w)$ for two SThM tips, from separate batches procured commercially, are shown in fig. S2(f). To make our calibration useful for arbitrary temperature profiles and to remove the effect of r_{th} and topography features for narrow heaters measured above, we used a deconvolution approach described in Section S3.

S3. SThM Resolution Limitations and Deconvolution Approach

Because the electrical data obtained from near-isothermal metal heater lines have limited applicability, we model the tip voltage dependence on arbitrary temperature profiles using two deconvolution approaches. At the same time, because the topography features of our heater lines caused their own parasitic footprint, we decouple the effects of topography within the same deconvolution approach. We assumed the tip-sample interface to be a linear, spatially invariant (LSI) heat exchange system. Two enforced measurement constraints were important in justifying this assumption – first, the tip deflection (electronically controlled by tuning the laser photodetector output to -0.5 V), scan speed (set to 0.7 Hz for every 2 to 6 μm scan line with 256 pixels, more than few ms per pixel) and voltage gain (1000x) were kept constant across *all* scans shown in this work; second, the top sample material in direct physical contact with the SThM tip in every case was Al_2O_3 deposited at the same atomic layer deposition (ALD) process conditions, i.e. 200°C using trimethylaluminum (TMA) and deionized (DI) water as precursors in the same tool, as described in the Methods section. The SThM measurement ambient had a temperature close to room temperature (~20°C) and similar humidity levels for most scans. The effects of different scan rates are not explored in this work, i.e. we assume the effect of the scan rate does not interfere with our analysis because the scan rate remains the same, ~ 0.7 Hz, throughout our work.

We let the initial thermal drift (with the initial heating of the tip to its steady state upon starting the setup) in the SThM signal stabilize over the course of 5 dummy scans before acquiring data. Subsequent measurements on the three narrowest lines are shown in figs. S3(a-t) below. Note that figs. S3(b, h, n) show artificially elevated SThM signal at the line edges due to SThM tip movement over the changing topography, despite 0 μW applied electrical power in those panels. Figure S3(t) shows that our measurement can resolve signals from edges of the narrowest line. These images and data show that our measurement resolution is sub-100 nm, which we further refine using a deconvolution approach as detailed below.

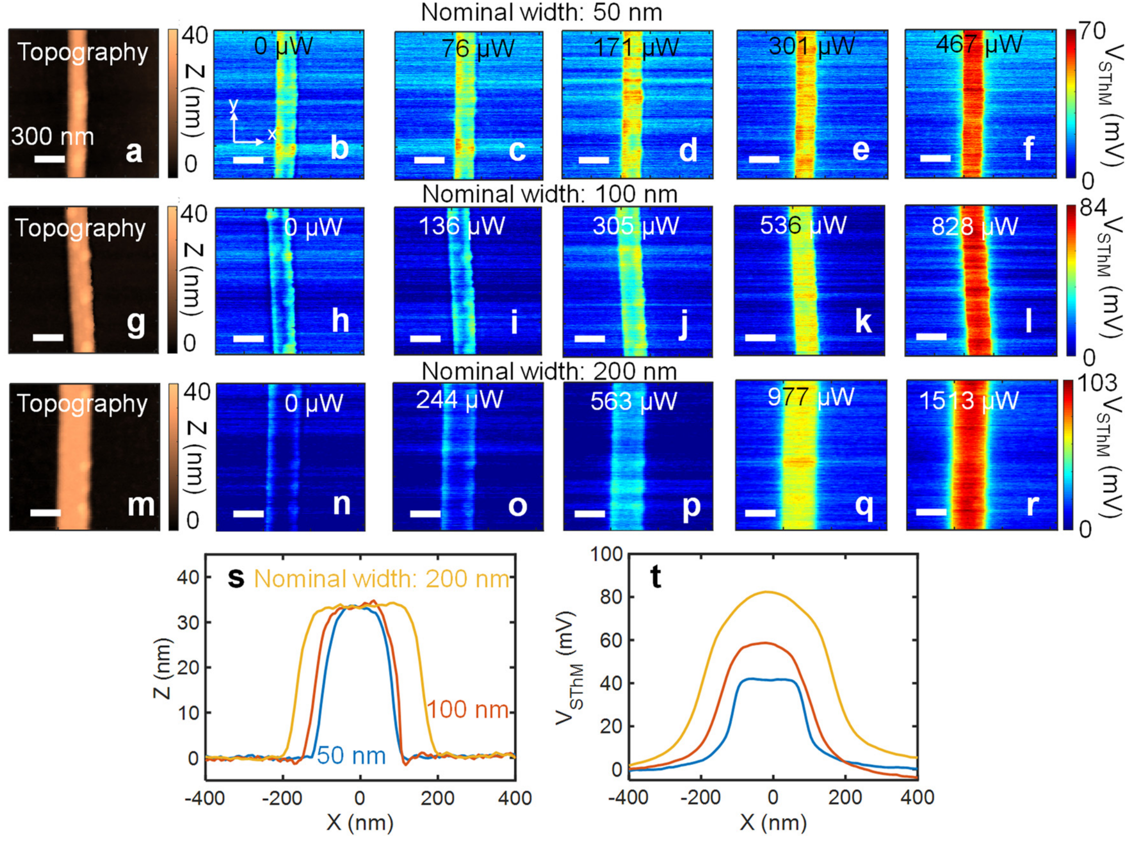

Fig. S3. SThM Measurements on metal lines with nominal width ~ 200 nm or less. (a-f) Topography and SThM scans for lines with nominal width 50 nm, **(g-l)** 100 nm, and **(m-r)** 200 nm. Apparent width can be higher for atomic force microscopy measurements in contact mode. Scale bar: 300 nm. **(s)** Topography scan line for all 3 nominal widths. Lines appear wider in the topography scans, which are acquired with the SThM cantilever in contact mode, so the width of the tip convolves with the actual line width to show an apparently wider line. **(t)** SThM signal at highest power for 3 lines, from panels (f, l, r) as blue, red and orange lines respectively, averaged over the length of the line measured.

For the first deconvolution approach (also referred to here as a "simple approach"), we assumed the following relation between the V_{SThM} scan, topography (or z scan) and actual sample temperature above ambient (or ΔT values) using eq. 3 below:

$$V_{SThM}(x,y) = g_T(x,y) * \Delta T(x,y) + g_z(x,y) * z(x,y) \quad (3)$$

Here, x and y are the spatial coordinates at the top surface (in the plane of the measurement), g_T is a calibration matrix that is the spatial heat exchange footprint of the tip-sample interface and g_z is a calibration matrix that is the spatial footprint of the topography parasitic in the corresponding V_{SThM} scan. We also evaluate an inverse relation which models the SThM voltage contributing to the sample temperature as:

$$\Delta T(x,y) = g_V(x,y) * V_{SThM}(x,y) + g_z'(x,y) * z(x,y) \quad (4)$$

Here g_V is a similar calibration matrix modeling the effect from V_{SThM} and g_z' is a corresponding matrix for topography parasitics. We note from these two separate solutions that the tip-sample heat exchange, in our scans (leading to a change in the V_{SThM} values), is from the nearest 4-5 pixels on either side for a 1.5 × 1.5

μm scan (corresponding to a heat exchange diameter ~ 100 nm). However, with this approach, our LSI assumption is likely to be too strong for the SThM-sample interface under consideration. The LSI assumption's validity is further challenged in the case of topography parasitics for e.g. near metal line edges. At the same time, to achieve a good fit to the $F(w)$ values, we need multiple scans on different line widths to estimate the best deconvolution solution (yielding least error across the entire measurement set).

