Novel Topologies for Highly-Scalable Non-Volatile NEMS Relays

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

James T. Best

B.S. Engineering, Harvey Mudd College (2014)

Carnegie Mellon University

Pittsburgh, PA

December, 2019

Acknowledgements

I would like to thank my advisor, Professor Gianluca Piazza, for his continual support throughout my Ph.D. He has helped me grow both academically and personally as he guided me through the various challenges that come during a Ph.D. I would also like to thank Jim Bain, Larry Pileggi, and Maarten De Boer for being on my defense committee. Their suggestions and guidance have been invaluable in developing and presenting my work.

I would like to thank the National Science Foundation DMREF project (award# DMREF-1334572) and the Kavcic-Moura Endowment Fund for supporting my research projects on the Pulse Activated Piezo Switch and the Phase Change NEMS Relay, respectively. I am thankful for the support of the Carnegie Mellon University Claire and John Bertucci Nanotechnology Laboratory staff members, Matt Moneck, Norm Gottron, Mason Risley, Mark Weiler, James Rosvanis, and Dante Boni for their knowledge, support, and help with fabrication.

I would like to thank the past and present members the Piazza lab as well as other nanofabrication colleagues. I would especially like to thank Ayaz who has been working with me on the Phase Change NEMS Relay project since starting his Ph.D.

I would like to thank my friends who have always provided entertainment and support throughout my Ph.D. program. I would especially like to thank my life-long friend, Eugene, who has offered support and competition since I was 7 years old. I would also like to thank my girlfriend, Catherine, whose love and support have been instrumental to completing my Ph.D.

I would like to thank Rick Dill, who encouraged me to do a Ph.D. at the Carnegie Mellon Electrical and Computer Engineering department. His guidance was a key factor in deciding to continue my education and work towards a Ph.D.

Finally, I would like to thank my parents, John and Maggie, who have always encouraged me to continue my education and never stop pushing my boundaries. Additionally, their support for my passion of engineering, ranging from mentoring FIRST robotics to even installing machine tools in the garage, has been key to my success. I wouldn't have been able to finish my Ph.D. without their support.

Abstract

Complementary metal oxide semiconductors (CMOS) are nearing the limits for supply voltage and energy efficiency, motivating the search for alternative technologies. Emerging technologies like resistive memory (ReRAM), magnetoresistive memory (MRAM), phase change memory (PCRAM), and ferroelectric memory (FeRAM) offer scalable alternatives to CMOS, but are limited in either reliability, threshold voltage, or power consumption. Microelectromechanical system (MEMS) relays have gained interest as an energy efficient alternative to solid-state technologies due to the low leakage, abrupt switching, and high on/off ratios associated with mechanical contacts. A wide range of actuation methods are used in MEMS relays including electrostatic attraction, piezoelectric expansion, and thermal expansion. Common to these actuation methods is the use of a flexure to provide constraint or restoring force to overcome contact adhesion. These flexures prevent MEMS relays from scaling to dimensions comparable to CMOS devices.

This dissertation focuses on the development of the Phase Change NEMS Relay (PCNR), a novel non-volatile mechanical relay that eliminates the use of flexures and addresses the limited scalability of traditional MEMS relays. Phase change materials like GeTe can support crystalline and amorphous states at room temperature, which differ in material properties. Converting GeTe from the crystalline to amorphous state requires melting and quenching the GeTe. The PCNR is the first device to exploit the large 10 % volume change between the crystalline and amorphous states of GeTe in an actuator for mechanical displacement.

The PCNR is fabricated and tested with heater dimensions as small as 1 µm wide by 3 µm long and an air-gap of 20 nm. Actuator expansion is measured to be 26 nm, and switching is demonstrated with on and off times of 300 ns and 600 ns, respectively. On state resistance is measured to be 260 Ω and non-volatility is demonstrated for over 24 hours. Off state leakage is measured as low as 10^{-14} A. Scaling

analysis shows a path towards CMOS comparable device sizes (5 nm wide by 20 nm long heater) as well as lower than CMOS actuation voltage (0.52 V). Actuation energy is predicted to be as low as 1.8 pJ with a 1.8 ns actuation time.

Table of Contents

List of Tables

List of Figures

1 Introduction

Complementary metal oxide semiconductors (CMOS) have dominated all aspects of computing from logic to memory. Moore's law, which predicted aggressive scaling of transistors, lead to the complex, high speed, and efficient computers of today [1]. Scaling of transistors has become more difficult and is expected to saturate around 2028 [2]. Further scaling will need to be through 3D integration, which increases device density but does not address power efficiency. Concerns with energy efficiency and energy consumption have grown with the advent of the internet of things (IoT).

Leakage is becoming a dominant factor in energy consumption as CMOS is continued to be scaled down [3]. A tradeoff exists between actuation voltage and leakage [4]. The sub-threshold slope describes the current flow through a transistor as a function of voltage applied to the gate (Figure 1). In particular the sub-threshold slope defines the transistor response before the threshold voltage. A steep slope allows for greater on/off ratios for a given gate voltage. Switching voltage, and therefore switching energy for CMOS is limited by the sub-threshold slope [5]. An ideal energy efficient technology will have a steep sub threshold slope to maximize gain and minimize actuation voltage. Additionally, "more than Moore" devices have gained attention where additional functions may be built into a device. One particular feature is non-volatility where a device can hold a state without applied power. Non-volatility opens the use of a device in memory type applications. Lower voltage, lower power, and non-volatility are desired features of a CMOS alternative.

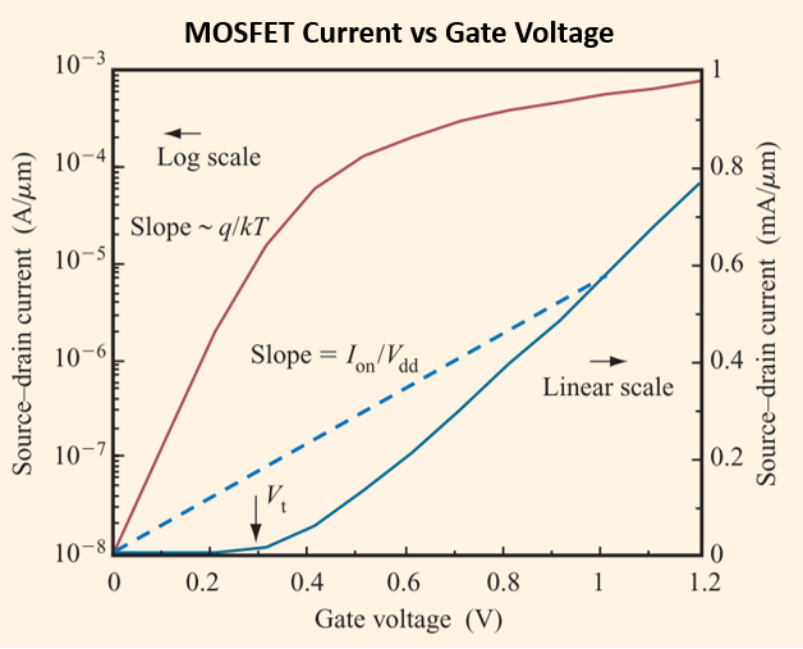

Figure 1. Current vs gate voltage for a MOSFET [5]*. Sub-threshold slope and desired on-off ratio sets the threshold voltage of a transistor.*

1.1 Emerging Technologies

Many emerging CMOS alternatives are targeted towards non-volatile memory applications. Of these new technologies, the most common are: resistive memory (ReRAM), phase change memory (PCRAM), and magnetoresistive memory (MRAM) [6] [7]. These emerging memories work by changing the resistance of a material through various actuation methods. Additionally they are all characterized as two-terminal devices, where actuation and state sensing are performed through the same set of terminals. Another emerging technology is the ferroelectric field effect transistor (FeFET) which is a four-terminal device that can be in memory or logic circuits.

ReRAM (Figure 2) works by controlling the resistance of a metal-insulator-metal structure [8][9][10]. If sufficient voltage is applied to the metal electrodes that sandwich the insulator, a conductive path is formed, thus changing the insulator to a conductor. Forming the conductive path is referred to as

the “SET” action. Further application of voltage will destroy the conductive path, thus resetting the device to a high resistance state. Returning the device to a high resistance state is referred to as the “RESET” action. ReRAM devices are small and can be formed into dense memory arrays. In addition, resistance can be controlled leading to ReRAM use in fast and efficient neuromorphic computing applications [11][12]. The main disadvantage of ReRAM is reliability and consistency of switching.

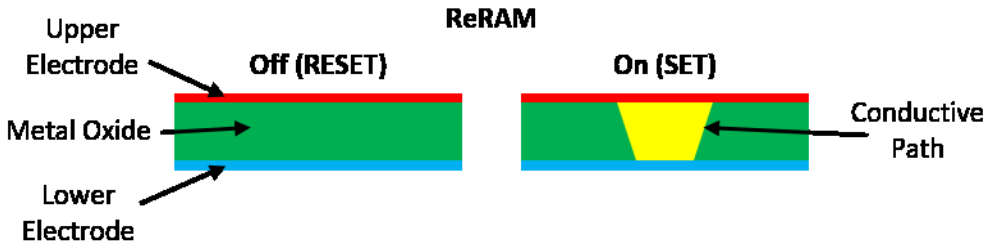

Figure 2. Schematic of a ReRAM device. A single device can be used to store either a digital value (1/0) or an analog value (relative resistance).

MRAM (Figure 3) devices are built using a magnetic tunnel junction (MTJ) [13][14][15]. The MTJ is a stack of two ferroelectric materials separated by a thin (< 2 nm) insulator. Switching the magnetic moments of the ferroelectric layers to be either aligned or misaligned will adjust the resistance through the MTJ as their alignment effects the tunneling through the thin insulating layer. Resistance is low for aligned moments, while resistance is high for misaligned moments. Switching is performed by the application of high magnetic fields. A few disadvantages of MRAM are limited scalability and high required currents associated with generating the fields. An alternate approach to MRAM is spin-transfer torque MRAM (STT-MRAM) [13]. In STT-MRAM current is flowed directly through the MTJ to switch alignment of the magnetic moments in the ferroelectric layers. The main disadvantages to STT-MRAM are fabrication difficulties associated with etching the MTJ pillars, endurance of the MTJ, and tradeoffs between error rates and switching energy.

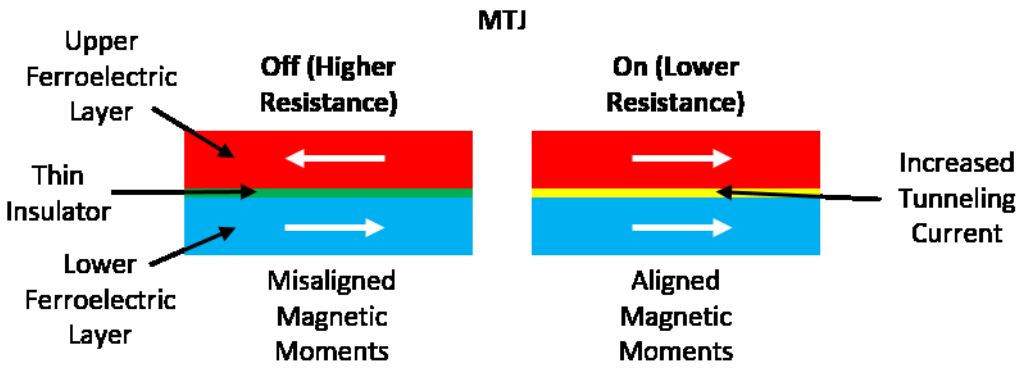

Figure 3. Schematic of an MTJ. Change in magnetic pole alignment of the two ferroelectric layers varies tunneling current through a thin insulator.

PCRAM (Figure 4) is built around the change in resistivity when switching between crystalline (conductive) and amorphous (resistive) states of chalcogenide glasses [16][17][18][19][20]. Switching states is a thermal process. A RESET operation will melt and quench a section of phase change material to set it in the amorphous state. A SET operation will heat the amorphous material to foster crystal growth, bringing the phase change material into the crystalline state. PCRAM has already seen success in commercial memory applications and will continue to be developed. The advantages of PCRAM are its high scalability, relatively high speed for a memory, and relatively high on-off ratio. The major disadvantages of PCRAM are the higher currents required to resistively heat and melt the phase change material, and limited ability to maintain non-volatility in elevated temperature environments.

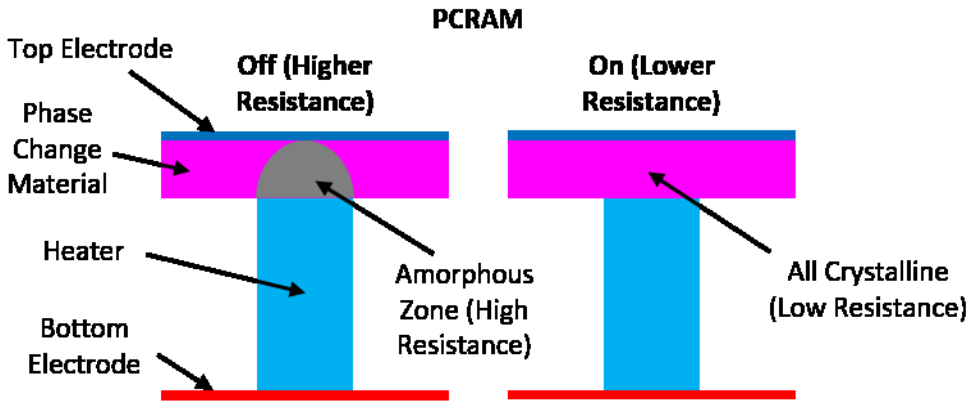

Figure 4. Schematic of a PCRAM cell. A small section of phase change material is converted between amorphous and crystalline states to control device resistivity.

The FeFET (Figure 5) is a modified field effect transistor where the gate is augmented by the addition of a ferroelectric layer like HfO_2 [21][22]. The threshold voltage of the FET is shifted by the polarization direction of the ferroelectric layer. When the ferroelectric layer is polarized upward, the threshold voltage of the FET is high. When the ferroelectric layer is polarized downward, the threshold voltage of the FET is low. Gate voltage can be set between the threshold voltages associated with the different ferroelectric polarization directions. With the proper gate voltage, the FET will turn on and off depending on the ferroelectric polarization direction. Unlike the other emerging technologies, the FeFET is a four-terminal device, which makes possible to use it as logic element. FeFETs are fast and scalable, but require relatively high voltage (around 4 V) to switch states.

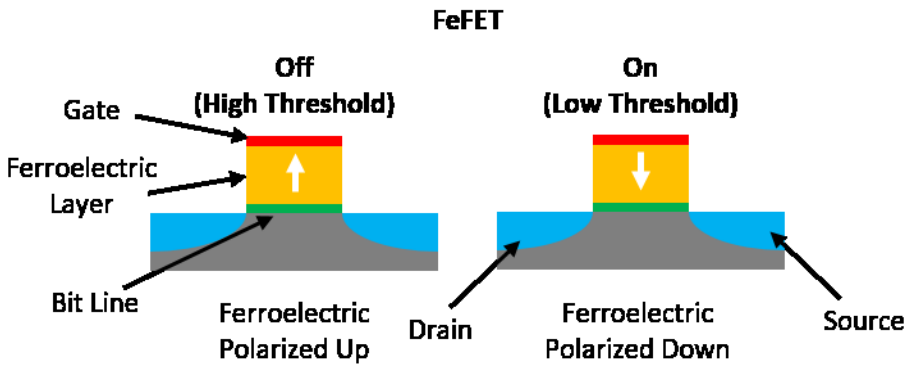

Figure 5. Schematic of a FeFET. The polarization direction of a ferroelectric layer changes the threshold voltage of a FET.

1.2 Background on MEMS Relays

MEMS relays have come out as an alternative technology to CMOS to improve energy efficiency in computing and memory [23][24][25]. MEMS relays are built around contacts separated by an "air-gap" which is opened and closed by an actuator. When the contacts are open and separated by an air gap there is near zero leakage. When the contacts are closed, current can flow only limited by the contact resistance. These relays are characterized by a large on-off ratio. In addition, switching energy can be dramatically reduced as the transition from on to off states is determined by physical connection of the contacts [26]. MEMS relays differ from other emerging technologies in that they are not two terminal devices and can be used in logic. The high energy efficiency and resistance to harsh environments has led to MEMS relays being proposed as a replacement to CMOS for power efficient computing [27]. One particular target application for MEMS relays is integration with CMOS in field programmable gate arrays (FPGAs). FPGAs are reconfigurable logic, which can benefit from the zero leakage current associated with the air gap of MEMS relays and the potentially low on state resistance. A study by Altera, an FPGA manufacturer, showed integration of MEMS relays in an FPGA could reduce total area by 43.6 %, reduce leakage power by 37 %, and reduce critical path delay by up to 28 % [28].

1.2.1 Current State of MEMS Relays

A wide range of MEMS relays have been proposed ranging in function, actuation, and fabrication. The common design aspect among MEMS relays is the use of a flexure. This flexure provides constraint for components and in most cases a restoring force. Common actuation methods include electrostatic attraction, thermal expansion, and the piezoelectric effect.

One of the most popular actuation methods for a relay uses electrostatic attraction to pull a pair of contacts together or apart. The simplest implementation of an electrostatic MEMS relay is a cantilever beam actuated by a single electrostatic pad (Figure 6) [29][30][31][32][33]. Beams are mostly commonly fabricated through common lithographic processes, but some are grown as nano-wires [34]. Electrostatic cantilevers have even been integrated to form a ring oscillator [35]. To date, the smallest by area demonstrated MEMS relays are simple electrostatic cantilever beams [36][37]. The cantilever beam devices typically rely on the flexure to provide adequate restoring force to break adhesion. Electrostatic actuation is typically characterized by high actuation voltages required to generate required forces over large gaps. The high actuation voltage limits the actuation energy reduction of this relay design compared to CMOS.

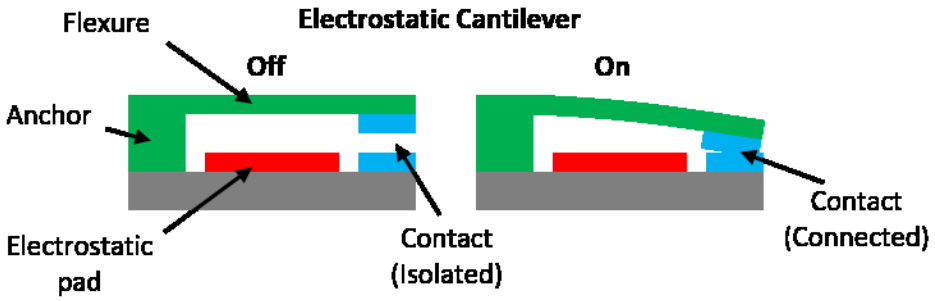

Figure 6. Schematic of a simple electrostatic cantilever MEMS relay. Electrostatic force pulls the contact pair together, allowing current to flow.