To overcome these limitations, we also use a second deconvolution in tune with a more familiar Wiener approach. Within this approach, we model the tip-sample heat exchange as a point spread function (PSF) around the tip-sample contact: $g_{T,W}(x, y)$. This PSF can be extracted by comparing ΔT_{top} for our Ti/Pd heaters against corresponding SThM scans, on a pixel-by-pixel basis. A PSF obtained using a simple inverse filtering of the heater line temperature shows a good fit to the measured calibration factors at line widths > 200 nm. With a Wiener deconvolution approach, we treat interference from topography changes at the line edges as a correlated noise in the thermal signal. We use the SThM scan on every line at 0 W electrical power applied as our noise prior image: $n_w(x, y)$ for that line. The final temperature maps are calculated by noise removal and PSF deconvolution performed in the Fourier domain using 2-dimensional (2D) fast Fourier transform (FFT), in accordance with the set of equations shown below:

$$F_{VSThM} = FFT\left[V_{SThM}\left(x,y\right)\right]; F_N = FFT\left[n_w\left(x,y\right)\right]; F_{GT} = FFT\left[g_{T,W}\left(x,y\right)\right] \tag{5}$$

For every element in row i and column j, corresponding to an individual FFT frequency component, we then calculate the power spectral density:

$$PSD_{VSThM,ij} = f_{VSThM,ij} \times f'_{VSThM,ij}; PSD = f_{N,ij} \times f'_{N,ij} \tag{6}$$

Here, $PSD_{VSThM,ij}$ and $PSD_{N,ij}$ are the values of the power spectral density for the frequency in the ith row and j^{th} column for the V_{SThM} scan being evaluated and the corresponding noise prior image respectively, $f_{VSThM,ij}$ and $f_{N,ij}$ are the elements in the ith row and j^{th} column in the matrices F_{VSThM} and F_N respectively, and $f_{VSThM,ij}'$ and $f_{N,ij}'$ are their respective complex conjugates. The Wiener matrix elements are defined according to eq. 7 below:

$$W_{ij} = \frac{1}{f_{GT,ij}} \times \frac{PSD_{VSThM,ij}}{PSD_{VSThM,ij} + PSD_{N,ij}} \tag{7}$$

Here W_{ij} is the Wiener matrix element in the ith row and j^{th} column, $f_{GT,ij}$ is the element in the ith row and j^{th} column for the matrix F_{GT} defined above in eq. 5 and other symbols are as defined above. The calculation of the corresponding temperature map is carried out in accordance with eq. 8 below:

$$f_{T,ij} = F_{VSThM,ij} \times W_{ij} \tag{8}$$

Here, $f_{T,ij}$ is the element in the ith row and j^{th} column for the matrix F_T corresponding to the FFT of the calculated temperature map $T(x, y)$. Our calculated $T(x, y)$ now simply corresponds to the 2D inverse-FFT of F_T. Note that the system of eqs. 5-8 is solvable for either $f_{T,ij}$ or $f_{GT,ij}$, provided that one of these two parameters is known for every (i, j). In order to calculate the value of $g_{T,W}(x, y)$ for a particular SThM tip, we solve this system of equations after the measurements performed in Supplementary Information Section S2. In order to calculate RRAM device top surface temperatures, we then use the best fitted $g_{T,W}(x, y)$ from this procedure to calculate $T(x, y)$ for each case.

S4. Raman Thermometry

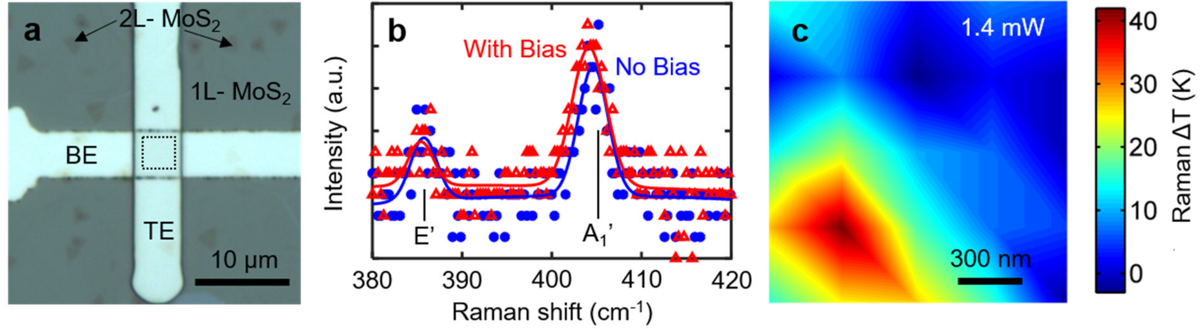

Fig. S4. Raman measurements on MoS_2-covered RRAM device. (a) Optical image of a single RRAM crossbar with MoS_2 transferred on top. The MoS_2 is largely single-layered (1L-MoS_2), with some bilayer regions (2L-MoS_2) visible. **(b)** Measured Raman signal on top of MoS_2 without (blue filled circles) and with (red empty triangles) 1.4 mW electrical bias to the formed filament in the device in panel (a). Lines show Gaussian peaks fitted to Raman data. A red shift in the A_1' and E' peak positions with bias indicates MoS_2 heating. **(c)** Measured and extracted temperature of the MoS_2 layer using the procedure outlined by Yalon *et al.* (*33*) for an applied electrical power of 1.4 mW to the RRAM device.

Because Raman thermometry with our laser has a diffraction-limited spatial resolution ≈ 600 nm, we need a wider hot spot than the one measured in Fig. 1F for comparison. We use the increased heat spreading in an RRAM device with a thick TE (schematic shown in Fig. 2C) to compare temperatures between SThM and Raman thermometry. Because neither the metals nor the oxides employed in switching (HfO_2) and capping (Al_2O_3) have a strong Raman signal, we use a single layer of MoS_2 (*33*) as an ultrathin Raman transducer, transferred on top of our complete device stack as shown in fig. S4(a). The corresponding Raman signal measured with and without electrical bias is shown in fig. S4(b). The data show signatures of A_1' and E' peaks in the single layer MoS_2 on top of our RRAM device with and without electrical bias. The data are somewhat noisy because signal from MoS_2 on metal is poor, and Raman laser power was kept low at 5% (corresponding to ≈ 150 μW incident or < 10 μW absorbed power) to minimize optical heating of the device and MoS_2. Upon fitting Gaussians to estimate peak position, we observe a red shift in peak positions (both A_1' and E') with 1.4 mW electrical power, indicating a temperature rise in the MoS_2 due to the Joule heating in the filament of the RRAM device underneath. A quantitative extraction of temperature is performed by the procedure outlined in previous work (*33*).

The Raman thermometry map on the same device at an electrical power of 1.4 mW shows a Raman temperature of 48 K above ambient at the top surface as shown in fig. S4(c). This is about 2× the electrical power than the SThM measurement in Fig. 2D.

S5. Thermal Healing Length and Hot Spot Width

To understand the heat spreading seen in our devices, we note that the hot spot full width at half maximum (FWHM) is approximately the sum of d_{CF} and twice the TE thermal healing length (L_H). The L_H refers to the distance from the hot spot at which the temperature decays to 1/e of the peak temperature, as (*40*):

$$L_H = \sqrt{\frac{k_{th,TE} \times t_{TE}}{G_{TE-ox}}} \tag{9}$$

where $G_{TE\text{-}ox}$ is the thermal boundary conductance (TBC) between the TE and the underlying HfO_2, while $k_{th,TE}$ is the lateral TE thermal conductivity as shown in fig. S5(a). The t_{TE} is the thickness of the TE, e.g. 0.34 nm for SLG, 0.68 nm for 2-LG, and 15 nm for TiN. With different hot spot widths seen for different TE devices, as plotted in fig. S5(b), we can estimate interfacial properties using eq. 9. (We note these hot spot widths are obtained from the calculated temperature profiles, not the raw SThM data.)

For our estimates, we use $k_{th,TE} \approx 11$ $Wm^{-1}K^{-1}$ for TiN (*41*) and $k_{th,TE} \approx 300$ $Wm^{-1}K^{-1}$ for oxide-encased SLG (*42*). The thermal conductivity for TiN thin films was reported between 11 $Wm^{-1}K^{-1}$ and 25 $Wm^{-1}K^{-1}$ (*43*). The electrical conductivity for our TiN was in the range 0.5 to 1.25 × 10^5 S/m, suggesting an electronic contribution to thermal conductivity < 1 $Wm^{-1}K^{-1}$ using the Wiedemann-Franz Law. Because our measured electrical conductivity is near the low end for TiN thin films (ostensibly due to grain boundary and surface scattering in the ~15 nm thin films), we take the thermal conductivity near the low end of the range above.

Given that RRAM filaments are extremely narrow ($d_{CF} \ll L_H$), the hot spot FWHM ≈ $2L_H$, allowing us to estimate $G_{TE\text{-}ox} \approx 50$ $MWm^{-2}K^{-1}$ for the TiN-HfO_2 interface and ≈ 9 $MWm^{-2}K^{-1}$ for the SLG-HfO_2 interfaces in Fig. 3B and 3D, respectively. By this approach, we underestimate $G_{TE\text{-}ox}$ due to the assumption that the thermal resistance of the device stack is dominated by the TE-oxide interface, so it is not surprising that

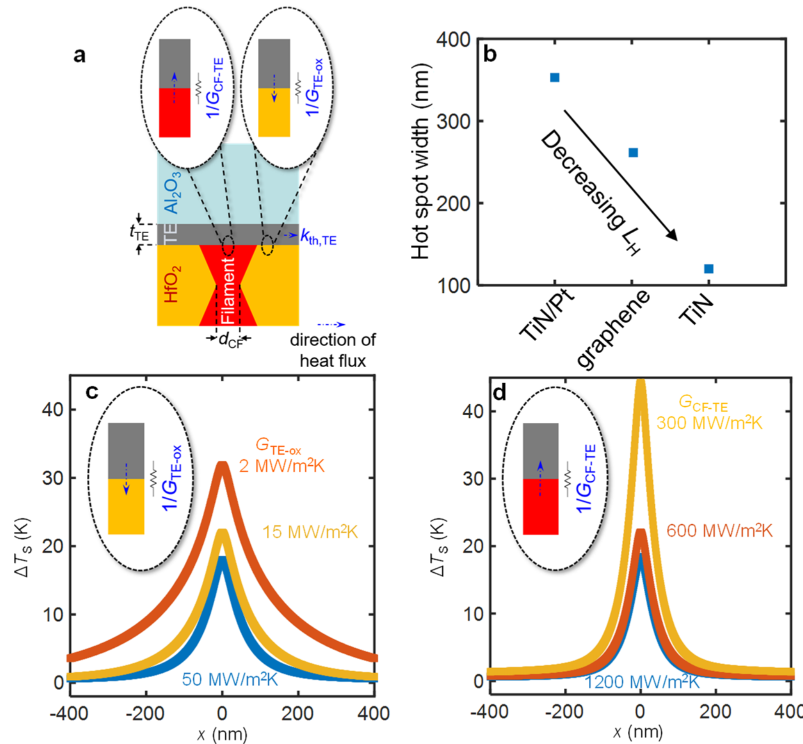

Fig. S5. TBC considerations for hot spot width. (a) Schematic of a single RRAM filament contacted by top electrode (TE), capped with Al_2O_3. The TBC at the TE-filament ($G_{CF\text{-}TE}$) interface, and the TE-HfO_2 (surrounding the filament) interface ($G_{TE\text{-}ox}$), thermal conductivity ($k_{th,TE}$) and filament diameter (d_{CF}) determine heat spreading. **(b)** Measured surface hot spot width vs. TE type (corresponding to Fig. 3E) showing least heat spreading with TiN TE. **(c-d)** Electro-thermal simulations of top temperature profiles (ΔT_S) for SLG TE devices (at the same applied electrical power) show changing the assumed TBCs alters ΔT_S and heat spreading. In (c), making the TE-oxide interface less thermally conductive broadens the hot spot FWHM and increases ΔT_S. In (d), making the TE-filament interface more thermally conductive leads to lower hot spot temperatures due to less heat confinement within the filament.

these TBC values are slightly lower than previous independent measurements of TiN and SLG interfaces (*35, 44*). Nevertheless, because the hot spot widths correlate with the thermal conductivity and inversely with the TBC of the TE, these results highlight the importance of choosing a TE with low $k_{th,TE}$ and high G_{TE-ox}, to minimize thermal crosstalk (*45*) in a highly scaled RRAM array. For example, the simulated hot spot can be wider or narrower depending on the G_{TE-ox} [simulated curves in fig. S5(c)]. The TBC at the filament-TE interface (G_{CF-TE}) is also important, as seen in fig. S5(d). These simulations are described in detail in the following sections. We also address the effects of a non-negligible d_{CF} and different $k_{th,TE}$ in Supplementary Information Sections S8 and S9, respectively.

S6. Electro-Thermal Finite-Element Model for RRAM Devices

The finite-element method (FEM) model for extracting filament temperature rise (ΔT_{CF}) by modeling electrical heating in an RRAM device is implemented in COMSOL. For the purposes of this work, we implement a 4 × 4 × 20 μm model, as shown in fig. S6(a). The structure is assumed to be cylindrically symmetric about the central axis of the filament, as indicated in fig. S6(a). The model geometry has open boundary conditions for heat flux at all surfaces, with equations:

$$T = T_{amb},\ if\ \vec{\phi}\cdot\vec{u} < 0 \text{ and } -\nabla T\cdot\vec{u} = 0,\ if\ \vec{\phi}\cdot\vec{u} \geq 0$$

Here, $\vec{\phi}$ is the heat flux vector, and $\vec{u}$ is a unit vector perpendicular to the surface pointing outwards from the geometry. In effect, if there is a net inward flow of heat, the surface is assumed to be at ambient T_{amb} =

Table S1. COMSOL model simulation parameters assumed across multiple devices simulated.

Parameter	Value at Room Temperature	Temperature dependence
σ_{Au}	4.56×10^7 S/m	
σ_{Pt}	8.9×10^6 S/m	
σ_{TiN}	1.25×10^5 S/m	
σ_{SLG}	10^7 S/m at $n \sim 10^{12}$ cm^{-2} (*46*)	
$k_{th,c\text{-}HfO2}$ (at T = 296 K)	1.2 Wm^{-1}K^{-1}	See Supplementary Information S7
$\sigma_{c\text{-}HfO2}$	10^{-15} S/m	
$k_{th,a\text{-}HfO2}$ (at T = 296 K)	0.4 Wm^{-1}K^{-1} (*47*)	
$\sigma_{a\text{-}HfO2}$	10^{-15} S/m	
$k_{th,Pt}$	50 Wm^{-1}K^{-1}	T^0
$k_{th,Au}$	158 Wm^{-1}K^{-1} (*48*)	T^0
$k_{th,SiO2}$	1.4 Wm^{-1}K^{-1}	T^0
$k_{th,Si}$	150 Wm^{-1}K^{-1}	T^0
$k_{th,Al2O3}$	3 Wm^{-1}K^{-1}	T^0
$k_{th,TiN}$	11 Wm^{-1}K^{-1} (Section 5)	T^0
$k_{th,SLG}$	300 Wm^{-1}K^{-1} (*42*)	T^0
$G_{TiN\text{-}Al2O3}$	150 MWm^{-2}K^{-1} (*49*)	T
$G_{Au\text{-}SiO2} \sim G_{Pt\text{-}SiO2}$	50 MWm^{-2}K^{-1} (*50*)	
$G_{SiO2\text{-}Si}$	434 MWm^{-2}K^{-1} (*51*)	T
$G_{SLG\text{-}Al2O3}$	150-200 MWm^{-2}K^{-1} (*52, 53*)	
$R_{C,CF\text{-}TE}$	10^{-9} Ω-cm^2 (*54*)	
$G_{TE\text{-}ox}$	50 MWm^{-2}K^{-1} (for TE: TiN); 9 MWm^{-2}K^{-1} (for TE: SLG, see Section 5)	T

296 K. If there is a net outward flow of heat, the change in temperature in a direction perpendicular to the surface is assumed to be 0 K. Also, only the temperature of the bottommost surface is fixed at T_{amb} = 296 K. These boundary conditions are most realistic for our device, since we expect negligible heat conduction through the SThM tip in contact with our samples and through convection, compared to the total heat generated. In order to simulate Joule heating in our geometry (described above), we perform a complete electro-thermal simulation. The geometry also assumes that all surfaces are electrically insulating, except for the TE and BE edges. The TE edge is assumed to be a current source with a fixed DC value of current I_m (see fig. S6(a)). The BE edge is assumed to be the ground (GND) terminal, hence we forced the potential at that terminal to 0 V. The current I_m, in every case is at the edge of the top electrode in our model, with the bottom electrode being electrically grounded. In order to fit to the measured electrical data, we change filament level parameters in the model until the simulated voltage (V_{sim}) at the edge of the top electrode is