Another prominent example of an electrostatically actuation relay is the crab leg based four terminal electrostatic relay [38][39]. These relays (Figure 7) have been developed to minimize actuation voltage. A bias force is introduced by building a four terminal design. The bias force is generated by a constant voltage applied to the relay and allows for the relay to be kept close to the threshold for switching. The bias voltage does not switch, so no energy is consumed in the relay by the bias. The actuation voltage is dramatically reduced, lowering switching energy associated with charging and discharging both the actuator of the relay and the capacitance of the traces connecting to it. Other relays have also adopted the bias technique to reduce the switching energy consumption to near the theoretical limit [26]. The crab leg actuators have been used in logic circuits and a demonstration of a simple microprocessor [27]. The large actuator area of these relays limits scalability and requires significantly larger areas than other relays.

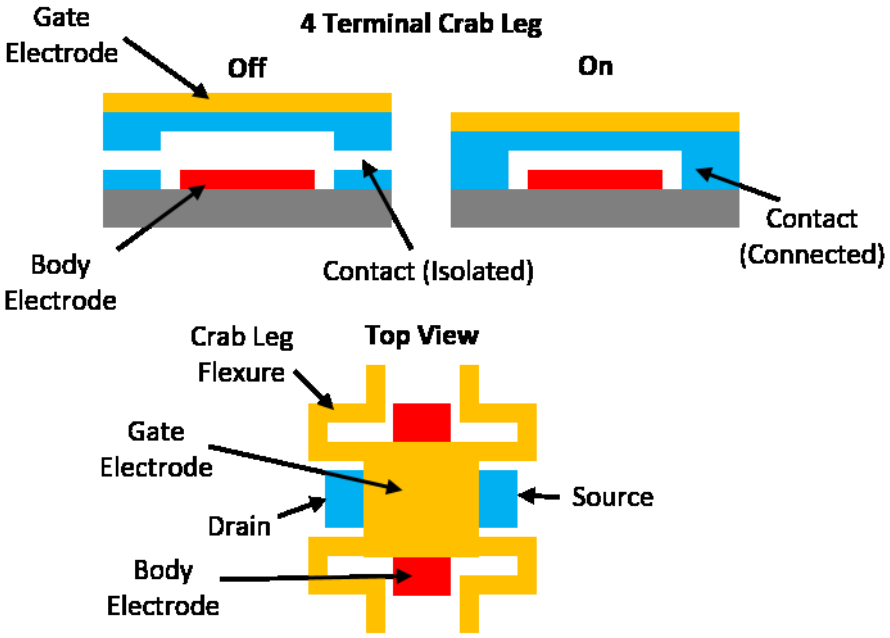

Figure 7. Schematic of the four terminal crab leg relay design. A bias can be applied to the body electrode, while actuation voltage is applied to the gate electrode. Drain and source contacts are separated from the electrodes.

Another alternative electrostatic MEMS relay design is based on the back end of the line (BEOL) CMOS process where the relay is built within the interconnects. In this relay design (Figure 8) an electrostatic cantilever is built to stand vertically like a pillar and deform laterally [40]. The pillar configuration reduces the area required by the cantilever as it is now limited by the cross-section of the beam and the lithography for forming the air gap. Additionally, the relay is non-volatile and can be used as memory. The beam is held in a particular state by surface adhesion. Although this design shows promise for scalability and CMOS compatibility, tuning the residual stress in the layers and the defining high aspect ratio nanoscale airgaps will be a challenge for device yield.

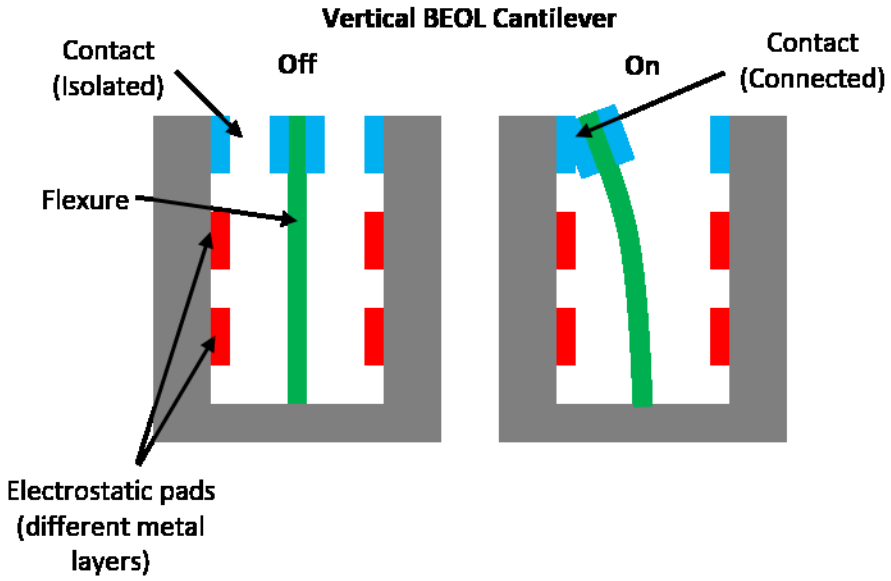

Figure 8. Schematic of vertical electrostatic cantilever relay built in a back end of the line process. The beam and contacts are made from multiple interconnect layers from a CMOS process.

Thermal actuation is less commonly used in relays as it generally requires high power to maintain a state. One example of a thermal relay (Figure 9) uses a chevron type structure to close and open an air gap [41]. Thermal expansion as an actuation method generates high forces but small displacement. The chevron design amplifies the relatively small strains generated by thermal expansion into useable

displacement. This thermal relay example is designed for characterizing relay contact materials. Since the relay is primarily used for laboratory measurements power consumption and speed are not important, only robustness and consistency.

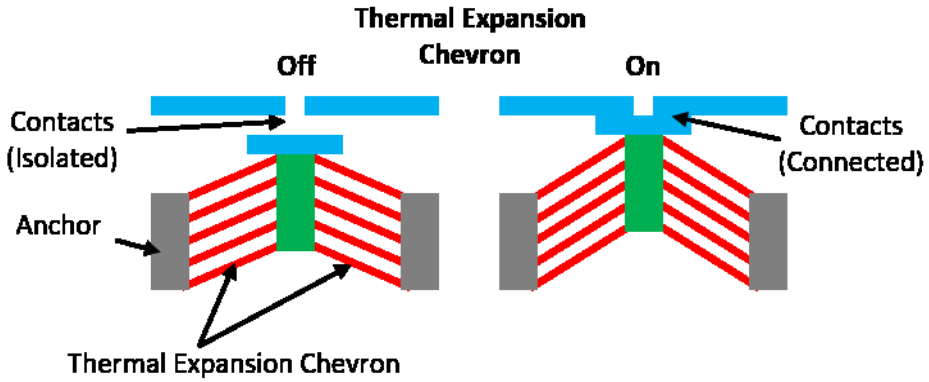

Figure 9. Schematic of a switch actuated by thermal expansion. A chevron structure is used to amplify the small displacements of the expanding beams.

Piezoelectric actuation uses piezoelectric materials like aluminum nitride (AlN) to generate strains from an applied electric field. Piezoelectric relays are typically unimorph or bimorph cantilevers (Figure 10) where the piezoelectric is off center from the neutral axis of the beam. When the piezoelectric expands or contracts, it induces a bending moment in the beam and deflects the tip. Residual stresses and the fabrication of thin piezoelectric layers are challenges in the fabrication of piezoelectric cantilever relays. 10 nm thin piezoelectric layers have been synthesized for relays [42][43]. Piezoelectric cantilevers have demonstrated sub 1V actuation with ultra-thin piezoelectric layers and stress compensated beams [44]. Alternatively, buckling piezoelectric actuators have been demonstrated as an alternative geometry that allows for low voltage operation and high-scalability [45][46]. These demonstrated piezoelectric relays are all volatile devices, and use flexures which prevent scaling to sizes comparable to CMOS components.

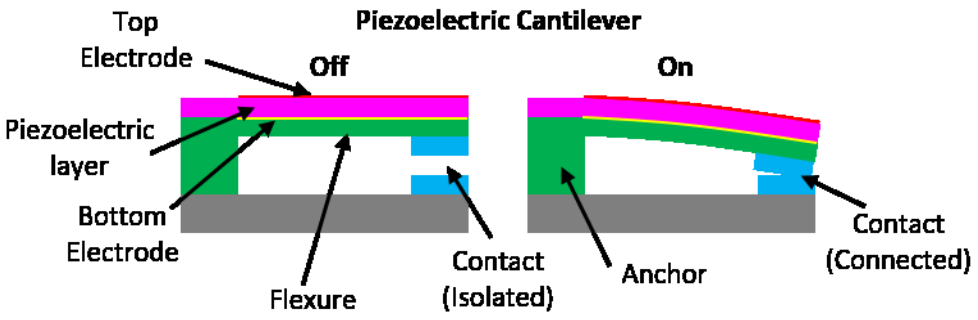

Figure 10. Schematic of a piezoelectric cantilever. Expansion of the piezoelectric layer will bend the flexure to make connect the contacts.

Non-volatile relays have gained recent interest for their potential use in memory as well as logic for computing with an intermittent power source. A wide range of geometries have been explored for non-volatile relays, but the majority of designs rely on electrostatic actuation and surface adhesion [36][40][47][48][49][50]. Demonstrated non-volatile relays use surface adhesion of the contacts to hold a specific state when actuation power is not applied. Relay adhesion, which is typically seen as a challenge to scaling, is used as a feature to hold a state. Restoring forces from the beam are small, so actuation voltage must overcome adhesion rather than rigidity of the beams.

Key limits to the adoption of MEMS relays in memory and logic are reliability and scalability [23]. Failure of the contact is one of the most common failure modes in relays [51]. A range of failure mechanisms have been investigated ranging from mechanical failure of the contact material to tribopolymer buildup on the contact [52]. Careful selection of contact materials and processing methods have shown improvements in the cycles to failure [53]. While high cycle count is required for logic, some memory based applications can tolerate lower cycles to failure.

1.2.2 Scaling of Flexure Based MEMS Relays

Although there have been attempts to aggressively scale flexure based relays, the smallest flexure based relay to date is only 200 nm by 300 nm based on the dimensions of the beam [36][37]. This excludes anchors and traces, so the full footprint is even larger. The general consensus is that these cantilever based relays cannot be scaled down to CMOS comparable dimensions with low actuation voltage and the scaling of flexure based relays is primarily limited by the air gap and contact adhesion [23].

To fully understand the limits of flexure based relay scaling an electrostatic cantilever device is evaluated for scalability. An electrostatic cantilever relay is chosen because it is the smallest NEMS relay demonstrated to date. Figure 11 shows the schematic for the evaluated design. Key dimensions are the cross-section of the beam, the length of the beam, and the air gap. The air gap sets the required deflection of the cantilever beam.

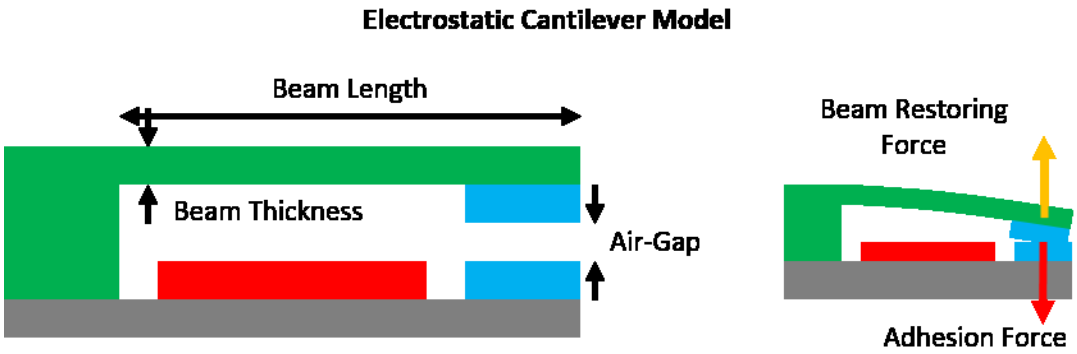

Figure 11. Schematic of an electrostatic cantilever beam relay used for scaling analysis.

Beam stresses are the primary limiting factor in miniaturizing the flexure based relays. For a cantilever beam with an end load, deflection (d) can be related to beam geometry and force by:

$$d = \frac{FL^3}{3EI} \qquad (1)$$

where F is the force at the tip of the cantilever beam, L is the length of the beam, E is the young's modulus of the beam material, and I is the second moment of area of the beam cross-section. For a rectangular cross-section, I can be modeled as:

$$I = \frac{wt^3}{12} \tag{2}$$

where w is the beam width and t is the beam thickness. Maximum stress on a bending beam is at the surface and is dependent on the applied bending moment. Maximum stress (σ_{max}) can be calculated by:

$$\sigma_{max} = \frac{Mt}{2I} \tag{3}$$

where M is the bending moment in the beam. Bending moment varies along the length of a cantilever beam with an end load. The highest bending moment is at the anchor and can be calculated by:

$$M = FL \tag{4}$$

For a viable relay, maximum stress must be below the fatigue strength of the beam material. Rearranging maximum stress to be a function of beam displacement (air-gap) gives:

$$\sigma_{max} = \frac{3Edt}{2L^2} < \sigma_{fatigue} \tag{5}$$

When isotropically scaling a beam, where all dimensions scale by the same factor (K_v), maximum stress will follow:

$$\sigma_{max} \propto \frac{dt}{L^2} \propto 1 \tag{6}$$

An aggressively scaled beam will be limited by the fatigue strength of the beam material. Assuming a minimum viable air-gap (fixed d), beam thickness and beam length must scale by:

$$t \propto L^2 \tag{7}$$

If a beam is optimized such that stress is close to the fatigue limit, it cannot be scaled after reaching the minimum air-gap. Beam thickness is limited by fabrication and sets the minimum beam length. For the purpose of this scaling analysis only electrostatic beams are considered. An alternate actuation method

uses a piezoelectric material in the beam to induce a bending moment with an applied voltage. While piezoelectric beams are generally considered a great actuation material for low power relays, they limit the scalability of the beam due to minimum layer thicknesses. The thinnest layer of piezoelectric AlN deposited for actuators to date is 10 nm, and film quality is shown to decrease with thickness [44][43] [42]. Electrostatically actuated relays only require the beam to be a metallic electrode and are therefore able to scale to thinner layers. Thinner beams allow for shorter beams, leading to smaller relays. Electrostatic actuation will cause a switch to turn on at the pull-in voltage. The pull in voltage (V_{pullin}) for a cantilever can be calculated by:

$$V_{pullin} = \sqrt{\frac{16t^3d^3E}{81\epsilon_o L^4}} \qquad (8)$$

where ϵ_o is the dielectric constant for vacuum [36]. The actuation voltage limit of electrostatic cantilevers can be estimated based on the current smallest demonstrated cantilevers to date. Actuation voltage (pull-in voltage) and device area scaling are shown in Figure 12. The smallest beam is 20 nm wide and 30 nm long and actuates at nearly 2 V. Figure 13 compares these cantilever beams to CMOS and shows that even aggressive scaling results in devices that are significantly larger than CMOS and use high actuation voltages.

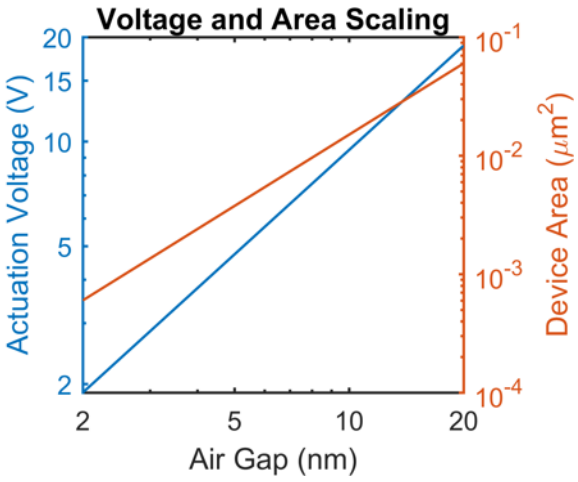

Figure 12. Isotropic scaling of voltage and area based on scaling of the smallest demonstrated MEMS relay to date.

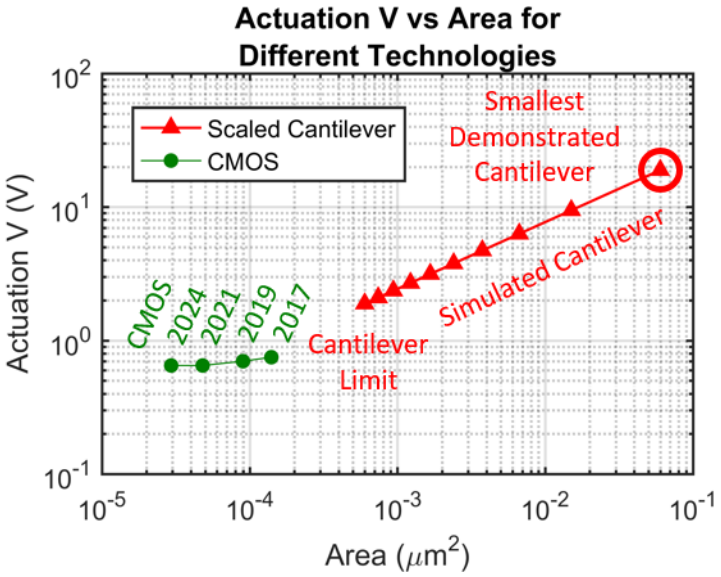

Figure 13. Comparison of actuation voltage and area between the scaled electrostatic cantilever MEMS relay and state of the art CMOS. CMOS data is pulled from [1] *and* [2].

Adhesion force between contacts is present in all relay designs and can limit the scalability of a particular design. In a non-volatile relay design, where adhesion holds a contact in place, actuation forces

must be able to overcome adhesion. In a volatile relay design, where actuation voltage must be applied to maintain a state, the restoring force of the cantilever must be able to break the adhesion of the beam. The restoring force can be calculated by:

$$F = \frac{Edwt^3}{4L^3} > F_{adhesion} \quad (9)$$

where $F_{adhesion}$ is the contact adhesion force. Contact adhesion is primarily Van der Waals forces between the contact surfaces assuming the contacts do not weld together from current flow during switching [54][55][56]. Van der Waals adhesion force (F_{VDW}) can be modeled in two parts:

$$F_{adhesion} = F_{VDW} = F_{VDW_{surface}} + F_{VDW_{asperity}} \quad (10)$$

where $F_{VDW_{surface}}$ is the parallel plate contribution from two close surfaces and $F_{VDW_{asperity}}$ is the contribution from direct contact of asperities in the contact. The parallel plate surface adhesion can be modeled as:

$$F_{VDW_{surface}} = \frac{A_m A}{D_{rms}^3 6\pi} \quad (11)$$

where A_m is the Hamaker constant for the contact material pair, A is the overlapping area, and D_{rms} is the average roughness of the two contact surfaces. Roughness keeps the two surfaces separated, limiting adhesion between the parallel surfaces. In cases of high roughness, a few asperities will hold the surfaces apart. These asperities can have a large effect on adhesion at high surface roughness or small contact area, and can be modeled as:

$$F_{VDW_{asperity}} = \frac{A_m R_A N_A}{6D_A^2} \quad (12)$$

where R_A is the radius of the asperity, N_A is the number of asperities, and D_A is the distance separating the contacting asperities. Three asperities will be in contact for a rigid plate with high roughness as it takes three points of contact to define a plane. In the case of a short cantilever one or two asperities may be in contact and provide the majority of adhesion. Figure 14 shows the restoring force and adhesion force

scaling of the smallest demonstrated cantilevers to date. Adhesion force and restoring force scale together until a gap of 5 nm, where adhesion does not scale as aggressively. This is due to asperity adhesion becoming dominant over rough surface adhesion. Depending on surface roughness of the contacts adhesion may be higher or lower than modeled. Direct measurement of the adhesion forces in a fabricated relay is possible and should be performed for an accurate force estimate [57]. If adhesion forces are lower than estimated, an aggressively scaled relay will be limited by a minimum viable air-gap.

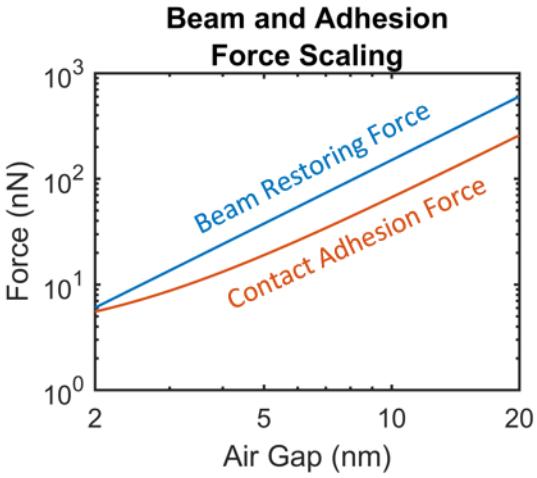

Figure 14. Isotropic scaling of beam restoring force and adhesion force based on scaling the smallest demonstrated MEMS relay to date.

In addition to physical limits of flexure based devices, fabrication becomes significantly more difficult. Stress gradients typically become harder to control with thin layers, and therefore can cause higher actuation voltages than predicted or require different aspect ratio beams to avoid generating too much stress in the flexures [58]. Overall electrostatic cantilevers are not able to meet the aggressive scaling required to compete with CMOS. An alternate design for NEMS relays is required to continue scaling for area and actuation voltage.

1.2.3 Non-Flexure Based MEMS Relays

An alternate approach to scaling NEMS relays is to eliminate the use of a flexure. This can dramatically reduce the footprint of a device as the anchor is eliminated and materials are no longer limited by the stress generated by bending. Non-flexure based relays are also generally non-volatile in nature and can be used as memory. One example of an anchorless relay is developed by Cavendish Kinetics [59][60]. In this relay (Figure 15), a "teeter-totter" is built from a single rigid beam of metal and a pivot. The relay is actuated electrostatically by tilting the beam from one side to another. Van der Waals adhesion is used to hold the beam in place, making the relay non-volatile. Non-volatility is shown to be resilient to a wide temperature range (-150 °C to 300 °C) and high acceleration forces up to 20,000 g [59]. The Cavendish Kinetics relays are fabricated in a back end of the line CMOS process – making them cheap and easy to integrate with CMOS. One drawback is a relatively high contact resistance on the order of 30 kΩ.

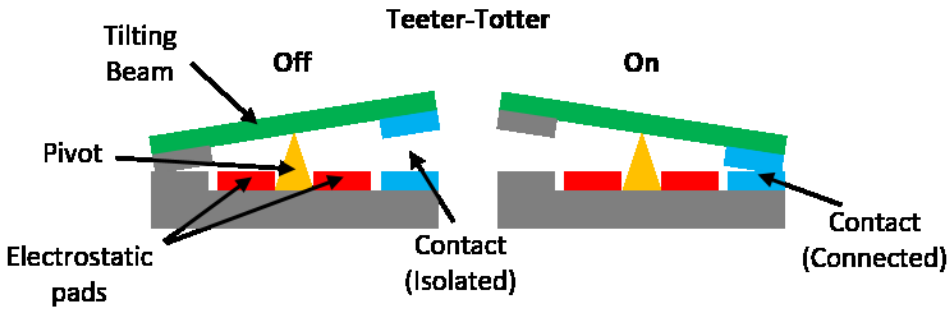

Figure 15. Schematic of the Cavendish Kinetics "teeter-totter" non-volatile relay. A rigid beam pivots to connect and disconnect a pair of contacts.

Another promising design to address the scalability of NEMS relays is the proposed Shuttle Nanoelectromechanical Nonvolatile Memory [61]. In this design (Figure 16) a free floating plate of metal, called the shuttle, is shifted within a cavity by electrostatic forces. The shuttle has two states, up and down, where the shuttle is connecting a set of contacts or sitting on the gate. Van der Waals adhesion

holds the shuttle in place, making the device non-volatile. Electrostatic bias is generated by applying a potential between the gate electrode and the drain and source electrodes. One notable problem with this design is electrostatic instability of the shuttle. The free floating shuttle transfers charge as it travels from one state to the other. Upon contact with the new electrodes, charge is transferred, and the shuttle is now attracted to the opposite state. If adequate actuation voltage is applied to overcome adhesion forces, the shuttle will oscillate between the two states. This oscillation requires the relay to use pulsed actuation. The actuation pulse must be precisely timed to avoid the shuttle bouncing back to its original state [62]. Additionally tilting of the shuttle during actuation can cause shorting between the source or drain and the gate electrode [63]. The high scalability of the electrostatic shuttle relay [64] makes it a promising device towards making CMOS competitive NEMS relays. Although fabricated, the device does not have published results demonstrating switching.

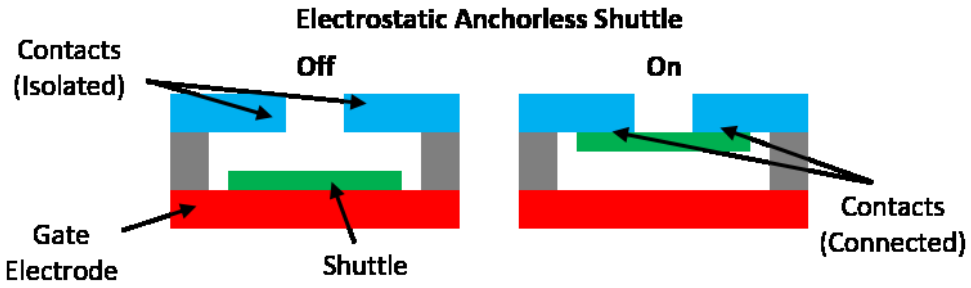

Figure 16. Schematic of the electrostatic shuttle relay. The shuttle is held in place by Van der Waals adhesion and is moved by electrostatic attraction.

1.3 Pulse Activated Piezo Shuttle Relay (PAPS)

An advanced implementation of the electrostatic shuttle relay, called the pulse activated piezo switch (PAPS), was developed to address the challenges associated with the electrostatic only actuation [65][66]. The PAPS (Figure 17) differs by adding a piezoelectric actuator to the electrostatically actuated shuttle relay. The PAPS relay consist of three main components: the actuator, shuttle, and contacts for

the drain and source. The PAPS is a 4 terminal device, different from the 3 terminal electrostatic shuttle relay. The addition of the electrostatic actuator is controlled by the body and gate terminals. The actuator expands or contracts when voltage is applied to the gate. Rapid expansion of the actuator launches and releases the shuttle through an acoustic pulse. Additionally, a body bias can be applied through electrostatic forces, similar to the electrostatic shuttle relay. These electrostatic forces can reduce the required piezoelectric actuation pulse. Analysis for the PAPS showed device performance was only dependent on layer thickness, preventing the design from suffering from the same scalability limits as the flexure based designs.

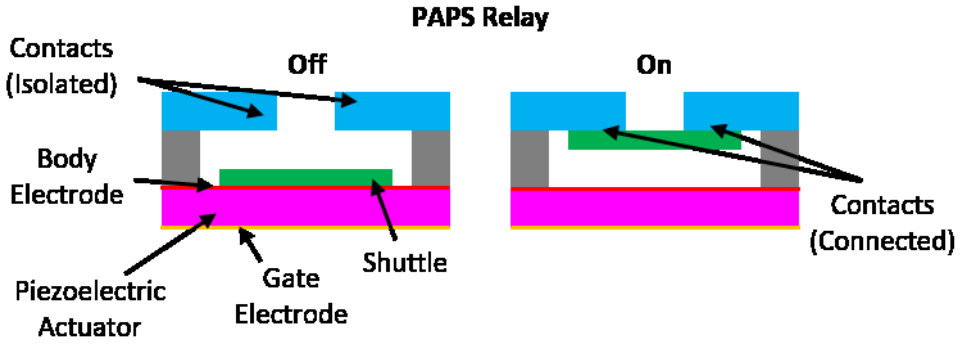

Figure 17. Schematic of the pulse activated piezo switch. A piezoelectric actuator is used to break surface adhesion and transfer a free floating shuttle between the actuator and the contacts.

The PAPS device is fabricated in a nine step process (Figure 18) and tested under both electrostatic and piezoelectric actuation. Key fabrication difficulties are associated with creating and maintaining the narrow air gap. The fabricated shuttle is 100 nm of platinum surrounded by 75 nm of sacrificial material. The narrow gaps and long etch distances make release of the shuttles difficult. Additionally the thin shuttle requires significant tuning of stress gradient to avoid the shuttle warping and shorting the contacts after release. A summary of electrostatic and piezoelectric testing is shown in Figure 19. Electrostatic only actuation is demonstrated to switch a PAPS relay between two states, with motion confirmed by SEM.

This switching is based on tilting of the shuttle, causing a short between the drain or source and the body electrodes. This instability is likely caused by varied adhesion along the actuator. The electrostatic actuation forces grow stronger as the shuttle approaches the other side, so if the shuttle begins to tilt, the tilting will be accelerated. The addition of the piezoelectric actuator can help alleviate this problem by supply additional energy to the shuttle where it touches the actuator, but piezoelectric testing still exhibits the tilting phenomenon.

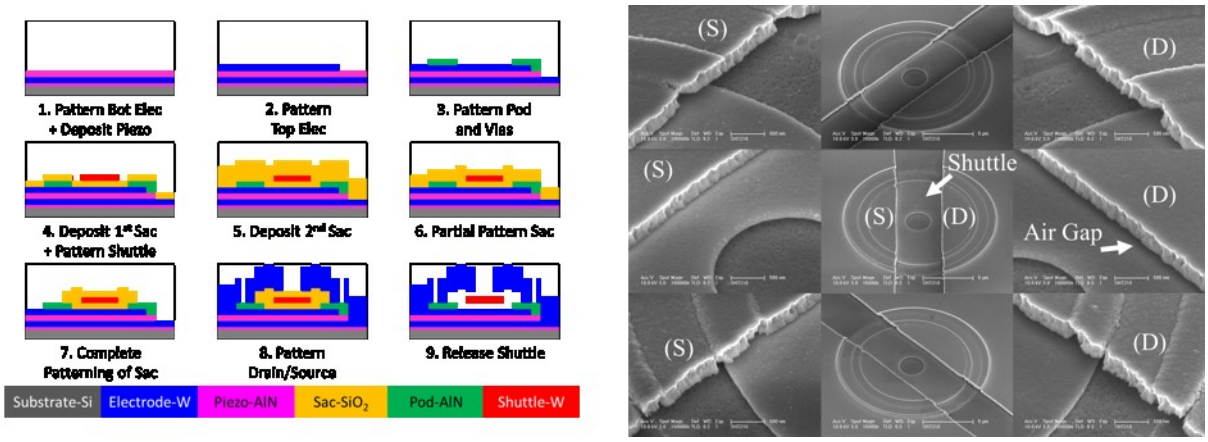

Figure 18. Fabrication of the PAPS relay. The relay is fabricated in 9 steps. SEM images show a fabricated and released PAPS relay from different angles. S and D represent the source and drain contacts.

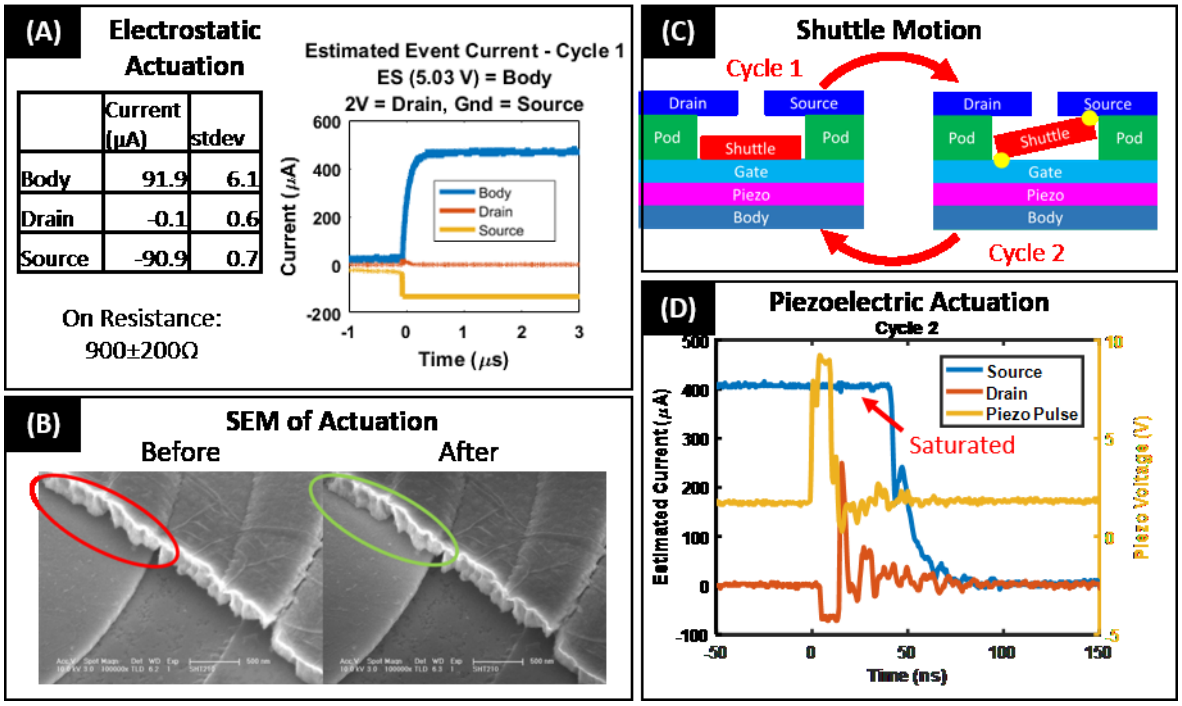

Figure 19. Testing summary for the PAPS relay highlighting (a) electrostatic only switching, (b) SEM images of shuttle motion, (c) schematics of shuttle motion, and (d) Piezoelectric actuation with electrostatic bias.

The PAPS device showed promise as a highly scalable NEMS relay but was unable to demonstrate reliable or predictable switching. An iteration of the PAPS device would require elimination of the tilting instability in the shuttle. One alternative to the shuttle design would be to directly open and close the air gap with vertical expansion of the actuator. The piezoelectric actuator of the PAPS, which expands in the vertical direction, is unable to generate the strain required to directly close the air gap. An alternative actuation material, capable of expanding and contracting enough to open and close the air gap directly, would be required for this iteration of the PAPS device.

1.4 Thesis Outline

This thesis will discuss the concept, design, fabrication, and testing of a novel iteration of the PAPS device, where the expansion of GeTe is used to directly actuate the air gap of a relay. This novel relay, the Phase Change NEMS Relay (PCNR) is a non-volatile device that addresses the limited scalability of traditional NEMS relays.

Chapter 2 will introduce the phase change NEMS relay concept. The operating principle of the relay will be described, including the theory on GeTe and other phase change materials. The modeling process will be described in detail, ranging from a 1D analytical model to varied COMSOL finite element models. Additionally, a scaling analysis will be presented to motivate an aggressively scaled PCNR. Scaling analysis will be used to predict the minimum dimensions and actuation metrics of a CMOS scale device for future iterations of the PCNR.

Chapter 3 will describe the fabrication process development for the PCNR. Key challenging fabrication steps will be described in depth. In particular, considerations for a useable GeTe film will be described, including ensuring proper composition, low roughness, and smooth sidewalls. Other features to be described include defect free deposition of an encapsulation layer and the formation of a narrow 20 nm air gap.

Chapter 4 will describe the testing of a phase change actuator. Optical measurements will be used to determine the presence of amorphous GeTe after actuation. Additionally, optical measurements will be used to rapidly test a range of actuators to determine the optimal actuation and recrystallization pulse. Finally, the actuator will be measured by AFM to determine the expansion and set the required air gap for full PCNR fabrication.

Chapter 5 will describe testing the fully fabricated PCNR. Testing will highlight key performance metrics of the PCNR including actuation voltage, actuation time, on resistance, non-volatility, reliability,

recrystallization voltage, recrystallization time, and leakage. Residual contamination in the air-gap will be characterized by TEM. Finally, the COMSOL model will be compared to the fabricated PCNR devices.

Chapter 6 will summarize the work of this thesis and state key contributions. Future work on the PCNR will be described including possible alternate materials for fabrication and an alternative geometry to improve the performance of an aggressively scaled PCNR.

2 Phase Change NEMS Relay (PCNR) Concept Design and Modeling

2.1 Abstract

The phase change NEMS relay (PCNR) is introduced as an alternative non-volatile relay that addresses the limited scalability of traditional mechanical relay designs while maintaining the high on-off ratio of mechanical contacts. The PCNR is built around the volumetric change observed in chalcogenide glass phase change materials when shifting between amorphous and crystalline states. These phase change materials support both amorphous and crystalline states at room temperature and are characterized by the rapid crystallization at elevated temperatures. Switching of the PCNR is performed by heating, melting, and quenching the phase change material into the expanded amorphous state. Expansion of the phase change material connects metallic contacts, leading to low on-state resistance. Turning the PCNR off is performed by heating, but not melting the phase change material to promote rapid crystallization into the contracted crystalline state. Contraction of the phase change material separates the metallic contacts, leading to a well isolated relay. GeTe is determined to be the best candidate phase change material for use in the PCNR due to its large volumetric expansion (10 %), higher temperature tolerance (amorphous for greater than 10 years at 125 °C), and rapid crystallization (less than 60 ns).