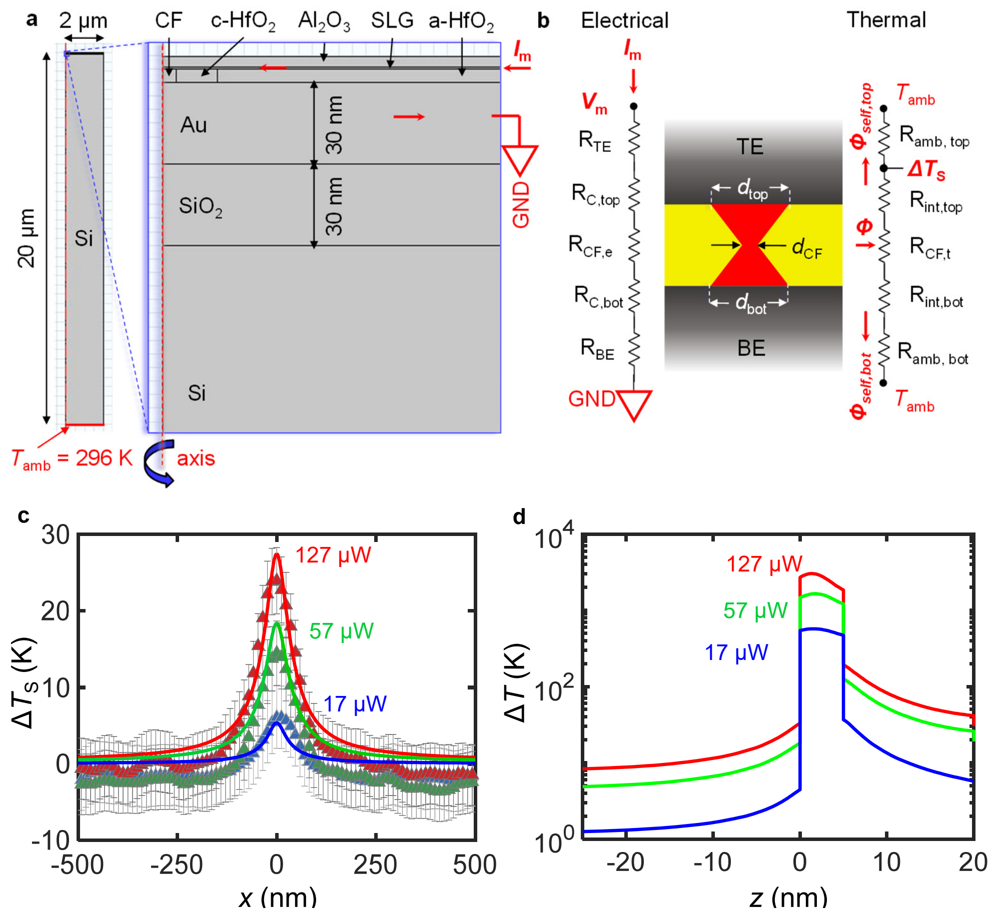

Fig. S6. FEM Simulation and electro-thermal model. (a) Cross-section of axisymmetric model used in FEM analysis for a 4×4×20 μm cell, with open thermal boundary conditions (see text) and fixed ambient temperature T_{amb} enforced 20 μm below the device. The boundaries are assumed to be electrically insulating, except for the TE (current source I_m) and BE (= 0 V) boundaries. Also shown is a cross-section of the geometry near the CF. A ~20 nm wide crystalline HfO_2 region is assumed concentric to the CF, in accordance with recent proposal in literature (*17*). **(b)** Electrical circuit, CF schematic and thermal circuit shown in a typical simulation. **(c)** Fitted top temperature profiles from FEM simulations (lines) for the same electrical power as the SThM measurements (symbols) for a 15 nm TiN TE device. **(d)** Temperature along the axis of the model in the vertical (z) direction for the fitted cases in panel (c).

equal to the measured voltage (V_m) in our SThM scans at the particular value of I_m. The parameters of interest are described further below.

We assume a cylindrical, hourglass-shaped filament geometry with a constriction in this model. The corresponding fig. S6(b) shows a schematic for the filament and all interfaces considered for the purposes of this model. The main parameters of interest for the filament are d_{CF}, d_{top}, d_{bot}, σ_{CF}, $k_{th,CF}$, $G_{CF\text{-}TE}$ and $G_{CF\text{-}BE}$. Out of these, we first adjust d_{CF}, d_{top} and d_{bot} for different assumed values of σ_{CF}, thus modeling for $R_{CF,e}$ (or filament electrical resistance) in fig. S6(b). The precise values of these parameters that show best agreement depend on the filament geometry and sub-stoichiometry. However, because the exact geometry and sub-stoichiometry are extremely challenging to determine even with destructive imaging of the device, our assumption of a double conical shape with a constriction is the most generic and valid assumption, taking into account recent results on filament shapes (*6*). We assume an electrical contact resistance $R_{C,top} = 10^{-9}$ $\Omega\cdot cm^2 \times \pi d_{top}^2/4$ in accordance with estimates of filament contact resistivity in previous work (*54*). We point out that this is an empirical approach because we are attempting to fit the measured electrical resistance and the measured SThM temperature profiles at the same time. The material parameters assumed for FEM simulations are summarized in Table S1. Note that we assume a T^0 dependence of the thermal conductivity for metal electrodes, which is a reasonable approximation.

From our complete thermal model for filament thermal conductivity (*55*) we observe the filament thermal conductivity $k_{th,CF} \approx 1$ $Wm^{-1}K^{-1}$ [for amorphous HfO_2 (*47*)] to be relatively insensitive to the filament electrical resistivity (see Section S7). This, combined with the radii estimated from the previous approach, model the $R_{CF,t}$ (or filament thermal resistance) from fig. S6(b). Besides the value of $R_{CF,t}$, the value of ΔT_S is dependent on the two thermal interface resistances $R_{int,top}$ and $R_{int,bot}$.

By matching the simulated profile to the measured profile as shown in fig. S6(c), we extract $d_{CF} \sim 4$ nm at 127 μW for TiN TE device. We can extract a range of $G_{CF\text{-}TE}$ for a given diameter d_{CF} on fitting to measured SThM data. The widest range of TBC for most known material interfaces is from 3 to 3000 $MWm^{-2}K^{-1}$ at room temperature (*40*). Over this range of TBC values, we simulate a range of d_{CF} to compare to our measured ΔT_{top} profile. We find that the range of d_{CF} is small over this entire range of TBC values for SLG TE devices. We observe a large temperature drop at the interface of the filament with the TE, as shown in fig. S6(d). We keep our maximum ΔT_{CF} limited by the maximum material limit (i.e. HfO_2 melting).

Detailed simulation parameters for Fig. 4(B-C) are shown in Table S2. The d_{CF} (CF constriction diameter) is gradually increased from 7.4 to 13.4 nm in simulations, with increase in power from 5 to 228 μW, in order to match measured temperature profiles. This indicates increased electrical conduction in the sub-stoichiometric oxide surrounding the filament at higher electrical bias and temperature.

Table S2. Simulation parameters for Fig. 4(B-C), with SLG TE. Parameters at 5 μW, 73.8 μW, and 228 μW correspond to panels (i), (ii), (iii) in Fig. 4B. The thermal conductances G are at 296 K.