PCNR devices are modeled with both a 1D analytical model and a full COMSOL finite element analysis model. Electrical and thermal physics are combined to determine switching pulse requirements for a range of scaled geometries. Additionally, 2D COMSOL models of the phase change material expansion show the effects of cap thickness and melted PCM width. Scaling trends are fit to both the analytical and COMSOL models. Analytical models do not match the full COMSOL simulation, however lateral heat flow provides an explanation for analytical model deviation. Models predict an aggressively scaled relay with a 5 nm x 20 nm heater and a 2 nm air-gap is capable of switching at only 0.52 V with a quench time of 1.8 ns and a total switching energy of 1.8 pj.

2.2 Operating Concept

The phase change NEMS relay (PCNR) is built to address the limited scalability drawback of existing relay designs, while maintaining the high on-off ratio of mechanical contacts, with the added feature of non-volatility. The PCNR is based around the mechanical expansion of phase change material (PCM) when converting from a crystalline state to an amorphous state. The PCNR consists of two main components (Figure 20): a contact pair, and a GeTe based phase change actuator. The contact pair is formed by a metallic contact adhered to the top surface of the actuator and a suspended pair of drain/source contacts. The actuator is comprised of a resistive heater, and a layer of phase change material encapsulated in an insulating cap. The actuator expands and contracts depending on whether the GeTe is in the crystalline (smaller) or amorphous (larger) states.

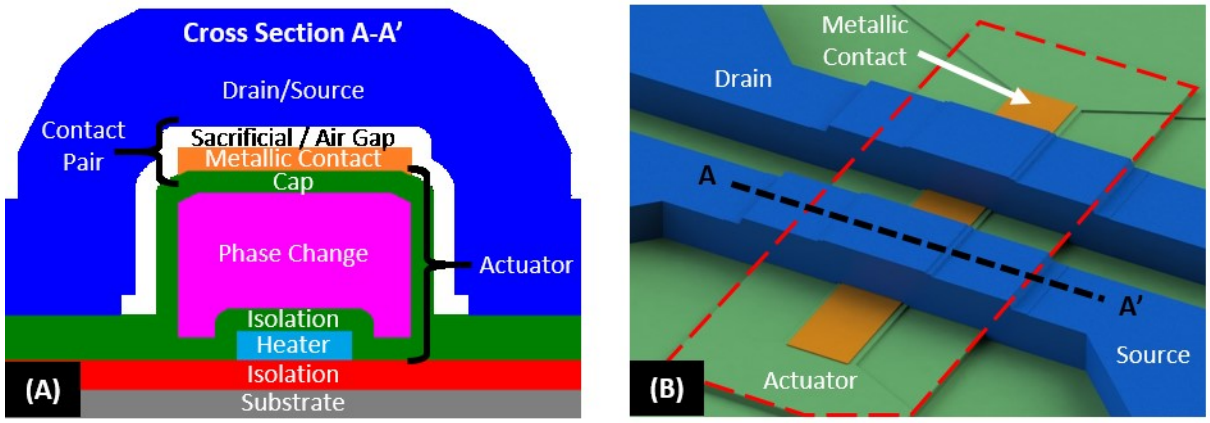

Figure 20. Schematic representation of the phase change NEMS relay in (a) 2D cross section and (b) 3D isometric views. The actuator and contact pair are the two main components of the PCNR.

The PCNR operates following the process shown in Figure 21. The PCNR is initially fabricated in the off state where the drain and source electrodes are isolated from the metallic contact by an air gap. In the off state the phase change material is in the crystalline state. To switch to the on state, the heater is used to melt a portion of the phase change material, which is then quenched, converting the PCM to the larger amorphous state. The PCM is liquid for a short period before quenching. Quench time is important to ensure the PCM remains in the amorphous state. If cooled too slowly, the PCM will crystallize and the relay will not switch to the on state. The expansion of the actuator presses the metallic contact into the drain and source, connecting the drain and source, and thus turning the relay on.

A similar process is followed to switch the PCNR back to the off state, however the PCM is not melted. The heater is used to raise the PCM temperature without melting to promote rapid crystallization. The PCM remains solid throughout the process of switching the PCNR to the off state. The crystallization process contracts the actuator, separating the metallic contact from the drain and source, creating an air gap and turning the relay off.

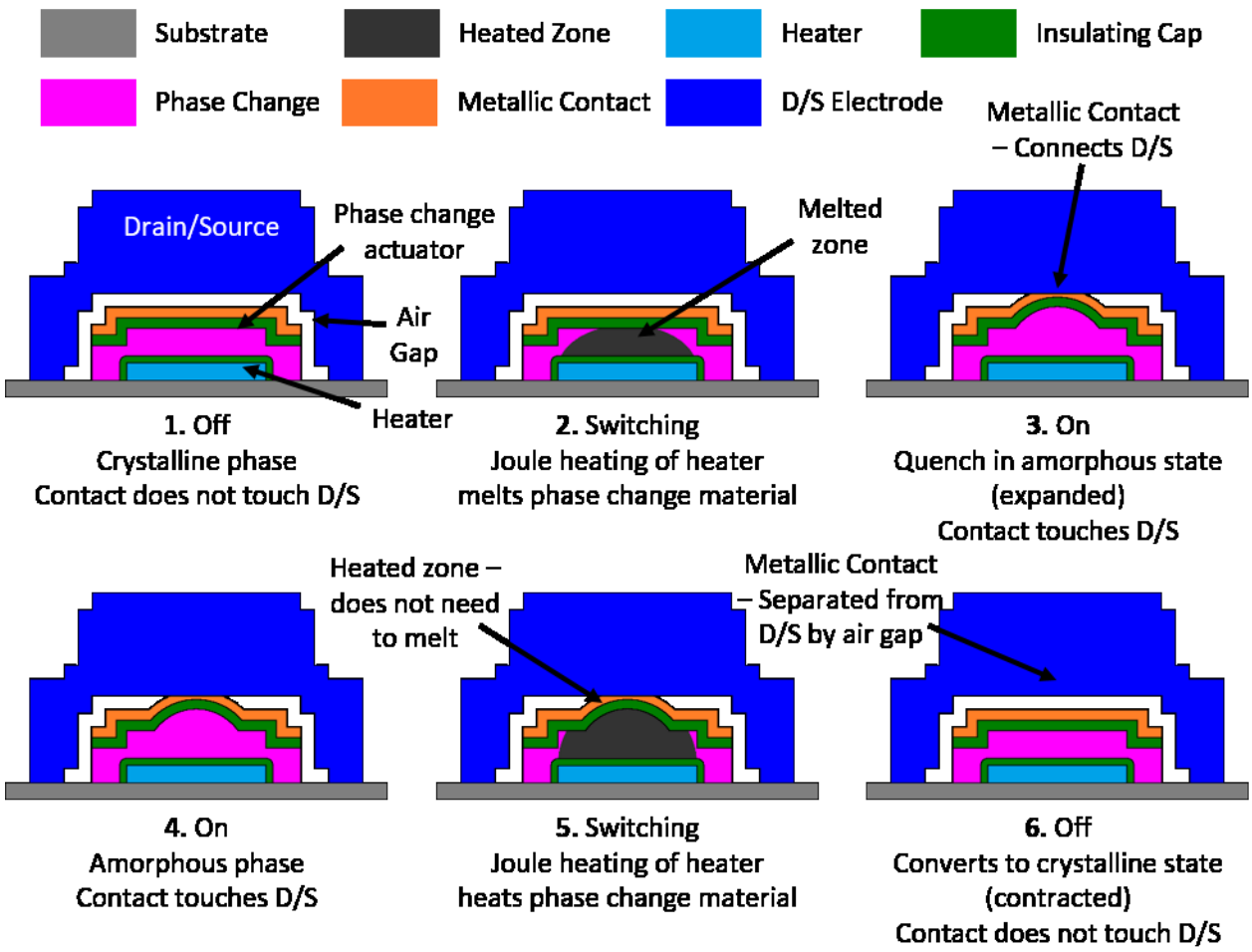

Figure 21. Switching process for turning the phase change NEMS relay on and off.

2.3 Phase Change Materials

Phase change materials (PCM) are a broad category of materials encompassing any material that undergoes a change in atomic structure following an outside force [67]. Many of these phase changes encompass a large change in volume which can be harnessed in a mechanical actuator [67]. Shape memory alloys, a subset of PCMs, have been demonstrated as the active element in a number of cantilever based actuators [68][69]. These shape memory based actuators do not exhibit non-volatile states and must have power or heat applied to maintain an actuated state. Another subset of phase change materials is chalcogenide glass. Chalcogenide glasses are glasses containing elements from the oxygen family of

elements: oxygen, sulfur, selenium, tellurium, and polonium. Common chalcogenide glass phase change such as GeTe and GST were originally developed for use in optical storage technologies [70][71], and later for resistivity based phase change memories [7]. These materials are characterized by room temperature stable amorphous and crystalline states, which can be used for non-volatile applications. When transitioning from one state to another, three material properties change: reflectivity, resistivity, and density. Optical storage takes advantage of the shift from reflective to opaque when switching from crystalline to amorphous states. PC-RAM takes advantage of the shift from conductive to resistive when switching from crystalline to amorphous. The change in volume has yet to be used in a non-volatile device as an actuator, although the change in stress has been used in FinFETs for stress enhancement [72][73]. When changing from crystalline to amorphous phase change materials expand by 5 – 10 % [74].

GeTe is a chalcogenide glass phase change material of particular interest for use in the PCNR. GeTe exhibits one of the largest changes in volume (10 %)[73] in any non-volatile PCM. Figure 22 compares intrinsic work density for GeTe and other materials commonly used in MEMS actuators. GeTe exhibits extremely high work density (200 J/cm^3) with both higher strain and displacement than other comparable materials. GeTe has a higher crystallization temperature compared to GST (another common PCM) and is expected to maintain non-volatility for over 10 years at 125 °C [75][76]. Phase change materials like GeTe are also characterized by their fast crystallization time (less than 60 ns) [77]–[81]. The large volume change, elevated temperature non-volatility, and fast crystallization time make GeTe an ideal material for the actuator in the PCNR.

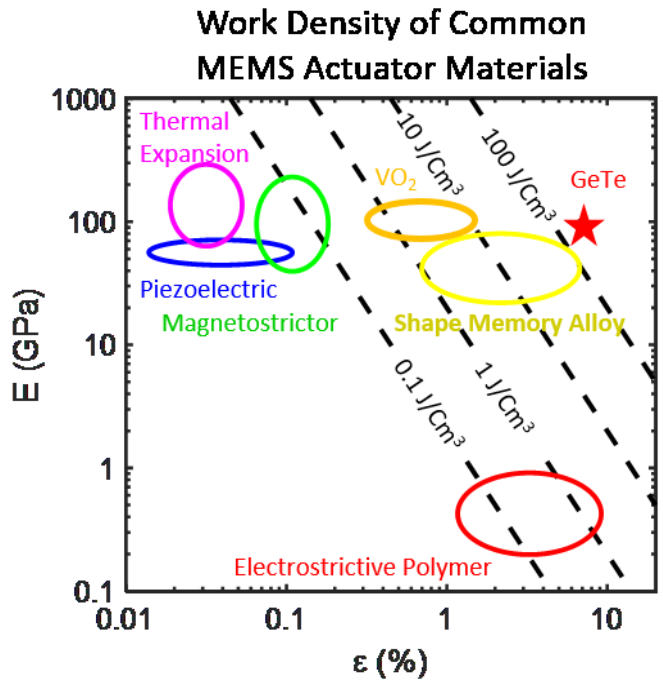

Figure 22. Comparison of work densities for common materials used in MEMS actuators. Adapted from [68].

2.4 Device Modeling

A 1D analytical model is used to quickly model operating voltages and scaling laws. Additionally, PCNR devices are modeled in COMSOL Multiphysics finite element analysis software to account for 3D effects and non-idealities in the geometry. Electrical and thermal effects are combined in one model to account for joule heating within the heater, while mechanical expansion and stress are separately modeled to optimize solving time.

2.4.1 Electrical Modeling

Electrical circuit modeling includes the pulse generator's 50 Ω internal impedance. Figure 23 shows the electrical circuit and probe locations for the simulated device. The probing pads are modeled in addition to the heater pads. Even though the pads and traces are substantially wider than the heater, the additional length leads to comparable resistances. Eventually, this additional resistance can be

eliminated in an advanced fabrication process. It is present in the currently fabricated devices and that is why it is included in the model. A thin isolation layer is conformally deposited over the heater layer to electrically separate the PCM from the heater. Capacitive coupling between the heater and PCM was found to be negligible, so only the heater is modeled for electric currents. Pulse generator voltage (V_{PG}) is related to heater voltage (V_{heater}) by:

$$V_{heater} = 2V_{PG}\left(\frac{R_{heater}}{50\Omega + \mathrm{R_{traces}} + R_{Heater}}\right) \tag{13}$$

where R_{heater} is the heater resistance, and R_{traces} is the total resistance of the traces and probes connected to the heater. V_{PG} is the output voltage of the pulse generator assuming a 50 Ω load, so the high impedance output of the pulse generator is double the voltage. Thermal energy is generated by joule heating, and is modeled by:

$$P_{heater} = \frac{V_{heater}^2}{R_{heater}} \tag{14}$$

In the analytical model heat is assumed to be dissipated evenly over the heater as temperature and resistivity are assumed to be uniform over the heater.

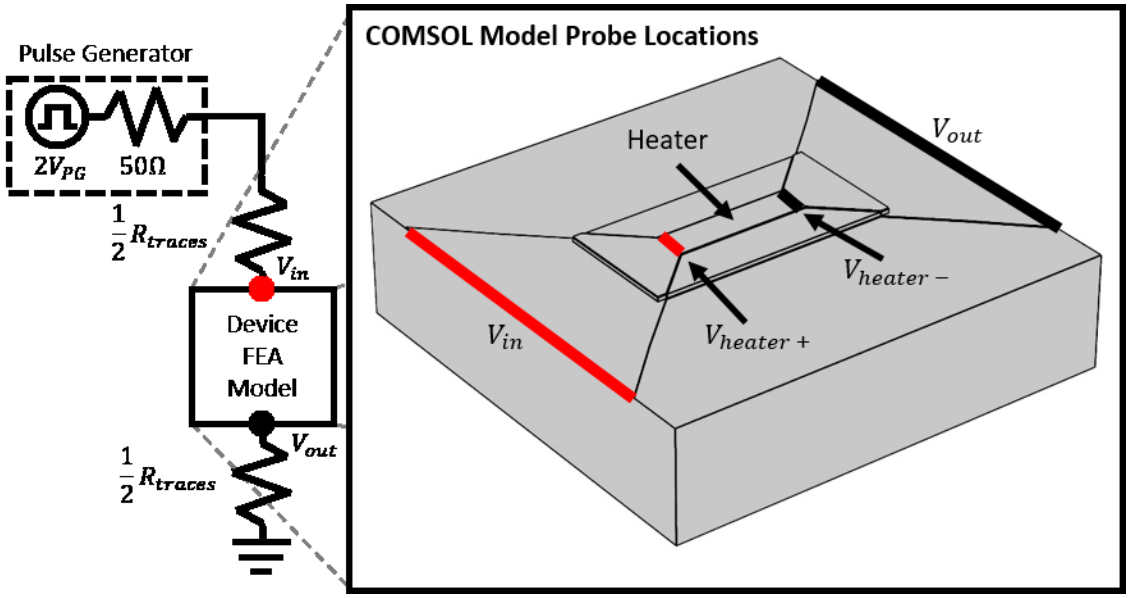

Figure 23. Heater actuation circuit model and probe locations in the COMSOL model.

Change in resistance as a function of temperature must be included for an accurate estimate of energy dissipation in the heater. Heater resistance accounting for elevated temperatures is modeled as:

$$R_{heater} = R_{0_{heater}}\left(1 + \alpha_{heater}(T_{heater} - T_0)\right) \quad (15)$$

where α_{heater} is the temperature coefficient of resistivity (TCR) of the heater material, T_{heater} is the heater temperature, and T_0 is the reference temperature for resistivity and TCR measurements. $R_{0_{heater}}$ is calculated by:

$$R_{0_{heater}} = \rho_{heater}\left(\frac{l_{heater}}{w_{heater}t_{heater}}\right) \quad (16)$$

where ρ_{heater} is the resistivity of the heater material and l_{heater}, w_{heater}, and t_{heater} are the length, width, and thickness respectively of the heater. Heater resistivity and TCR are measured for the tested fabricated device. Figure 24 shows resistivity measurements of the tungsten for a range of substrate temperatures. Resistance is measured as a function of trace length on a serpentine test structure and fit with a linear trend to account for probe contact resistance. Resistance was measured for temperatures ranging 170 K – 370 K in a Lakeshore Cryogenic probing station. TCR is estimated by fitting a linear trend to the resistance at each temperature. Resistivity and TCR referenced to 293 K are measured to be 1.67×10^{-7} Ωm and 0.17 %/K respectively. The TCR of tungsten has been shown to be constant up to the melting point [82], so measurements around the operating point were deemed unnecessary.

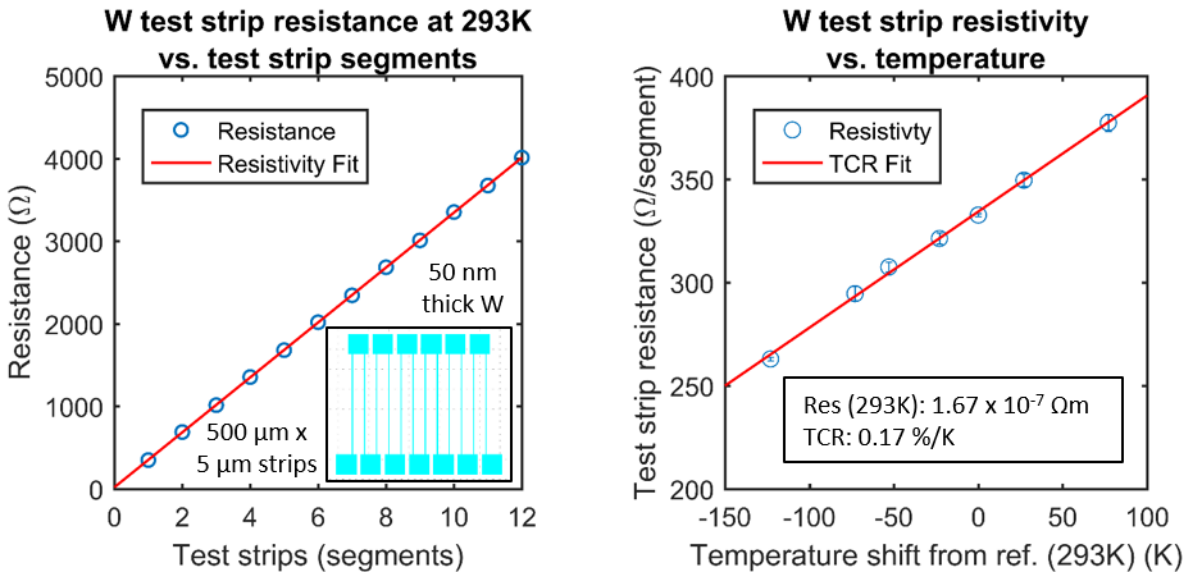

Figure 24. Resistivity and temperature coefficient of resistivity measurements for heater tungsten.