$P = I \times V$	5 μW = 20 μA × 0.25 V	29.2 μW = 40 μA × 0.73 V	73.8 μW = 60 μA × 1.23 V	150.4 μW = 80 μA × 1.88 V	228 μW = 100 μA × 2.28 V
d_{CF} (nm)	7.4	9	11	13	13.4
d_{top} (nm)	13	13	13.4	17	18
d_{bot} (nm)	16	16	16	17	18
$G_{CF\text{-}TE}$ ($MWm^{-2}K^{-1}$)	75	75	85	100	300
$G_{CF\text{-}BE}$ ($MWm^{-2}K^{-1}$)	100	100	150	150	200
$\rho_{C,CF\text{-}TE}$ (Ω-cm^2)	10^{-9}	10^{-9}	3.3×10^{-9}	1.1×10^{-8}	1.1×10^{-8}

S7. Filament Thermal Conductivity Model and Filament Shape

To model the RRAM filament electro-thermally, we assume diffusive thermal and electronic transport in the filament (*56*). This assumption is justified in the case where the filament is not single crystalline (*56, 57*). We had previously modeled the phononic (*55*) component of thermal conductivity for HfO_2. Figure S7(a) shows our modeled phononic (lattice) thermal conductivity for HfO_2 in comparison with previously measured and modeled values by Panzer *et al*. (*47*). In order to incorporate the electronic components of the thermal conductivity in a high defect density (n_D) RRAM oxide, we consider the defect- and temperature-dependent electronic conductivity model by Larentis *et al*. (*58*) reproduced here as eq. 10 below.

$$\sigma = \sigma_0 e^{\frac{-E_{AC}}{k_B T}} \tag{10}$$

Here, σ_0 is a pre-exponential factor, E_{AC} is the activation energy for conduction, k_B is the Boltzmann constant and T is the temperature. Both σ_0 and E_{AC} have an assumed piece-wise linear dependence on n_D (*58*). From this model of electronic conductivity, we calculate the equivalent electronic component of thermal conductivity ($k_{th\text{-}e}$) using Wiedemann-Franz law shown in eq. 11 below.

$$k_{th-e} = L_0 T \sigma \tag{11}$$

Here the Lorenz number $L_0 = 2.44 \times 10^{-8}$ WΩ/K^2, while other symbols are as defined above. Note that this value of L_0 is an assumption for the degenerate limit, in general, this value can be in the range 2 to 4 × 10^{-8} WΩ/K^2 for 1D-like metallic nanostructures (*59*). With this current model, we calculate the total defect-dependent thermal conductivity (k_{th}) for our RRAM oxide as:

$$k_{th} = k_{ph} + k_{th-e} \tag{12}$$

Calculated k_{th} is shown in fig. S7(b), with the calculated phononic and electronic thermal conductivities also displayed, for a mean free path (D) ~ 0.5 nm, as used previously by Niraula *et al*. (*56*). Our detailed model enables us to account for the temperature dependence of both the total electronic and thermal conductivities of our RRAM filament at a particular n_D. These complete thermal and electronic conductivity models are used in our electro-thermal simulations for a single RRAM device.

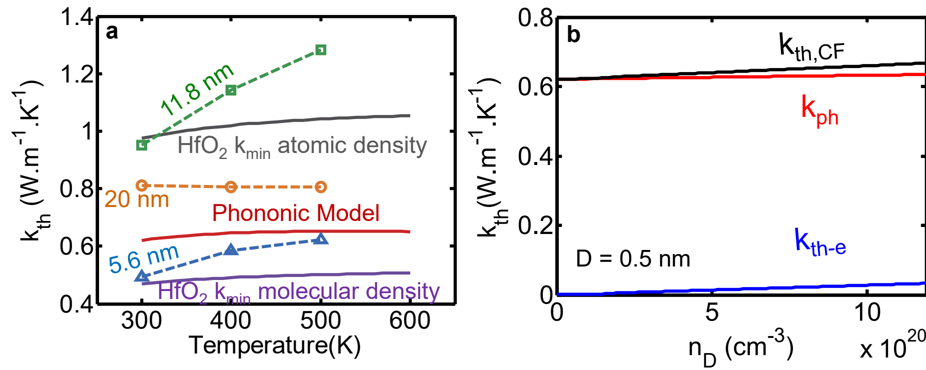

Fig. S7. Phononic and total thermal conductivity. (a) Thermal conductivity from phononic contribution in HfO_2 in this work (red) compared with other measured and simulated values by Panzer *et al*. (*47*) **(b)** Total thermal conductivity (black) is dominated by phononic component (red) at defect densities considered with an assumed electronic mean free path *D* = 0.5 nm [see Niraula *et al*. (*56*)].

S8. Contribution from TBC and Filament Thermal Conductivity

From electro-thermal simulations, we determine the effect of the thermal boundary conductance (TBC) vs. the filament thermal conductivity. For instance, as the electrical power applied to the CF increases, the temperature rise in the CF is higher, as driven by Joule heating. Thus, a higher heat flux flowing through the top and bottom electrodes leads to ΔT_S observed with *in operando* measurements. The temperature dependent $G_{CF\text{-}TE}$ trend seen for the fitted values agrees with a linear ($\propto T$) temperature dependence for TiN TE. Sensitivity to $G_{CF\text{-}TE}$ is ~20% of the extracted value, assuming all other parameters remain similar. Changing n_D in the filament can change the filament thermal conductivity over a range ~1 to 20 $Wm^{-1}K^{-1}$, but this does not substantially change our extracted $G_{CF\text{-}TE}$ beyond the 20% error. However, changing the diameter of the CF constriction [d_{CF}, see fig. S6(b)] drastically changes the ΔT_S value, and a corresponding change in the assumed $G_{CF\text{-}TE}$ is necessary to match the measured ΔT_S. Since the widest measured range of $G_{CF\text{-}TE}$ is 3 to 3000 $MWm^{-2}K^{-1}$ across all known material interfaces (*40*), this limits the d_{CF} values that can simulate the measured ΔT_S. A second set of constraints comes from the electrical resistance measured for the device. A wide filament would be electrically more conducting, hence ruled out.

With these two constraints (matching to measured ΔT_S and electrical resistance R_{LRS}), the extracted $G_{CF\text{-}TE}$ values are shown in fig. S8(a). It is interesting to note that the fitted $G_{CF\text{-}TE}$ for the SLG-HfO_x filament interface is higher than the TiN-HfO_x filament interface. However, we note that this extraction is one particular solution that fits our measurement. Detailed materials-level understanding of the filament and controlled forming measurements are necessary to precisely control filament diameters and extract a more accurate value of the $G_{CF\text{-}TE}$. The uncertainty in the actual value of the $G_{CF\text{-}TE}$ leads to a subsequent uncertainty in the filament temperature. Our simulated filament temperatures for the maximum power (~ 250 μW) case in Fig. 4A have resulted in a ΔT_{CF} from 800 to over 1200 K above ambient, with worst case thermal conductivity and TBC assumptions.

The shape of the filament can determine the contribution of filament thermal conductivity and TBC to total thermal resistance and hence, heat confinement. However, fig. S8(b) also reveals that for most reasonable

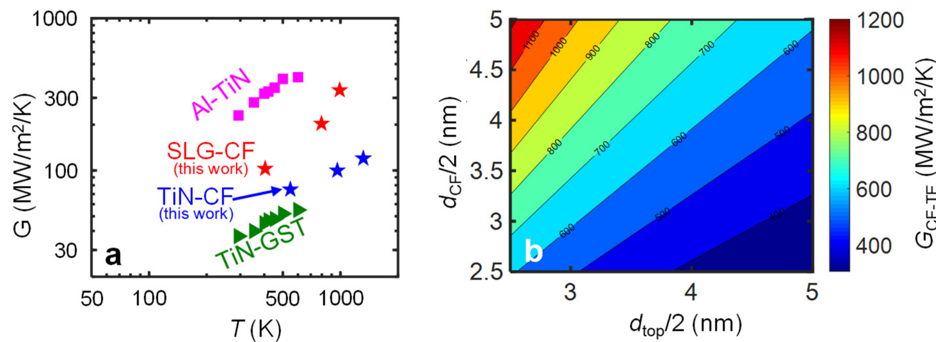

Fig. S8. Insights into TBC and filament temperature rise (ΔT_{CF}). (a) Thermal boundary conductance (TBC) for CF and TE ($G_{CF\text{-}TE}$) values for TiN-CF and SLG-CF interfaces from matching COMSOL simulations to SThM measurements. Also plotted are other TiN and memory interface TBC values from literature (*35*),(*49*). **(b)** Color plot showing $G_{CF\text{-}TE}$ values dependent on filament top radius ($d_{top}/2$) and constriction radius ($d_{CF}/2$) such that the top interface thermal resistance equals half the filament thermal resistance [$R_{int,top} = R_{CF,t}/2$ in fig. S6(b)]. Any value of TBC less than these values at the particular filament geometry would cause the interface to dominate total thermal resistance.

filament dimensions shown, the filament-TE TBC must be almost unphysically large (>400 $MWm^{-2}K^{-1}$ at room temperature) such that this interface thermal resistance ($1/G_{CF\text{-}TE}$) equals only half the total filament thermal resistance. In other words, this indicates that the filament-TE interface dominates the thermal resistance for heat spreading from a filament and demands further study.

S9. RRAM Model Sensitivity to Input Parameters

Figure S9 shows the sensitivity of 3 simulated outputs ($\Delta T_{S,max}$, V, and $\Delta T_{CF,max}$ from COMSOL simulations in Section S6) to electronic and thermal properties of a single filament. Other input parameters are assumed as specified in Table S1. Simulation over this entire broad range of parameters resulted in a consistent hot

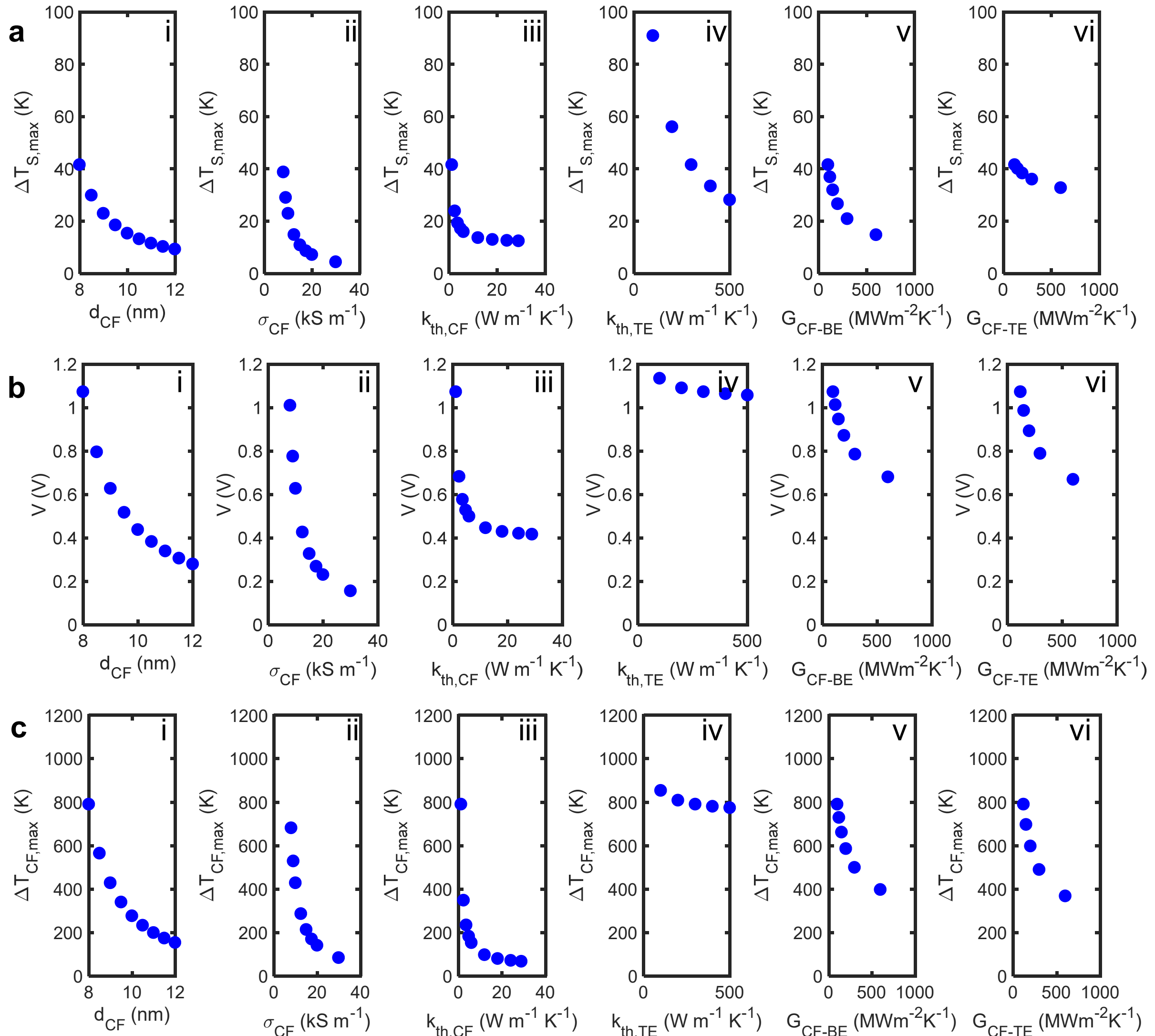

Fig. S9. Sensitivity analysis for electro-thermal parameters. Dependence of simulated **(a)** maximum surface temperature measurable by SThM: $\Delta T_{S,max}$; **(b)** voltage along the axis of the filament and **(c)** maximum filament temperature: $\Delta T_{CF,max}$ on (i) filament constriction diameter d_{CF}; (ii) filament electronic conductivity σ_{CF} (in units of 10^3 S m^{-1} or kS m^{-1}); (iii) filament thermal conductivity; (iv) top electrode thermal conductivity; (v) $G_{CF\text{-}BE}$ and (vi) $G_{CF\text{-}TE}$ – all parameters within their assumed ranges, all at 300 K and with a constant current = 40 μA through the device.

spot width ~100 to 120 nm. We also concluded that $\Delta T_{CF,max}$ and ΔT_S were very weakly dependent on the TE electronic conductivity and on the hourglass filament top diameter (negligible compared to parameters shown in fig. S9). In addition, reducing the TE thermal conductivity ($k_{th,TE}$), leads to an increase of $\Delta T_{S,max}$ with a small effect on $\Delta T_{CF,max}$ [figs. S9(a) vs. (c) (iv)]. This points to the ineffectiveness of a low $k_{th,TE}$ in determining heat confinement *within* an RRAM filament. This conclusion also suggests that thermal enhancement (i.e. confinement) in RRAM observed elsewhere (*10, 60*), is not determined by $k_{th,TE}$.

From fig. S9, we observe that simulated ΔT_S, $\Delta T_{CF,max}$ and V are very sensitive to assumed filament properties d_{CF}, σ_{CF} and $k_{th,CF}$. The values of the TBC at *both* top ($G_{CF\text{-}TE}$) and bottom ($G_{CF\text{-}BE}$) interfaces have a substantial effect on maximum filament temperature and also on the measured voltage. This is one of the first demonstrations of TBC at the nanoscale affecting overall RRAM device electrical properties, as a result of coupled electro-thermal behavior of the device. This observation suggests that the CF-electrode thermal interfaces are more likely to determine thermal confinement in RRAM devices, towards analog switching memory behavior.

S10. RRAM Device Measured Power

For all figures, the RRAM device power is estimated as: $P = IV - I^2R_{series}$. The R_{series} and the measured resistance R for devices with different TEs are shown in the Table S3 below. These values are determined from separate transfer length measurements (TLM) on these TE materials in separate test structures.

Table S3. Assumed series resistance values and measured device resistance for different TEs

TE	R_{series} (Ω)	R (Ω)
TiN/Pt	2×10^1	6.55×10^2
10 nm TiN	1×10^5	5.86×10^5
SLG	5×10^3	1.11×10^4
2-LG	5×10^3	1.40×10^4

REFERENCES AND NOTES

1. M. M. S. Aly, M. Gao, G. Hills, C. S. Lee, G. Pitner, M. M. Shulaker, T. F. Wu, M. Asheghi, J. Bokor, F. Franchetti, K. E. Goodson, C. Kozyrakis, I. Markov, K. Olukotun, L. Pileggi, E. Pop, J. Rabaey, C. Ré, H.-S. P. Wong, S. Mitra, Energy-efficient abundant-data computing: The N3XT 1,000x. *Computer* **48**, 24–33 (2015).