2.4.2 Thermal Modeling

Thermal modeling is performed to characterize two key aspects of the actuator: temperature distribution and quench time within the PCM. Temperature distribution is important to control and limit the melted area of the PCM. Melted PCM must be rigidly encapsulated in the lateral dimensions to force the isotropic volume changes into only the vertical direction of the actuator. The melted PCM can be laterally encapsulated by solid PCM. Quench time is important to ensure the PCM does not crystalize before cooling to room temperature.

Thermal resistance is generated from two sources: intrinsic material thermal conductivity and thermal boundary resistance (TBR). For the 1D analytical model, resistance through a layer can be modeled as:

$$R_{th} = \frac{t_{layer}}{l_{layer} w_{layer} k_{th}} \tag{17}$$

where R_{th} is the thermal resistance of the layer, k_{th} is the thermal conductivity of the layer material, and l_{layer}, w_{layer}, and t_{layer} are the length, width, and thickness of the layer respectively. Thermal boundary

resistance (TBR) also known as thermal interface resistance is modeled as a fixed thermal resistance per area at the interface between each layer. TBR can be quite substantial and even a dominant factor in thin materials with a large lattice mismatch [83]–[85]. Thermal resistance through a boundary between two layers (R_{TBR}) can be modeled as:

$$R_{TBR} = \frac{TBR}{A_{layer}} \tag{18}$$

where A_{layer} is the lateral area of the layer and TBR is the thermal boundary resistance. Figure 25 shows the analytical thermal circuit model. The substrate is modeled as a fixed temperature heat sink, with TBRs separating the heater from the substrate and the PCM from the heater. Thermal resistance of the layers is neglected due to the high values of the TBRs.

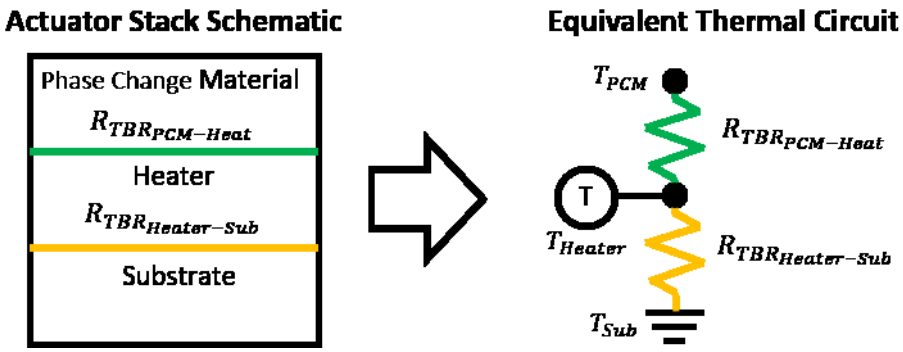

Figure 25. Thermal circuit model used for 1D analytical steady state modeling.

A thermal model was also built using COMSOL Multiphysics finite element analysis software. The model is a 3D geometry that includes the heater, PCM, and insulating layers. Figure 26 shows the geometry in 3D and 2D cross-section views. The cap and metallic contact are not modeled as the latent heat of the PCM is dominant for thermal timing. GeTe has a thermal conductivity on the order of 0.2-2 W/mK [86][87], while ALD Al_2O_3 and sputtered W have thermal conductivities of 2 W/mK [88] and 40-70 W/mK [89], respectively. Reducing the number of thin layers in the model dramatically reduces the number of elements required to accurately represent the model, thus reducing simulation time.

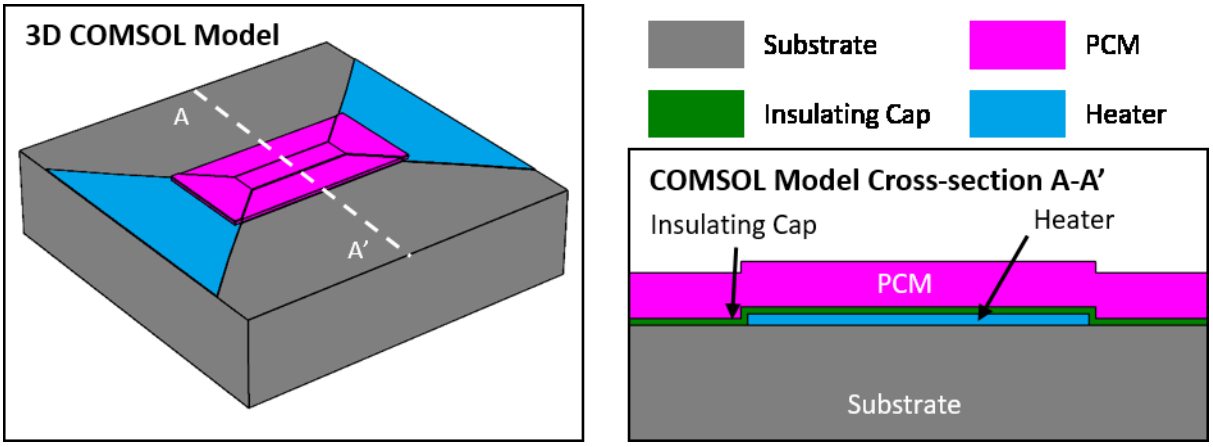

Figure 26. 3D and 2D cross-section views of the COMSOL model. Sections are colored by material.

As a material is melted, it absorbs large amounts of heat due to the latent heat of fusion while maintaining a constant temperature. Melting can be modeled as a sharp increase in heat capacity starting at the melting temperature such that the added heat equals the latent heat of fusion [90]. Melting is modeled to occur over a narrow temperature range, rather than a fixed temperature to ensure the simulation converges. The area of melted PCM is defined as areas of PCM above the melting transition temperature. Figure 27 shows the progression of the melted zone during an actuation pulse. The steady state melted zone is tightly confined to the heater. Substrate thermal conductivity is substantially higher than the PCM thermal conductivity, so heat is transferred down into the substrate before it can travel laterally to melt a wide area of the PCM. The melted zone boundary starts at the heater and travels upwards to the top surface of the PCM. The cooling and quenching process cools from the bottom up, as heat stored in the PCM must transfer through itself to the heatsinking substrate.

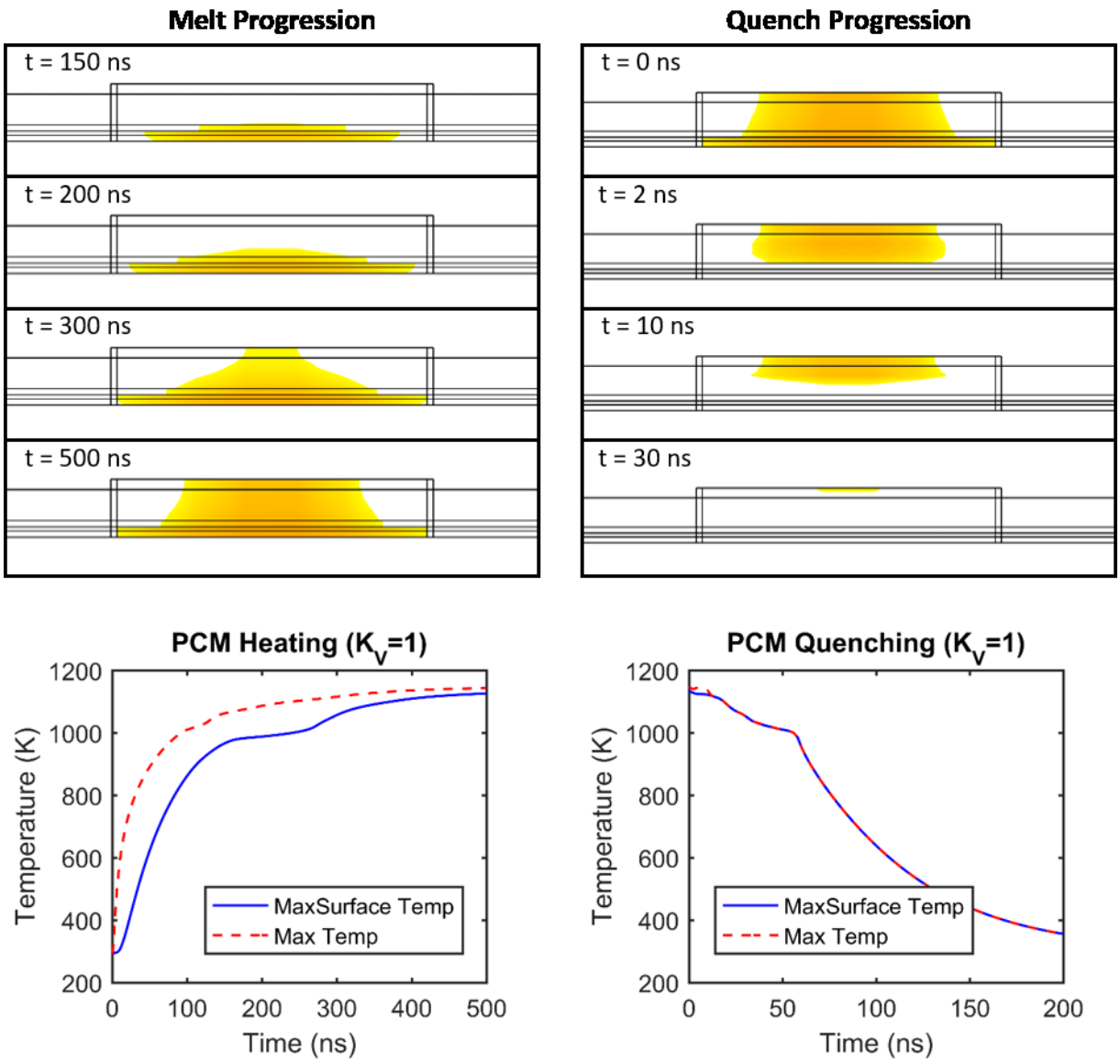

Figure 27. Progression of melted PCM during heating and cooling of the PCNR actuator. Colored sections indicate melted PCM. Graphs show maximum temperatures of the actuator surface and internal PCM.

2.4.3 Mechanical Modeling

The PCNR actuator is designed to confine melted PCM within a cap layer to ensure volumetric expansion leads to maximum vertical expansion of the actuator. Mechanical modeling was limited to expansion of the PCM within the cap layer and does not account for contact between the actuator and the drain and source. Models are built in 2D rather than 3D due to the long aspect ratio of the melted

zone. Areas in the center, where the drain and source are located, can be modeled as a plane strain 2D model. Plane strain specifies no deflection in the out of plane direction (equivalent to a fixed boundary). Figure 28 shows the mechanical actuator model built in COMSOL. Phase change material is surrounded by a cap layer and is attached to the substrate. The expansion of the PCM is modeled as a liquid in steady state. When the PCM is modeled as a liquid, a pressure is generated in the cavity corresponding to the peak pressure generated by the volumetric expansion of the PCM. Two cases of liquid PCM are modeled: PCM within the cap is partially melted and surrounded by solid PCM; and all PCM inside the cap is liquid.

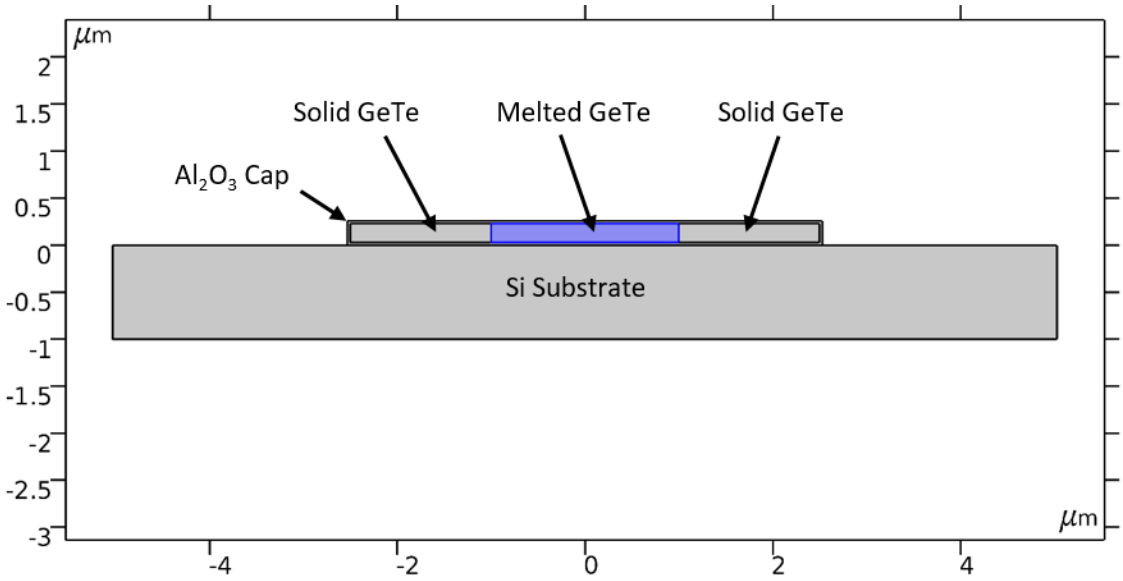

Figure 28. Mechanical actuator 2D model. GeTe is encapsulated in Al_2O_3 and is on top of an Si substrate.

Figure 29 shows the mechanical actuator model setup for partially melted PCM surrounded by solid PCM. A pressure is specified at the boundary between the liquid PCM and the solid PCM and cap to model the hydrostatic pressure of melted PCM. Hydrostatic pressure ($P_{HS_{PCM}}$) is related to the bulk modulus (K_{PCM}) of the PCM and the volumetric strain ($\frac{dV_{PCM}}{V_{PCM}}$) of the PCM by:

$$P_{HS_{PCM}} = \frac{dV_{PCM}}{V_{PCM}} K_{PCM} \tag{19}$$

The initial conditions of the simulation assume the PCM is compressed by the volumetric expansion associated with the change from crystalline to amorphous states (10 % for GeTe) – modeling PCM that has just been melted. As volume inside the cap expands to the amorphous state volume, pressure inside the cap approaches zero.

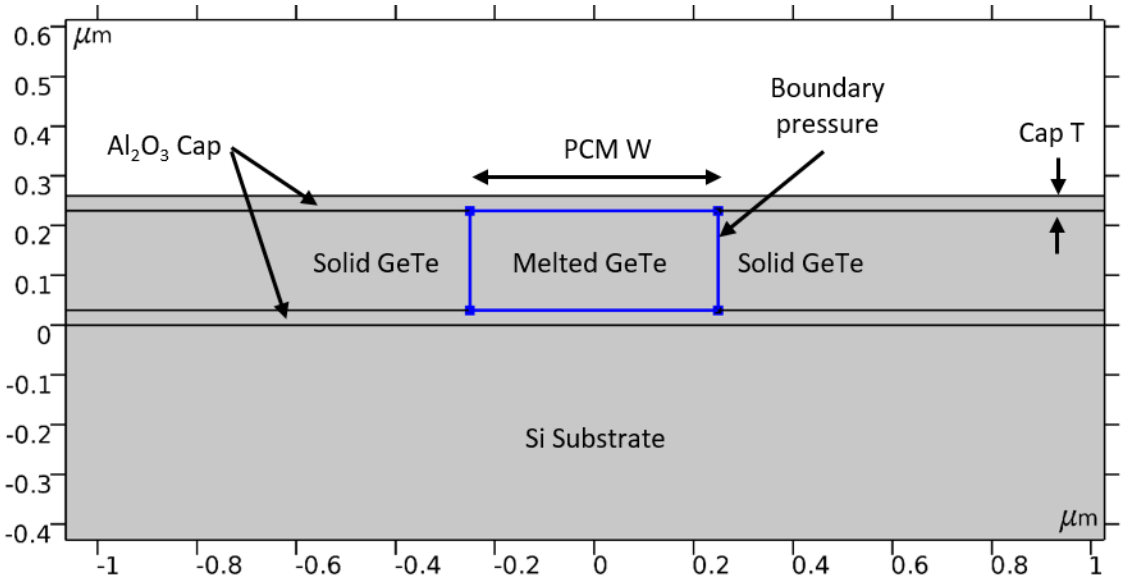

Figure 29. Mechanical actuator model setup for partially melted GeTe. A boundary load (blue) is applied at the boundary of the melted GeTe to simulate hydrostatic pressure.

Figure 30 shows hydrostatic pressure and maximum displacement for sweeps of melted PCM width and cap thickness. Low pressure is desirable to limit the stress on the cap and enable maximum displacement of the actuator. Since pressure is proportional to volumetric strain an actuator will exhibit the largest displacement with the lowest hydrostatic pressure. A thin cap layer resists the expansion of the PCM the least and therefore exhibits the largest displacement, although at the scales of this device the cap layer and melted PCM width have little effect on the maximum displacement. Hydrostatic pressure generated by the PCM expansion scales as:

$$P_{HS_{PCM}} \propto \left(\frac{t_{cap}}{w_{cap}}\right)^3 \quad (20)$$

where t_{cap} and w_{cap} are the thickness and width of the cap layer respectively. This intuitively makes sense as the stiffness of a clamped-clamped beam with a uniformly distributed load follows the same scaling.