2. H. S. P. Wong, S. Salahuddin, Memory leads the way to better computing. *Nat. Nanotechnol.* **10**, 191–194 (2015).

3. H. S. P. Wong, H. Y. Lee, S. Yu, Y. S. Chen, Y. Wu, P. S. Chen, B. Lee, F. T. Chen, M. J. Tsai, Metal-oxide RRAM. *Proc. IEEE* **100**, 1951–1970 (2012).

4. S. Yu, Resistive random access memory (RRAM). *Synth. Lect. Hum. Lang. Technol.* **2**, 1–79 (2016).

5. M. Yu, Y. Cai, Z. Wang, Y. Fang, Y. Liu, Z. Yu, Y. Pan, Z. Zhang, J. Tan, X. Yang, M. Li, R. Huang, Novel vertical 3D structure of TaO_x-based RRAM with self-localized switching region by sidewall electrode oxidation. *Sci. Rep.* **6**, 21020 (2016).

6. U. Celano, L. Goux, R. Degraeve, A. Fantini, O. Richard, H. Bender, M. Jurczak, W. Vandervorst, Imaging the three-dimensional conductive channel in filamentary-based oxide resistive switching memory. *Nano Lett.* **15**, 7970–7975 (2015).

7. D.-H. Kwon, K. M. Kim, J. H. Jang, J. M. Jeon, M. H. Lee, G. H. Kim, X.-S. Li, G.-S. Park, B. Lee, S. Han, M. Kim, C. S. Hwang, Atomic structure of conducting nanofilaments in TiO_2 resistive switching memory. *Nat. Nanotechnol.* **5**, 148–153 (2010).

8. C. Pan, Y. Ji, N. Xiao, F. Hui, K. Tang, Y. Guo, X. Xie, F. M. Puglisi, L. Larcher, E. Miranda, L. Jiang, Y. Shi, I. Valov, P. C. McIntyre, R. Waser, M. Lanza, Coexistence of grain-boundaries-assisted bipolar and threshold resistive switching in multilayer hexagonal boron nitride. *Adv. Funct. Mater.* **27**, 1604811 (2017).

9. S. Choi, S. H. Tan, Z. Li, Y. Kim, C. Choi, P.-Y. Chen, H. Yeon, S. Yu, J. Kim, SiGe epitaxial memory for neuromorphic computing with reproducible high performance based on engineered dislocations. *Nat. Mater.* **17**, 335–340 (2018).

10. W. Wu, H. Wu, B. Gao, N. Deng, S. Yu, H. Qian, Improving analog switching in HfO_x-based resistive memory with a thermal enhanced layer. *IEEE Electron Dev. Lett.* **38**, 1019–1022 (2017).

11. Y. Shi, X. Liang, B. Yuan, V. Chen, H. Li, F. Hui, Z. Yu, F. Yuan, E. Pop, H.-S. P. Wong, M. Lanza, Electronic synapses made of layered two-dimensional materials. *Nat. Electron.* **1**, 458–465 (2018).

12. H. Chen, S. Yu, B. Gao, P. Huang, J. Kang, H. S. P. Wong, HfO_x based vertical resistive random access memory for cost-effective 3D cross-point architecture without cell selector, in *IEEE International Electron Devices Meeting (IEDM)* (2012), pp. 20.27.21–20.27.24.

13. I. M. Datye, M. M. Rojo, E. Yalon, S. Deshmukh, M. J. Mleczko, E. Pop, Localized heating and switching in $MoTe_2$-based resistive memory devices. *Nano Lett.* **20**, 1461–1467 (2020).

14. D. Ielmini, Resistive switching memories based on metal oxides: Mechanisms, reliability and scaling. *Semicond. Sci. Technol.* **31**, 063002 (2016).

15. A. Padovani, L. Larcher, O. Pirrotta, L. Vandelli, G. Bersuker, Microscopic modeling of HfO_x RRAM operations: From forming to switching. *IEEE Trans. Electron Dev.* **62**, 1998–2006 (2015).

16. S. Menzel, M. Waters, A. Marchewka, U. Böttger, R. Dittmann, R. Waser, Origin of the ultra-nonlinear switching kinetics in oxide-based resistive switches. *Adv. Funct. Mater.* **21**, 4487–4492 (2011).

17. J. Kwon, A. A. Sharma, C.-Y. Chen, A. Fantini, M. Jurczak, A. A. Herzing, J. A. Bain, Y. N. Picard, M. Skowronski, Transient thermometry and high-resolution transmission electron microscopy analysis of filamentary resistive switches. *ACS Appl. Mater. Interfaces* **8**, 20176–20184 (2016).

18. Z. Jiang, Z. Wang, X. Zheng, S. Fong, S. Qin, H. Chen, C. Ahn, J. Cao, Y. Nishi, H. S. P. Wong, Microsecond transient thermal behavior of HfOx-based resistive random access memory using a

micro thermal stage (MTS), in *IEEE International Electron Devices Meeting (IEDM)*. (2016), pp. 21.23.21–21.23.24.

19. A. A. Sharma, M. Noman, M. Skowronski, J. A. Bain, High-speed in-situ pulsed thermometry in oxide RRAMs, in *Proceedings of Technical Program - 2014 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)*. (2014), pp. 1–2.

20. S. Kumar, Z. Wang, N. Davila, N. Kumari, K. J. Norris, X. Huang, J. P. Strachan, D. Vine, A. L. D. Kilcoyne, Y. Nishi, R. S. Williams, Physical origins of current and temperature controlled negative differential resistances in NbO_2. *Nat. Commun.* **8**, 658 (2017).

21. E. Yalon, S. Deshmukh, M. M. Rojo, F. Lian, C. M. Neumann, F. Xiong, E. Pop, Spatially resolved thermometry of resistive memory devices. *Sci. Rep.* **7**, 15360 (2017).

22. J. M. Goodwill, G. Ramer, D. Li, B. D. Hoskins, G. Pavlidis, J. J. McClelland, A. Centrone, J. A. Bain, M. Skowronski, Spontaneous current constriction in threshold switching devices. *Nat. Commun.* **10**, 1628 (2019).

23. E. Puyoo, D. Albertini, Conductive filament localization within crossbar resistive memories by dcanning joule expansion microscopy. *IEEE Electron Dev. Lett.* **41**, 848–851 (2020).

24. S. Deshmukh, M. M. Rojo, E. Yalon, S. Vaziri, E. Pop, Probing Self-Heating in RRAM Devices by Sub-100 nm Spatially Resolved Thermometry, in *76th Device Research Conference (DRC)*. (2018), pp. 1–2.

25. Y. Zhang, W. Zhu, F. Hui, M. Lanza, T. Borca-Tasciuc, M. M. Rojo, A review on principles and applications of scanning thermal microscopy (SThM). *Adv. Funct. Mater.* **30**, 1900892 (2019).

26. F. Menges, P. Mensch, H. Schmid, H. Riel, A. Stemmer, B. Gotsmann, Temperature mapping of operating nanoscale devices by scanning probe thermometry. *Nat. Commun.* **7**, 10874 (2016).

27. A. A. Wilson, M. M. Rojo, B. Abad, J. A. Perez, J. Maiz, J. Schomacker, M. Martín-Gonzalez, D.-A. Borca-Tasciuc, T. Borca-Tasciuc, Thermal conductivity measurements of high and low thermal

conductivity films using a scanning hot probe method in the 3ω mode and novel calibration strategies. *Nanoscale* **7**, 15404–15412 (2015).

28. M. M. Rojo, S. Grauby, J. M. Rampnoux, O. Caballero-Calero, M. Martin-Gonzalez, S. Dilhaire, Fabrication of Bi_2Te_3 nanowire arrays and thermal conductivity measurement by 3ω-scanning thermal microscopy. *J. Appl. Phys.* **113**, 054308 (2013).

29. L. Shi, S. Plyasunov, A. Bachtold, P. L. McEuen, A. Majumdar, Scanning thermal microscopy of carbon nanotubes using batch-fabricated probes. *Appl. Phys. Lett.* **77**, 4295–4297 (2000).