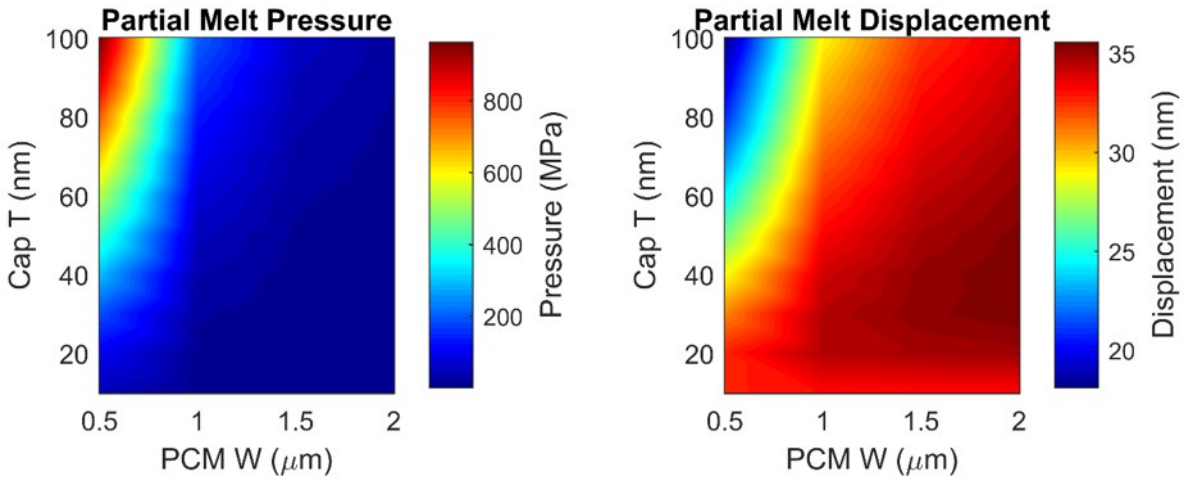

Figure 30. Hydrostatic pressure and actuator displacement for varied melted phase change material widths and cap thicknesses. The melted PCM is constrained on the sides by solid PCM. PCM W refers to the width of the melted PCM.

Additional modeling shows a cap with fully melted PCM (Figure 31). In this scenario all PCM within the cap is melted, so the side walls are the flexible cap layer instead of the rigid solid PCM. Figure 32 shows hydrostatic pressure and maximum displacement for sweeps of actuator width and thickness. Much like the solid PCM constrained case, a thinner cap layer resists the expansion of the PCM the least, leading to maximized expansion of the actuator. There is little difference in expansion compared with liquid PCM constrained by solid PCM. Melting all PCM inside the cap is still less preferred due to the cap likely being thinner and weaker at step coverage over the PCM – increasing the likelihood of cap failure.

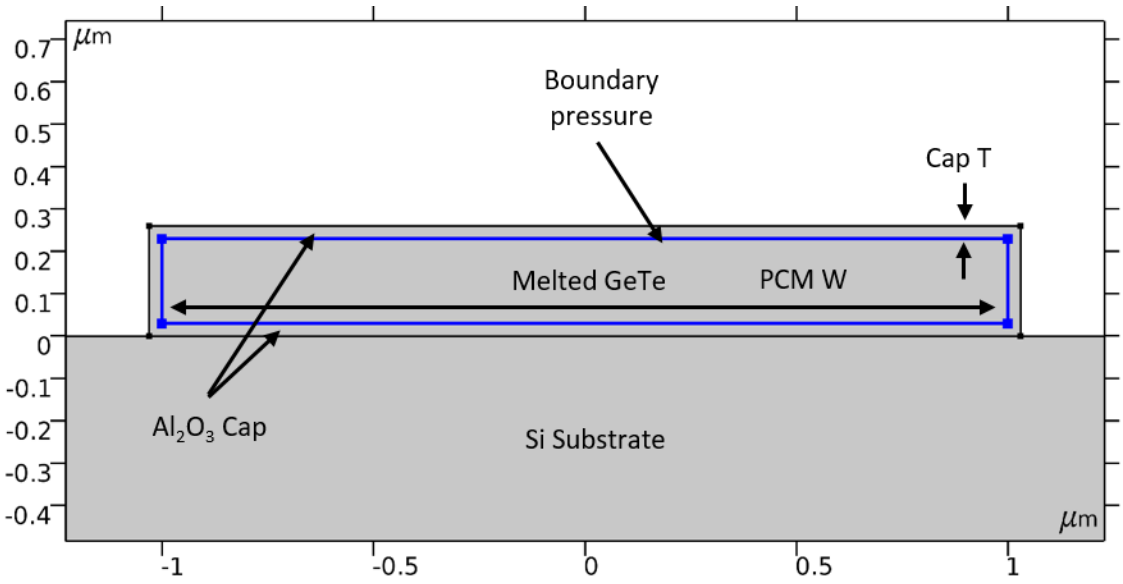

Figure 31. Mechanical actuator model setup for fully melted GeTe. A boundary load (blue) is applied to the entire cap to simulate the hydrostatic pressure of the entirely melted GeTe.

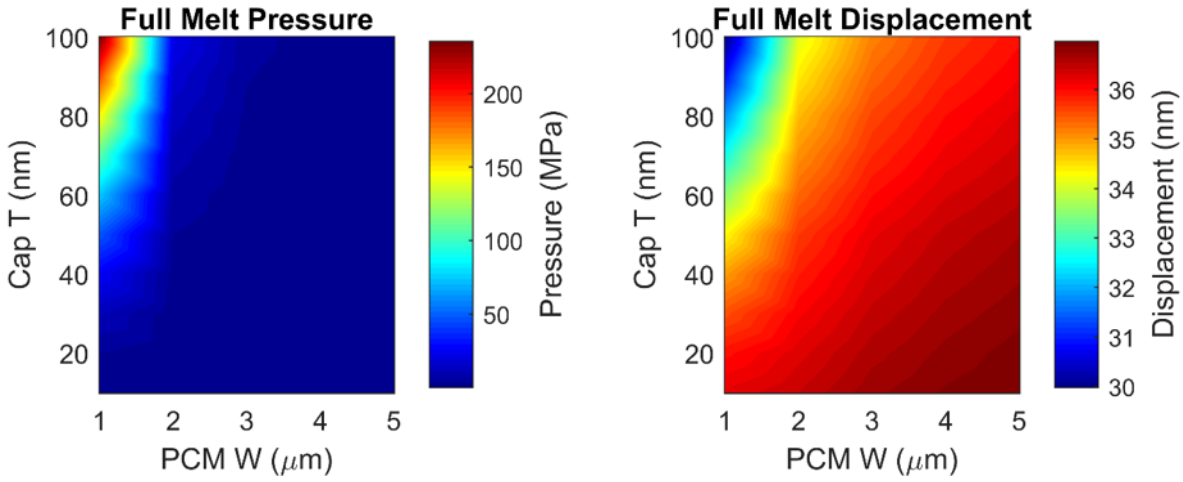

Figure 32. Hydrostatic pressure and actuator displacement for varied PCM widths and cap thicknesses. All PCM is melted, so PCM W refers to the entire actuator width.

2.5 Scaling Analysis

Scaling trends are derived based on a 1D thermal model of the actuator stack (Figure 25) and verified by simulation. A few key performance metrics were analyzed: actuation voltage, current, power,

and heater resistance. All parameters are measured over the heater, so voltage drop or power dissipated in the traces is not measured. Two models were created in COMSOL, a simple and full model. The simple model is a stack of layers representing the substrate, heater, isolation, and phase change material. Thermal boundary layers are included between each layer in the stack. The simple model closely represents the assumed conditions for the analytical model: all heat flow is vertical through the substrate, thermal boundary resistance is the dominant thermal resistance, and heater temperature is uniform. The full COMSOL model includes effects from lateral heat flow and heating of the substrate. Sweeps of applied voltage were used to determine the actuation voltage. Actuation voltage was found by determining the pulse required to bring the PCM surface temperature above 1100 K (100 K higher than the melting point of GeTe). All models are based on a set geometry with dimensions stated in Table 1. This geometry is representative of fabricated and tested devices. Other variations in geometry are possible, but not investigated in the following scaling study.

Table 1. Nominal Geometry used for scaling analysis of the PCNR.

PCNR Nominal Geometry

Geometry	Dimensions
Heater t x L x W	50 nm x 6 µm x 1.5 µm
PCM t x overlap	200 nm x 2 µm
Cap t	30 nm
Air Gap	20 nm

Isotropic scaling has all dimensions (length, width, and thickness) scale equally. If the scale factor (K_V) is 10, then all dimensions are 10 times smaller ($\frac{1}{K_V}$). For the analytical model, heater temperature is assumed to be constant for scaling. PCM is only connected to the substrate by the heater, so thermal

resistance (defined by the TBR between the heater and substrate) sets the power. Thermal resistance scales as:

$$R_{TBR} \propto K_V^2 \tag{21}$$

To reach a constant temperature, input power scales as:

$$P_{heater} \propto \frac{1}{K_V^2} \tag{22}$$

Heater resistance scales with the number of "squares" of material since the scaling assumes a constant heater temperature and can be modeled as:

$$R_{heater} \propto K_V \tag{23}$$

Voltage and current for the heater are set by the input power and heater resistance, so heater voltage scales as:

$$V_{heater} \propto \frac{1}{\sqrt{K_V}} \tag{24}$$

and heater current scales as:

$$I_{heater} = \frac{1}{K_V^{\frac{3}{2}}} \tag{25}$$

Figure 33 shows a comparison between simple and full COMSOL simulations for isotropic scaling. Power functions fit well for both the simple and full models, however exponents deviate for the two models. Heater resistance shows a greater increase with scaling for the full model compared to the simple model. This is caused by changes in temperature distribution over the heater. The full model exhibits non-uniform temperature over the heater caused by heat flow through the traces connecting to the heater. Additionally, the substrate does not act as a perfect heatsink. Required actuation power scales less aggressively for the full model. This is due to lateral heat flow in the actuator, preventing power from scaling proportionally to device area.

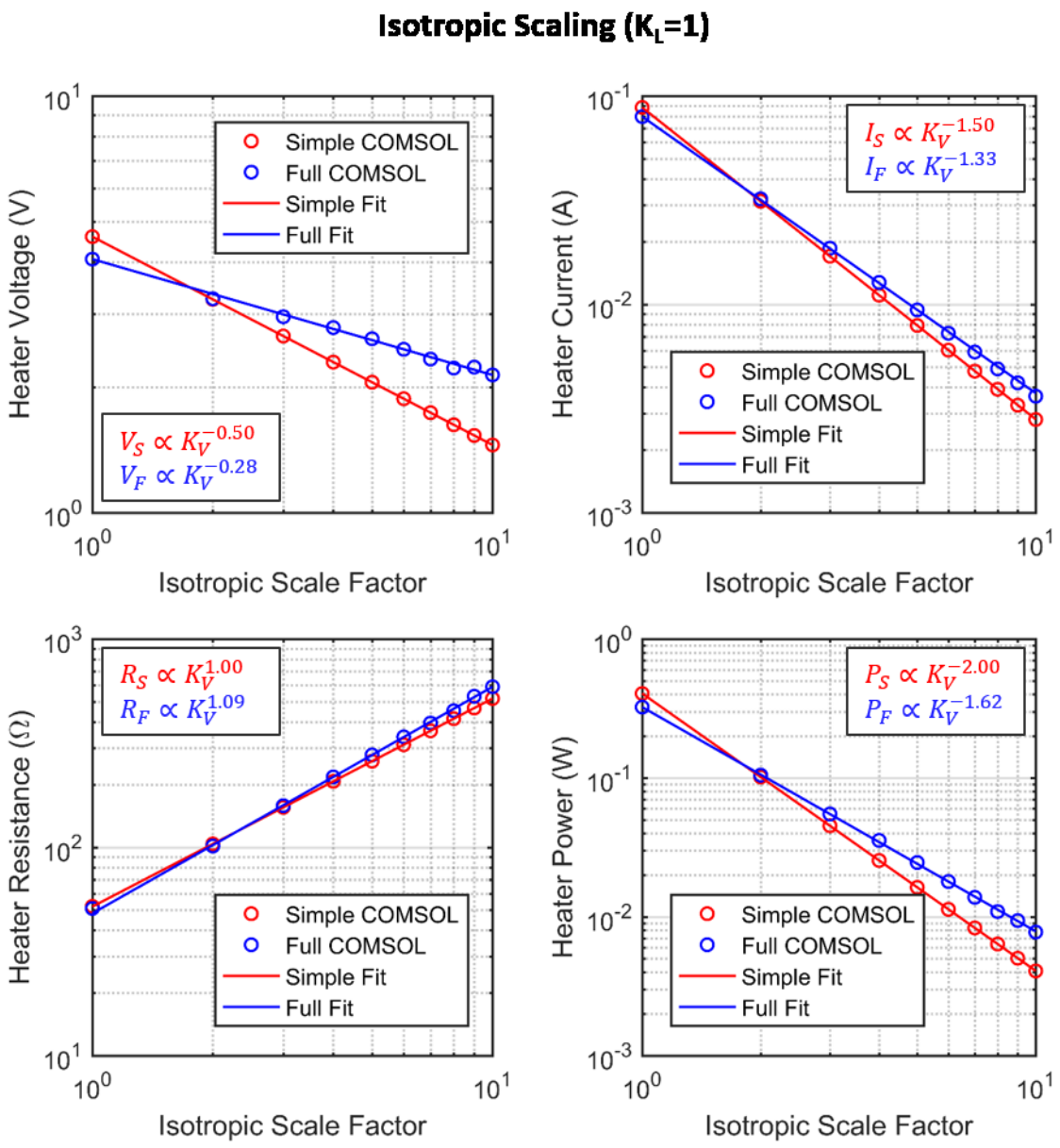

Figure 33. Isotropic scaling trend extraction for simple and full COMSOL models. Voltage, current, resistance, and power are all measured at the heater when heating the surface of the PCM to just above the melting point of the PCM.

Figure 34 compares extracted heater temperature and PCM temperature scaling for the full COMSOL simulation. PCM temperature remains constant as a PCM temperature above 1100 K is defined as switching. The difference between the heater and PCM temperatures increases as the device is scaled

in the full COMSOL simulation. This difference can be explained by changes in relative thermal resistance between the heater, PCM, and substrate. In the 1D analytical and the simple COMSOL models, the PCM is only over the heater, so the PCM temperature must match the heater temperature in the steady state as all heat from the PCM must flow through the heater layer. In the full model the PCM is connected to both the heater and the substrate, allowing the PCM to have a lower temperature than the heater. Heat can flow directly from the PCM to the substrate. The 2D schematic in Figure 34 illustrates the various heat paths and scaling of the thermal resistances. Lateral thermal resistance in the PCM is proportional to:

$$R_{lateral} \propto \frac{d}{t} \tag{26}$$

where $R_{lateral}$ is a thermal resistance per unit length, d is the distance traveled by the heat, and t is the thickness of the PCM. For isotropic scaling, lateral thermal resistance scales as:

$$R_{lateral} \propto 1 \tag{27}$$

while the TBR per unit length scales as:

$$R_{TBR} \propto K_V \tag{28}$$

As the device is scaled down the lateral thermal resistance in the PCM becomes less dominant, so a higher temperature in the heater is required to melt the PCM.

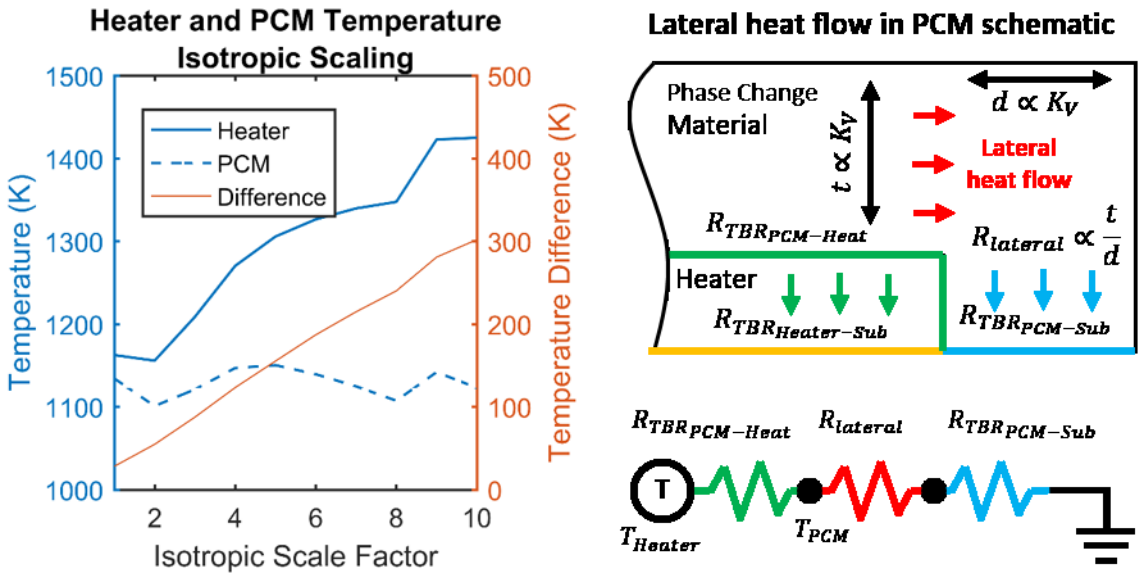

Figure 34. Heater and PCM temperature isotropic scaling. Colored arrows indicate heat flow through each thermal resistance. Deviation in heater and PCM temperature is explained by lateral heat flow (red arrows) in the 2D schematic of an actuator.

Isotropic scaling is limited by the air-gap. Further scaling must be lateral to reduce the footprint, while maintaining the air-gap. Lateral scaling has only planar dimensions (length and width, not thickness) scale equally. When the lateral scale factor (K_L) is 10, then length and width are 10 times smaller ($\frac{1}{K_L}$). For the 1D analytical model, input power is only dependent on the TBR between the heater and substrate. Since the TBR is independent of layer thickness, power scales as:

$$P_{heater} \propto \frac{1}{K_L^2} \tag{29}$$

and is the same as isotropic scaling. Heater resistance scales with the number of squares. Since thickness is constant, heater resistance scales as:

$$R_{heater} \propto 1 \tag{30}$$

Heater voltage and current are dependent on the input power and heater resistance. Heater voltage scales as:

$$V_{heater} = \frac{1}{K_L} \tag{31}$$

while heater current scales as:

$$I_{heater} = \frac{1}{K_L} \tag{32}$$

Although input power scales the same for isotropic and lateral scaling, heater resistance scales differently. The lateral scaling maintains constant resistance, leading to a faster decrease in required heater voltage for actuation.

Figure 35 shows a comparison between simple and full COMSOL simulations for lateral scaling on geometry that has already been scaled isotropically by a factor of 10. Power functions fit well for all parameters in the simple model, but poorly for heater voltage, current, and resistance in the full model. The poor fit can be attributed to two distinct scaling trends in the heater resistance, which transition at a lateral scale factor of four.

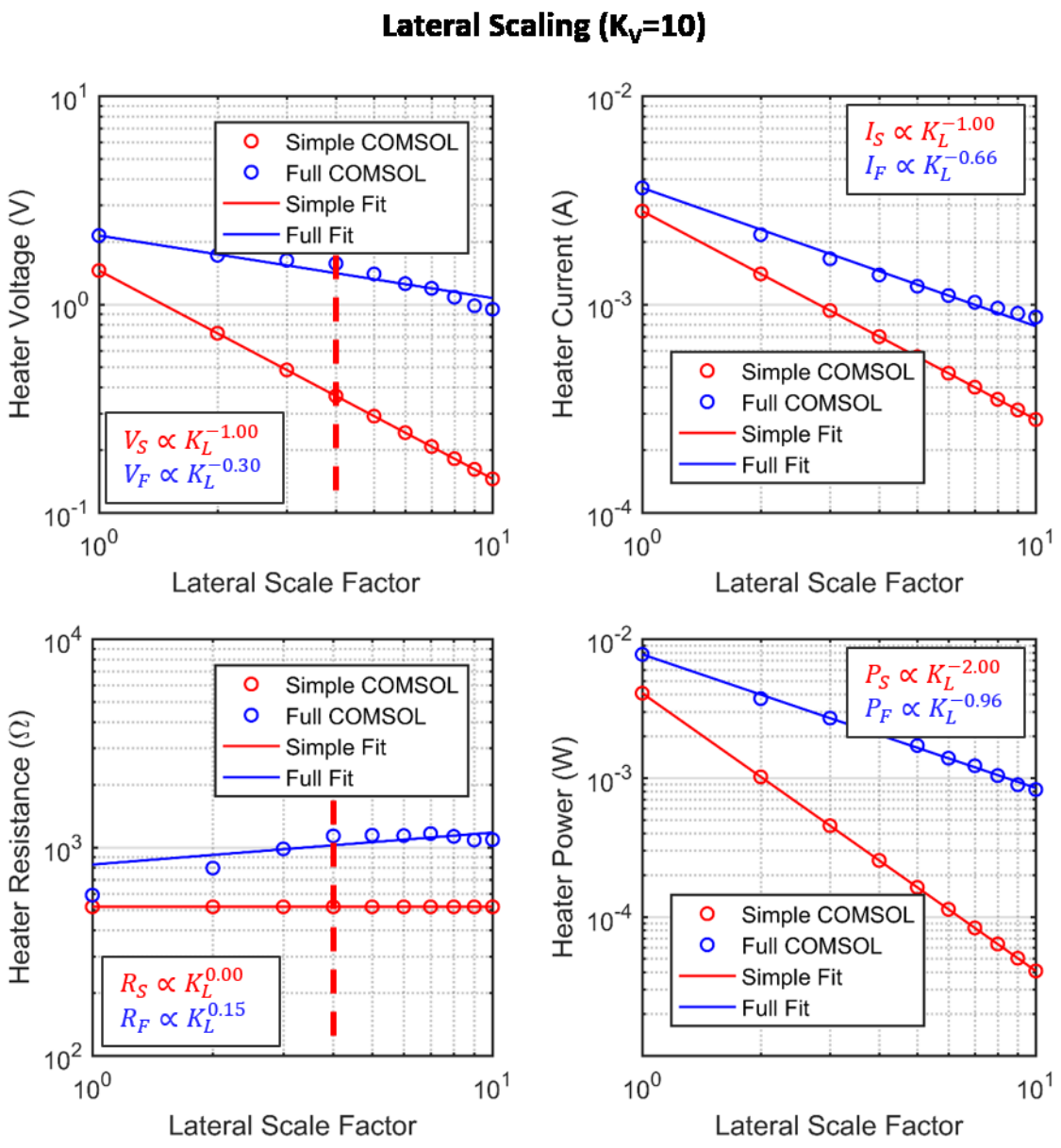

Figure 35. Lateral scaling trend extraction for simple and full COMSOL models. Voltage, current, resistance, and power are all measured at the heater when heating the surface of the PCM to just above the melting point of the PCM. Vertical red dashed lines indicate the boundary between two trends. Fits are for the overall scaling (not separated by the two trends).

Figure 36 compares the extracted heater temperature and PCM temperature scaling for the full COMSOL simulation. The two scaling trend regions are explained by a change in thermal conductance

from the PCM to the substrate and the PCM to the heater. For lateral scale factors below four, lateral thermal resistance of the PCM decreases relative to the vertical thermal conductivity of the heater. This transition at a lateral scale factor of four is not inherent to the PCNR design. Other geometry variations, like more or less overlap, may exhibit this behavior at a different scale factor. For lateral scale factors above four, thermal boundary resistance between the PCM and substrate becomes dominant, limiting the difference in temperature between the heater and PCM. Lateral thermal resistance per unit length in the PCM scales as:

$$R_{lateral} \propto \frac{1}{K_L} \tag{33}$$

since the PCM thickness remains constant. Lateral thermal resistance per unit length in the PCM decreases, while the TBR per unit length between the PCM and the substrate increases:

$$R_{TBR} \propto K_L \tag{34}$$

This comparative change in thermal resistance is faster than that seen in the isotropic scaling. Once the aspect ratio of the PCM results in a tall fin, the PCM exhibits a uniform temperature as thermal resistance in the PCM is significantly less than the thermal boundary resistance between the PCM and the heater or substrate.

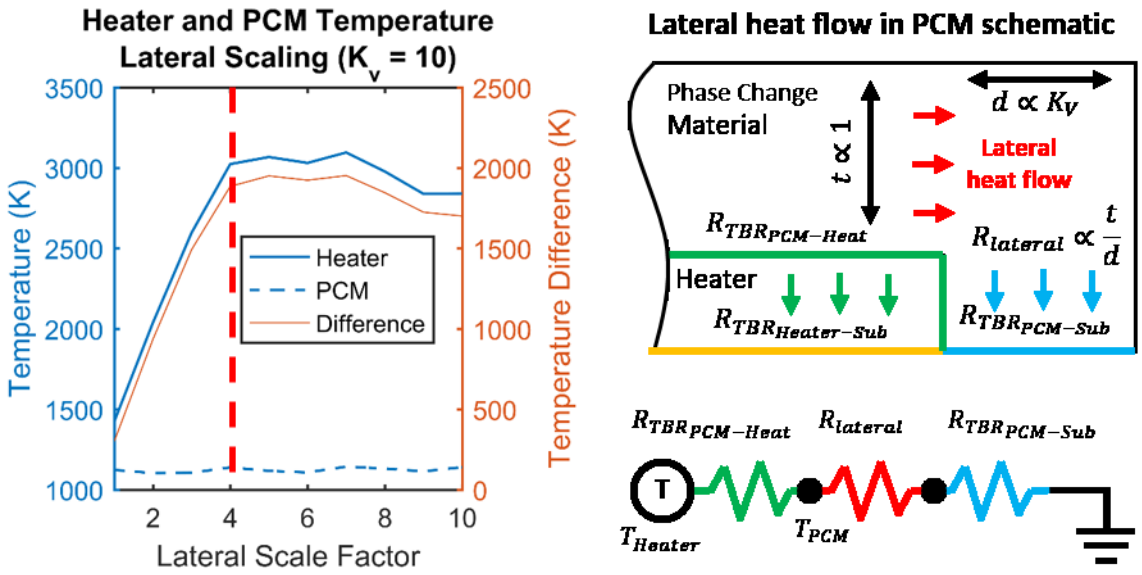

Figure 36. Heater and PCM temperature lateral scaling. Colored arrows represent heat flow through the corresponding resistance. Deviation in heater and PCM temperature is explained by lateral heat flow (red arrows) in the 2D schematic of an actuator. The TBR between the PCM and substrate (blue) becomes dominant past a scale factor of 4.

Quench time is scaled following the same isotropic and lateral scale factors. Figure 37 compares the simple and full COMSOL models for quench time scaling. Quench time is defined as the time for the maximum GeTe temperature to drop from 1000 K to 500 K. The quench time does not scale with lateral scaling as the layer thickness dominates the effect. Likewise both simple and full COMSOL models exhibit similar isotropic scaling of quench time as lateral effects play a minimal role in the quenching of the PCM.

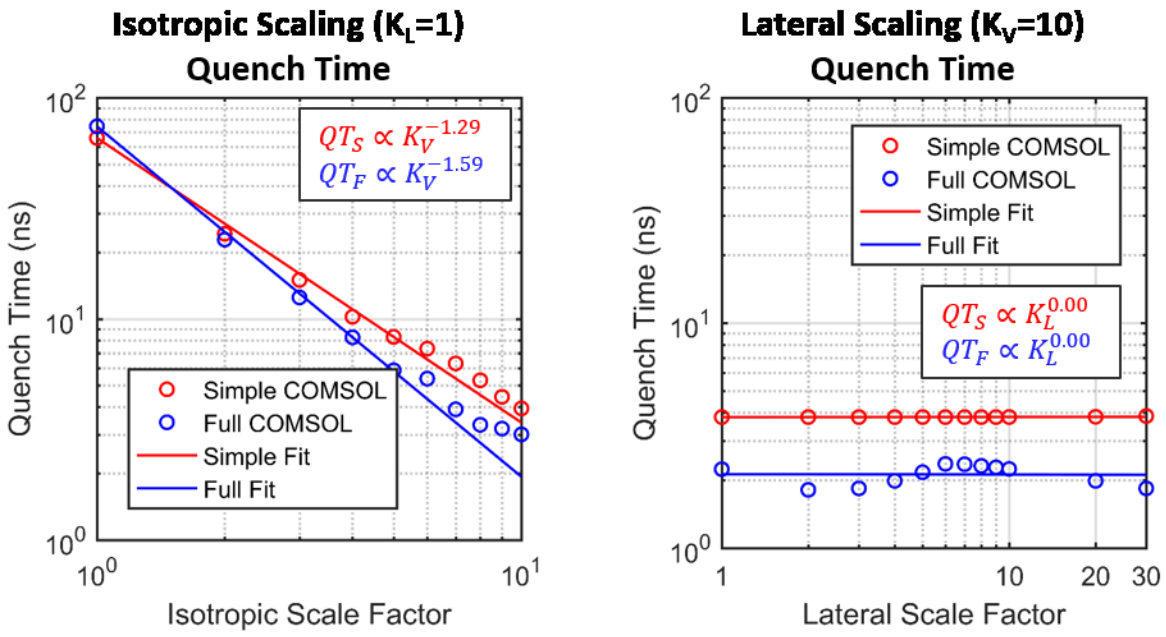

Figure 37. Quench time scaling trend extraction for simple and full COMSOL models. Quench time is defined as the time to drop from the melting temperature to the crystallization temperature.

Actuation energy is calculated by integrating the power to the heater over the time required to melt the PCM through the thickness of the actuator. Figure 38 shows the isotropic and lateral scaling of the actuation energy. Isotropic scaling of the simple and full models exhibit the same actuation energy scaling, however the full model strongly deviates for lateral scaling. This can be attributed to the changing aspect ratio of the device explained in the differing actuation power scaling analysis.

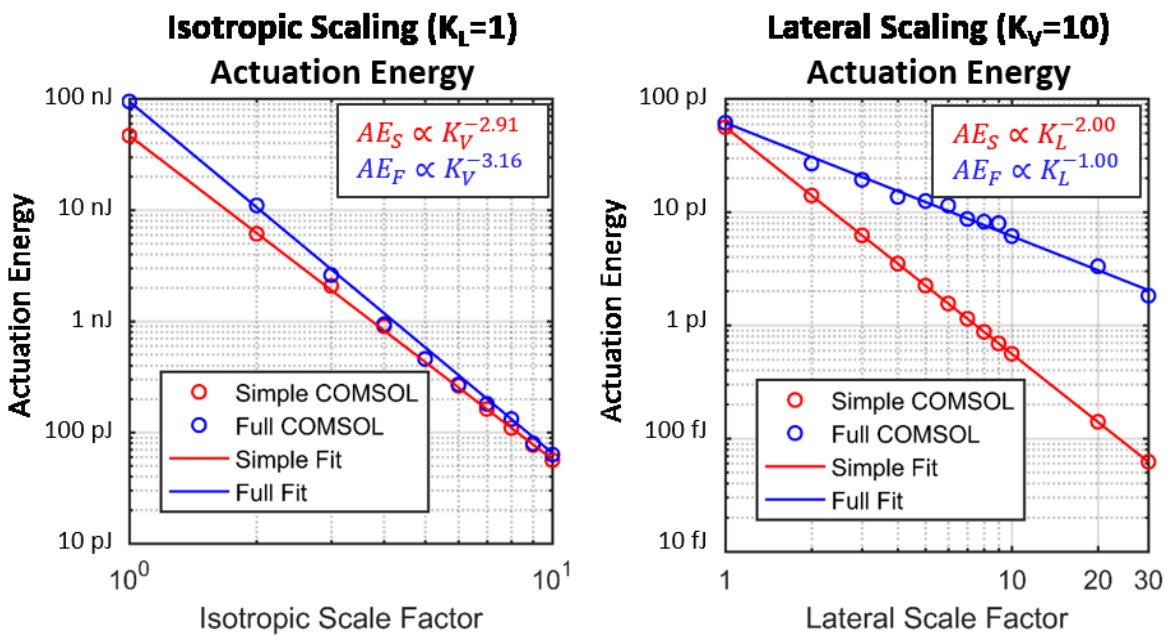

Figure 38. Isotropic and lateral scaling of actuation energy. Actuation energy is defined as the electrical energy dissipated in the heater while melting the PCM through the thickness of the actuator.

Table 2 compares the scaling trends for the 1D analytical, simple COMSOL, and full COMSOL models for isotropic and lateral scaling. The 1D analytical model matches the simple COMSOL model, but not the full COMSOL model. For this particular geometry a full 3D model is required to obtain accurate predictions of device performance as lateral effects are significant and change throughout scaling of the device. Moving to an alternate geometry may improve device performance when highly scaled, but it makes little difference at the scales proposed for initial fabrication and proof-of-concept. A future design should limit overlap of the PCM with the substrate to maximize heat transfer from the heater to the PCM.

Table 2. Extracted actuation scaling trends for PCNR devices.

PCNR Actuation Scaling Trend Comparison

	1D Analytical Isotropic Scaling	Simple COMSOL Isotropic Scaling	Full COMSOL Isotropic Scaling	1D Analytical Lateral Scaling	Simple COMSOL Lateral Scaling	Full COMSOL Lateral Scaling (K_V=10)
Heater V	$K_V^{-\frac{1}{2}}$	$K_V^{-0.50}$	$K_V^{-0.28}$	K_L^{-1}	$K_L^{-1.00}$	$K_L^{-0.30}$
Heater I	$K_V^{-\frac{3}{2}}$	$K_V^{-1.50}$	$K_V^{-1.33}$	K_L^{-1}	$K_L^{-1.00}$	$K_L^{-0.66}$
Heater P	K_V^{-2}	$K_V^{-2.00}$	$K_V^{-1.62}$	K_L^{-2}	$K_L^{-2.00}$	$K_L^{-0.96}$
Heater R	K_V	$K_V^{1.00}$	$K_V^{1.09}$	1	$K_L^{0.00}$	$K_L^{-0.15}$
Quench T	K_V^{-1}	$K_V^{-1.29}$	$K_V^{-1.59}$	1	$K_L^{0.00}$	$K_L^{0.00}$
Actuation E	K_V^{-3}	$K_V^{-2.91}$	$K_V^{-3.16}$	K_L^{-2}	$K_L^{-2.00}$	$K_L^{-1.00}$

Table 3 summarizes the scaling of the PCNR through isotropic and lateral scaling to a highly scaled PCNR. The highly scaled PCNR geometry (5 nm by 20 nm heater) can be compared to the previously analyzed cantilever and CMOS. Figure 39 compares the area and actuation voltage scaling of each device. The PCNR maintains significantly lower actuation voltages compared to the electrostatic relay. As the PCNR is scaled to fabrication limits, both area and voltage are able to compete with current CMOS.

Table 3. Predicted PCNR performance at different levels of scaling.

PCNR Scaled Geometry Performance

	Not Scaled (K_V=1 K_L=1)	Partially Scaled (K_V=10 K_L=1)	Highly Scaled (K_V=10 K_L=30)
Heater t x L x W	50 nm x 6 µm x 1.5 µm	5 nm x 600 nm x 150 nm	5 nm x 20 nm x 5 nm
PCM t x overlap	200 nm x 2 µm	20 nm x 200 nm	20 nm x 6.7 nm
Cap t	30 nm	3 nm	3 nm
Air Gap	20 nm	2 nm	2 nm
Heater V	4.1 V	2.1 V	0.52 V
Heater I	80 mW	3.6 mA	0.61 mA
Heater P	330 mW	7.8 mW	0.32 mW
Heater R	51 Ω	590 Ω	840 Ω
Quench T	74 ns	2.2 ns	1.8 ns
Actuation E	94 nj	61 pj	1.8 pj

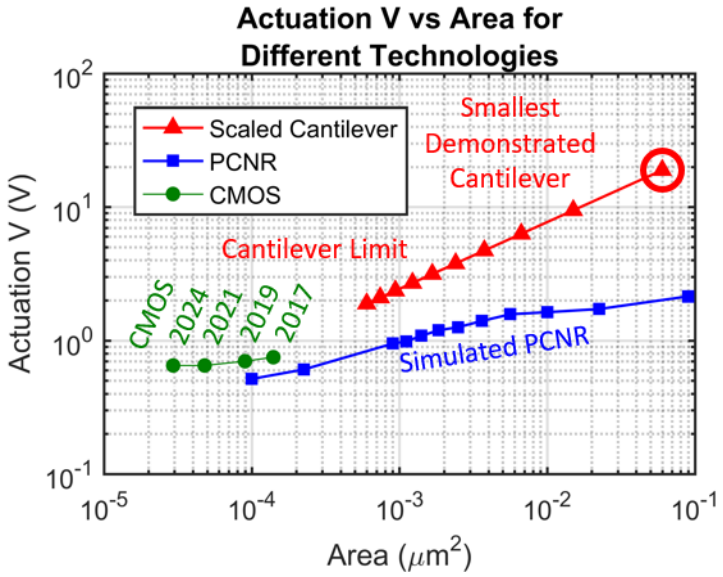

Figure 39. Comparison of actuation voltage and area between the highly scaled PCNR, the scaled electrostatic cantilever MEMS relay, and state of the art CMOS. CMOS data is pulled from [1] *and* [2].

2.6 Summary

In this chapter the phase change NEMS relay (PCNR) is introduced as an alternative non-volatile relay that addresses the limited scalability of traditional mechanical relay designs while maintaining the high on-off ratio of mechanical contacts. The operating concept and desired phase change material properties are described. Overall scaling predictions are made based on the analytical and COMSOL FEA device models for the device parameters: Actuation voltage, actuation current, heater resistance, actuation power, quenching time, and actuation energy. While scaling analysis shows lateral heat flow due to overlap of the phase change material can be a detriment to highly scaled devices, the fabricated proof of concept devices are minimally effected. These design considerations should be taken into account when scaling the device and the PCNR geometry properly modified to ensure maximum energy dissipation.