30. K. Kim, J. Chung, G. Hwang, O. Kwon, J. S. Lee, Quantitative measurement with scanning thermal microscope by preventing the distortion due to the heat transfer through the air. *ACS Nano* **5**, 8700–8709 (2011).

31. E. Puyoo, S. Grauby, J.-M. Rampnoux, E. Rouvière, S. Dilhaire, Thermal exchange radius measurement: Application to nanowire thermal imaging. *Rev. Sci. Instrum.* **81**, 073701 (2010).

32. N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications* (MIT press, Cambridge, MA, 1964), vol. 8.

33. E. Yalon, C. J. McClellan, K. K. H. Smithe, M. M. Rojo, R. L. Xu, S. V. Suryavanshi, A. J. Gabourie, C. M. Neumann, F. Xiong, A. B. Farimani, E. Pop, Energy dissipation in monolayer MoS_2 electronics. *Nano Lett.* **17**, 3429–3433 (2017).

34. E. Yalon, A. A. Sharma, M. Skowronski, J. A. Bain, D. Ritter, I. V. Karpov, Thermometry of filamentary RRAM devices. *IEEE Trans. Electron Devices* **62**, 2972–2977 (2015).

35. J. P. Reifenberg, K. Chang, M. A. Panzer, S. Kim, J. A. Rowlette, M. Asheghi, H. P. Wong, K. E. Goodson, Thermal boundary resistance measurements for phase-change memory devices. *IEEE Electron Dev. Lett.* **31**, 56–58 (2010).

36. A. Marchewka, B. Roesgen, K. Skaja, H. Du, C.-L. Jia, J. Mayer, V. Rana, R. Waser, S. Menzel, Nanoionic resistive switching memories: On the physical nature of the dynamic reset process. *Adv. Electronic Mater.* **2**, 1500233 (2016).

37. R. Degraeve, A. Fantini, N. Raghavan, L. Goux, S. Clima, Y. Y. Chen, A. Belmonte, S. Cosemans, B. Govoreanu, D. J. Wouters, P. Roussel, G. S. Kar, G. Groeseneken, M. Jurczak, Hourglass concept for RRAM: A dynamic and statistical device model, in *Proceedings of the 21th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*. (2014), pp. 245–249.

38. A. Kalantarian, G. Bersuker, D. C. Gilmer, D. Veksler, B. Butcher, A. Padovani, O. Pirrotta, L. Larcher, R. Geer, Y. Nishi, P. Kirsch, Controlling uniformity of RRAM characteristics through the forming process, in *2012 IEEE International Reliability Physics Symposium (IRPS)*. (2012), pp. 6C.4.1-6C.4.5.

39. A. Banerjee, H. Satoh, A. Tiwari, C. Apriono, E. T. Rahardjo, N. Hiromoto, H. Inokawa, Width dependence of platinum and titanium thermistor characteristics for application in room-temperature antenna-coupled terahertz microbolometer. *Jpn. J. Appl. Phys.* **56**, 04CC07 (2017).

40. E. Pop, Energy dissipation and transport in nanoscale devices. *Nano Res.* **3**, 147–169 (2010).

41. M. K. Samani, X. Z. Ding, N. Khosravian, B. Amin-Ahmadi, Y. Yi, G. Chen, E. C. Neyts, A. Bogaerts, B. K. Tay, Thermal conductivity of titanium nitride/titanium aluminum nitride multilayer coatings deposited by lateral rotating cathode arc. *Thin Solid Films* **578**, 133–138 (2015).

42. W. Jang, Z. Chen, W. Bao, C. N. Lau, C. Dames, Thickness-dependent thermal conductivity of encased graphene and ultrathin graphite. *Nano Lett.* **10**, 3909–3913 (2010).

43. R. E. Taylor, J. Morreale, Thermal conductivity of titanium carbide, zirconium carbide, and titanium nitride at high temperatures. *J. Am. Ceram. Soc.* **47**, 69–73 (1964).

44. S. Vaziri, E. Yalon, M. Muñoz Rojo, S. V. Suryavanshi, H. Zhang, C. J. McClellan, C. S. Bailey, K. K. H. Smithe, A. J. Gabourie, V. Chen, S. Deshmukh, L. Bendersky, A. V. Davydov, E. Pop, Ultrahigh thermal isolation across heterogeneously layered two-dimensional materials. *Sci. Adv.* **5**, eaax1325 (2019).

45. P. Sun, N. Lu, L. Li, Y. Li, H. Wang, H. Lv, Q. Liu, S. Long, S. Liu, M. Liu, Thermal crosstalk in 3-dimensional RRAM crossbar array. *Sci. Rep.* **5**, 13504 (2015).

46. V. E. Dorgan, M.-H. Bae, E. Pop, Mobility and saturation velocity in graphene on SiO_2. *Appl. Phys. Lett.* **97**, 082112 (2010).

47. M. A. Panzer, M. Shandalov, J. A. Rowlette, Y. Oshima, Y. W. Chen, P. C. McIntyre, K. E. Goodson, Thermal properties of ultrathin hafnium oxide gate dielectric films. *IEEE Electron Dev. Lett.* **30**, 1269–1271 (2009).

48. G. Chen, P. Hui, Thermal conductivities of evaporated gold films on silicon and glass. *Appl. Phys. Lett.* **74**, 2942–2944 (1999).

49. R. M. Costescu, M. A. Wall, D. G. Cahill, Thermal conductance of epitaxial interfaces. *Phys. Rev. B* **67**, 054302 (2003).

50. J. C. Duda, C. Y. P. Yang, B. M. Foley, R. Cheaito, D. L. Medlin, R. E. Jones, P. E. Hopkins, Influence of interfacial properties on thermal transport at gold:silicon contacts. *Appl. Phys. Lett.* **102**, 081902 (2013).

51. D. H. Hurley, M. Khafizov, S. L. Shinde, Measurement of the Kapitza resistance across a bicrystal interface. *J. Appl. Phys.* **109**, 083504 (2011).

52. D. Villaroman, X. Wang, W. Dai, L. Gan, R. Wu, Z. Luo, B. Huang, Interfacial thermal resistance across graphene/Al_2O_3 and graphene/metal interfaces and post-annealing effects. *Carbon* **123**, 18–25 (2017).

53. Z.-Y. Ong, M. V. Fischetti, A. Y. Serov, E. Pop, Signatures of dynamic screening in interfacial thermal transport of graphene. *Phys. Rev. B* **87**, 195404 (2013).

54. S. Deshmukh, E. Yalon, F. Lian, K. E. Schauble, F. Xiong, I. V. Karpov, E. Pop, Temperature-dependent contact resistance to nonvolatile memory materials. *IEEE Trans. Electron Dev.* **66**, 3816–3821 (2019).

55. S. Deshmukh, R. Islam, C. Chen, E. Yalon, K. C. Saraswat, E. Pop, Thermal modeling of metal oxides for highly scaled nanoscale RRAM, in *2015 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. (2015), pp. 281–284.

56. D. Niraula, V. G. Karpov, Heat transfer in filamentary RRAM devices. *IEEE Trans. Electron Dev.* **64**, 4106–4113 (2017).

57. K. T. Regner, J. A. Malen, Nondiffusive thermal transport increases temperature rise in RRAM filaments. *IEEE Electron Dev. Lett.* **37**, 572–575 (2016).

58. S. Larentis, F. Nardi, S. Balatti, D. C. Gilmer, D. Ielmini, Resistive switching by voltage-driven ion migration in bipolar RRAM—Part II: Modeling. *IEEE Trans. Electron Dev.* **59**, 2468–2475 (2012).

59. N. Stojanovic, D. H. S. Maithripala, J. M. Berg, M. Holtz, Thermal conductivity in metallic nanostructures at high temperature: Electrons, phonons, and the Wiedemann-Franz law. *Phys. Rev. B* **82**, 075418 (2010).

60. W. Wu, H. Wu, B. Gao, P. Yao, X. Zhang, X. Peng, S. Yu, H. Qian, A Methodology to Improve Linearity of Analog RRAM for Neuromorphic Computing, in *2018 IEEE Symposium on VLSI Technology*. (2018), pp. 103–104.