3 PCNR Fabrication

3.1 Abstract

A single process flow (Figure 40) is developed to fabricate PCNR devices. A partial fabrication yields phase change actuators, which are used to calibrate expansion and set the sacrificial material thickness (air-gap). Processes were characterized for depositing, patterning, and protecting the GeTe, depositing and patterning the drain and source, and releasing the air-gap.

GeTe processes are characterized to create smooth films with good adhesion. Ge and Te sputtering rates are characterized to form a 50-50 ratio of Ge to Te when co-sputtering GeTe. Proper GeTe composition is crucial for smooth films with good adhesion. Additionally, the quality of the seed layer (ALD Al_2O_3) strongly influences the subsequently deposited GeTe film roughness. Pinholes in the Al_2O_3 can cause different grain growth in the GeTe, leading to high roughness. GeTe must also be deposited in the

crystalline state, as GeTe deposited in the amorphous state and later crystalized exhibits poor adhesion and cracks at steps.

GeTe must be protected from the vapor HF release, so an ALD Al_2O_3 cap is used to encapsulate the GeTe. The cap must be pin-hole free, requiring the Al_2O_3 to be deposited at an elevated temperature. Additionally, sidewall roughness of the GeTe must be minimized to ensure the effectiveness of the cap. An argon physical etch was developed to pattern the GeTe, as the addition of reactive gas like chlorine causes rough, sponge-like side walls.

Etching the drain and source was found to be infeasible due to significant lateral etching around steps, which causes breaks in the drain and source. Lift-off using negative photoresist was used instead of etching. Undercut of the negative photoresist sidewall helps improve pattern resolution, but too much undercut can lead to fences of deposited material that does not lift-off. Release uses vapor HF to remove the sacrificial SiO_2 and form the air-gap. Etch rates are adequate to remove all SiO_2 under the drain and source.

3.2 Fabrication flow

A single fabrication flow is used to create actuators and relays, where actuators are a partially complete fabrication of the final relay. Initially designed layer thicknesses needed to be changed to account for fabrication difficulties like pinhole defects. This section details the development of the fabrication process of both the PCNR actuator and the full PCNR device.

Fabrication was completed using processes previously developed for GeTe based RF switches that harnessed the change in resistivity of GeTe between the crystalline and amorphous states. Device geometry is limited to the minimum feature size of the ASML5500/80 i-line stepper used in fabrication (0.5 µm line width). The eight step process that is used is shown in Figure 40. Devices are fabricated on a 4” Si substrate. First, alignment marks are etched into the substrate followed by the deposition of a

100 nm AlN isolation layer. A 50 nm layer of W is then sputtered on the isolation layer with the substrate at an elevated temperature of 800 °C. This elevated temperature is used to ensure lower resistance of the heater layer. The heater is then patterned by an SF_6 reactive ion etch. Heater widths range from 0.5 µm to 1.5 µm and heater lengths range from 3 µm to 11 µm. Second, a 30 nm layer of Al_2O_3 is deposited by atomic layer deposition (ALD) at 250 °C. ALD is required to ensure a defect free conformal coating for later GeTe deposition. Third, 200 nm of crystalline GeTe is deposited by co-sputtering at 400 °C, and is then patterned by an Ar plasma etch. Fourth, another 30 nm layer of Al_2O_3 is deposited by ALD to encapsulate the PCM. This cap layer protects the PCM from the release etch, contains the PCM while melted, and electrically isolates the metallic contact from the actuator. Actuators are completed by the end of step four and can be tested to determine actuation voltages and actuator expansion. Wafers used for actuator testing are not used to create complete devices.

In the fifth step, 30 nm of W is sputtered and patterned by SF_6 reactive ion etch to form the metallic contact. The W is not deposited at elevated temperature as this would melt the GeTe PCM. Sixth, 20 nm of SiO_2 is deposited ty ALD and patterned by a CHF_3 and O_2 reactive ion etch to form the sacrificial layer. Additionally before drain and source deposition, vias are etched through the Al_2O_3 to expose the heater pads with buffered hydrofluoric acid (BFH) and a hard-baked photoresist mask. Seventh, 1 µm of W is sputtered on a 20 nm Cr adhesion layer and patterned by lift-off to form the drain and source contacts as well as probing pads. The drain and source are thick relatively to the rest of the structure to minimize the effects of stress on the air gap after release. Finally, the device is released by etching the sacrificial SiO_2 in vapor HF.

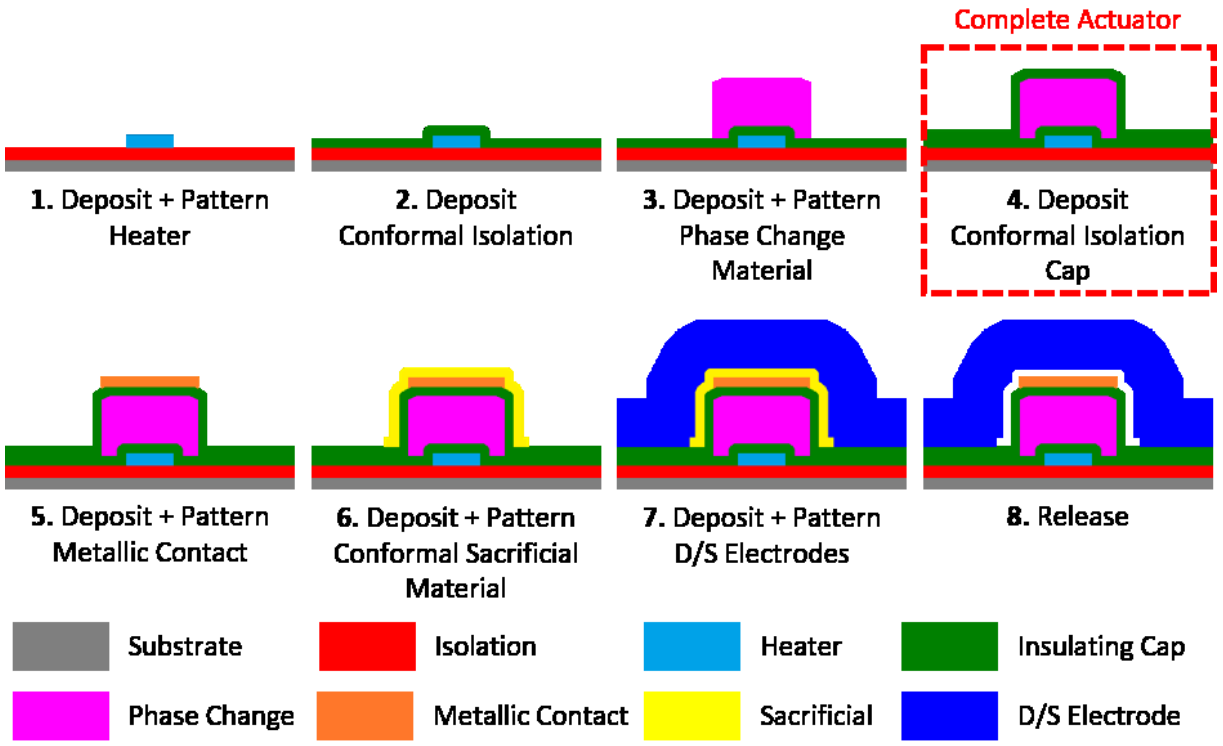

Figure 40. Fabrication flow for the Phase Change NEMS Relay. Fabrication of just actuators is completed after step four.

3.3 GeTe Processing

A number of different factors influence the quality of the GeTe layer and how well it performs as an actuator ranging from sputtering rates to seed layer quality. A good film is relatively smooth, exhibits good adhesion, and can be patterned with smooth sidewalls.

3.3.1 GeTe Composition

GeTe is co-sputtered to form a 50-50 ratio of Ge to Te. While composition was not thoroughly studied to see its effect on expansion, binary phase diagrams (Figure 41) [91] indicate a narrow region where the GeTe congruently melts. GeTe composition is set by measuring the relative deposition rates of the Ge and Te targets and setting the appropriate sputtering powers to produce 50-50 GeTe. The composition relationship is:

$$t_{Ge}\left(\frac{Density_{Ge}}{MolecularWeight_{Ge}}\right) = t_{GeTe}\left(\frac{Density_{GeTe}}{MolecularWeight_{GeTe}}\right) \quad (35)$$

Figure 42 shows the effects of a Te rich GeTe deposition where the ratio is 40-60 Ge-Te. The deposited film is extremely rough and exhibits poor adhesion. Roughness of the Te rich films makes measuring deposition rate difficult and can lead to inaccurate calibrations of the relative sputtering rates.

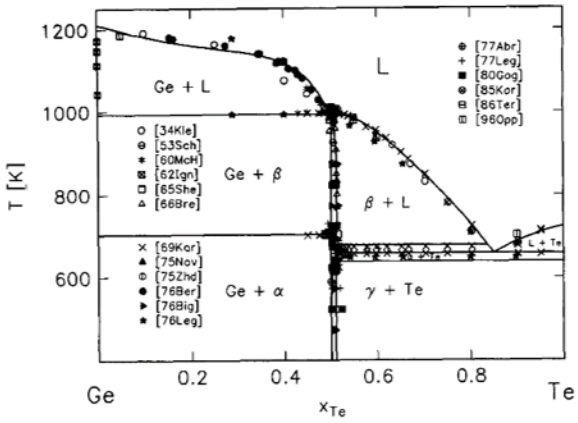

Figure 41: Ge-Te phase diagram [91]. *Ge and Te are specified by atomic ratio, not by weight.*

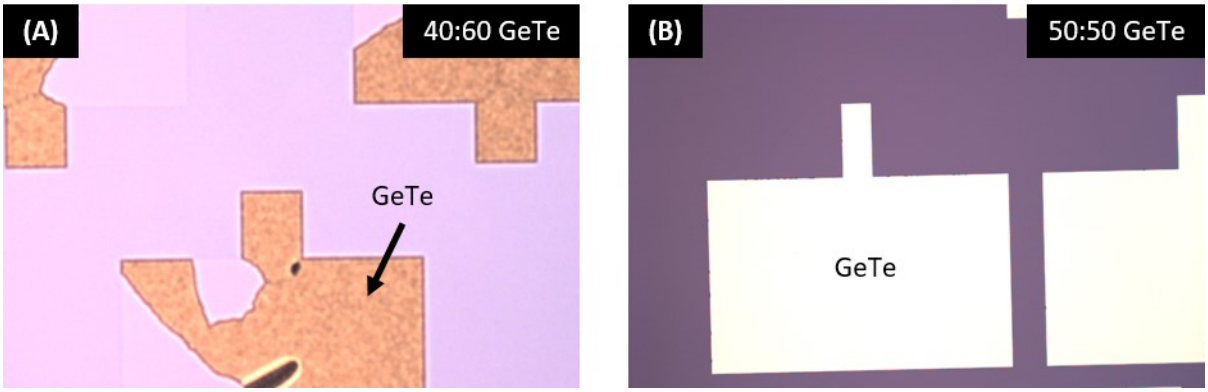

Figure 42. Optical microscope images comparing adhesion and roughness of GeTe with (a) Te rich 40:60 GeTe and (b) stoichiometric 50:50 GeTe.

3.3.2 GeTe Seed Layer

The isolation layer between the GeTe and the heater can greatly influence the quality of the GeTe grown. Figure 43 shows the effect of pinholes in the Al_2O_3 isolation layer. Large bumps in the GeTe layer grow at the location of pinholes in the isolation layer. 10 nm of Al_2O_3 deposited on a blank Si wafer has no defects, while the same deposition process on AlN and W creates a high density of defects. Increasing the thickness of the Al_2O_3 to 30 nm creates a pinhole free layer, which creates smooth GeTe. Additionally using a deposition temperature of 250 °C improves deposition time and reduces pinhole defects in the film as shown in the cap layer development tests. Figure 44 shows the difference between GeTe patterned on the original 10 nm 150 °C ALD Al_2O_3 seed layer and the improved 30 nm 250 °C ALD Al_2O_3 seed layer.

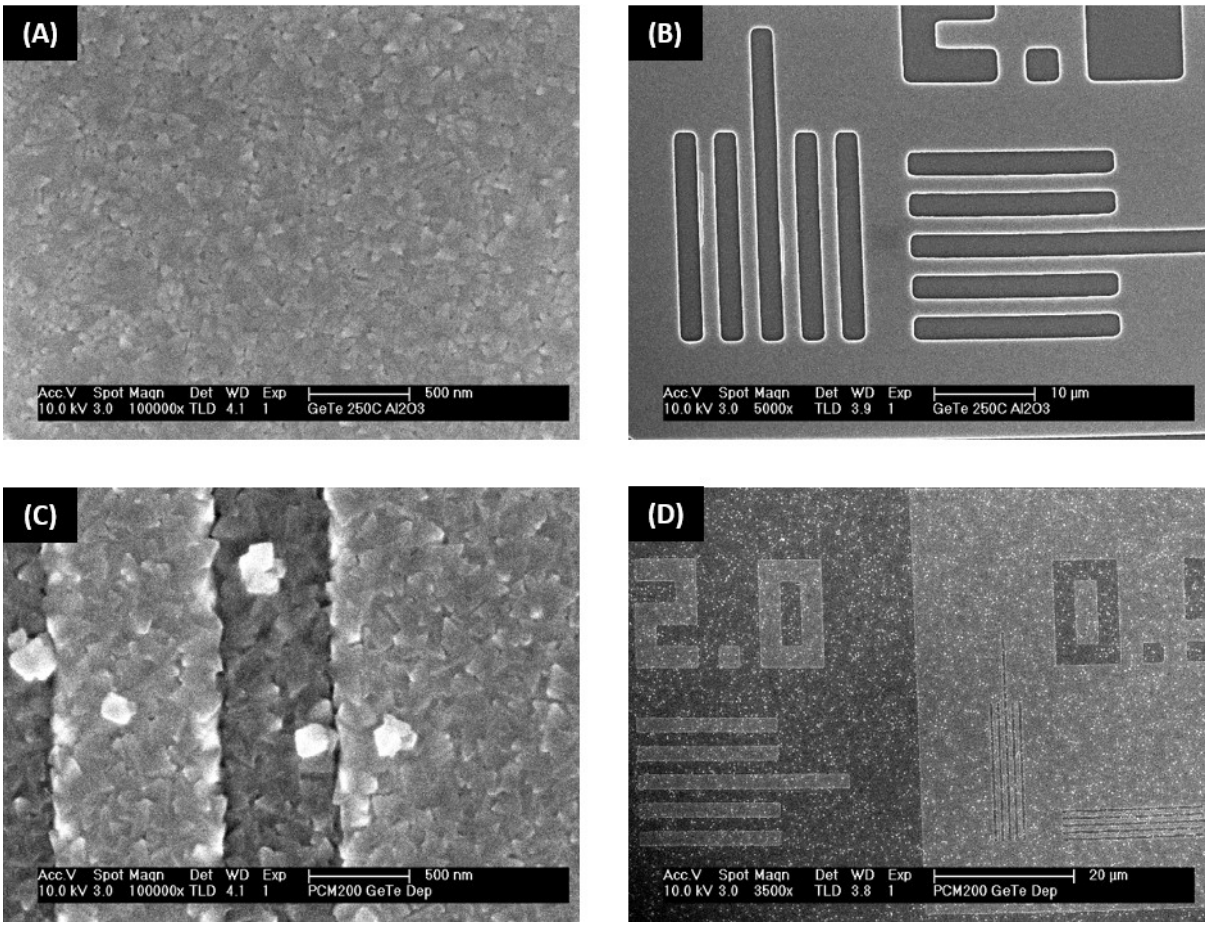

Figure 43. GeTe roughness comparison for crystalline GeTe deposited on (a,b) a test Si substrate with 10 nm Al_2O_3 deposited by ALD at 250 °C and (c,d) device substrates with 10 nm Al_2O_3 deposited by ALD at 150 °C.

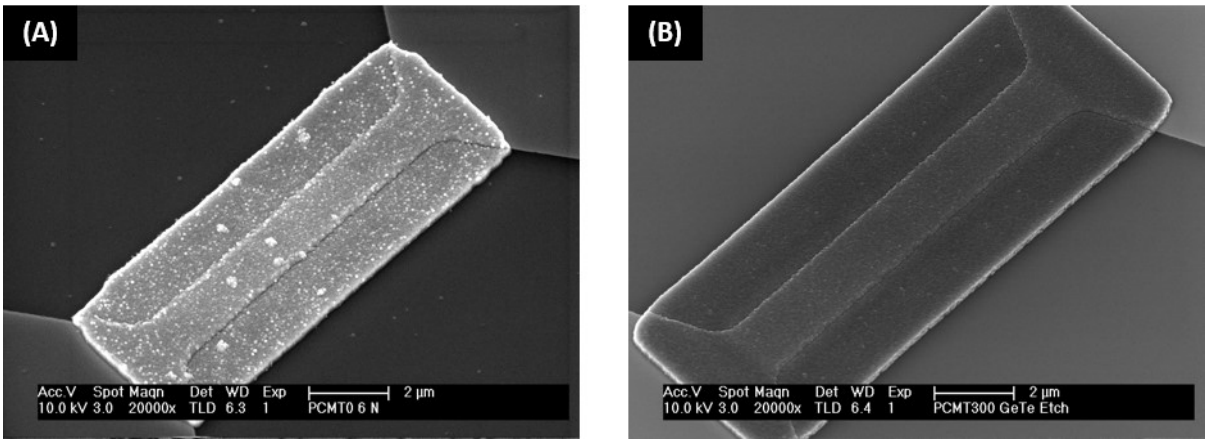

Figure 44. GeTe roughness comparison between devices fabricated with (a) 10 nm Al_2O_3 deposited by ALD at 150 °C and (b) 30 nm Al_2O_3 deposited by ALD at 250 °C.

3.3.3 Amorphous and Crystalline GeTe

Devices must be fabricated with crystalline GeTe for the actuator to expand and close the air gap when actuated. GeTe can be deposited in either the amorphous or crystalline states depending on the substrate temperature. GeTe deposited as amorphous can be crystalized after deposition by holding the film at elevated temperature – similar to the crystallization process used in the function of the relay. Figure 45 shows a patterned section of GeTe before and after conversion to the crystalline state. SEM images show the amorphous film is very smooth even after crystallization, offering an alternate solution to the thicker seed layer. All three (optical, electrical, and mechanical) of the property changes can be seen in the transition to the crystalline state. Microscope images show the section of GeTe shifting from opaque to reflective. Resistance measurements show the film resistivity changes from 5.4 Ωm to $9.1*10^{-6}$ Ωm (almost 6 orders of magnitude) when converting from amorphous to crystalline. Additionally SEM images show a crack formed in the GeTe where it covers the heater step. While this crack will be filled up during the capping process, it indicates a large tensile stress in the film that can cause delamination of the actuator. Figure 46 shows the effect of the large tensile stress. Actuators tend to peel off the substrate when used, causing poor device yield.

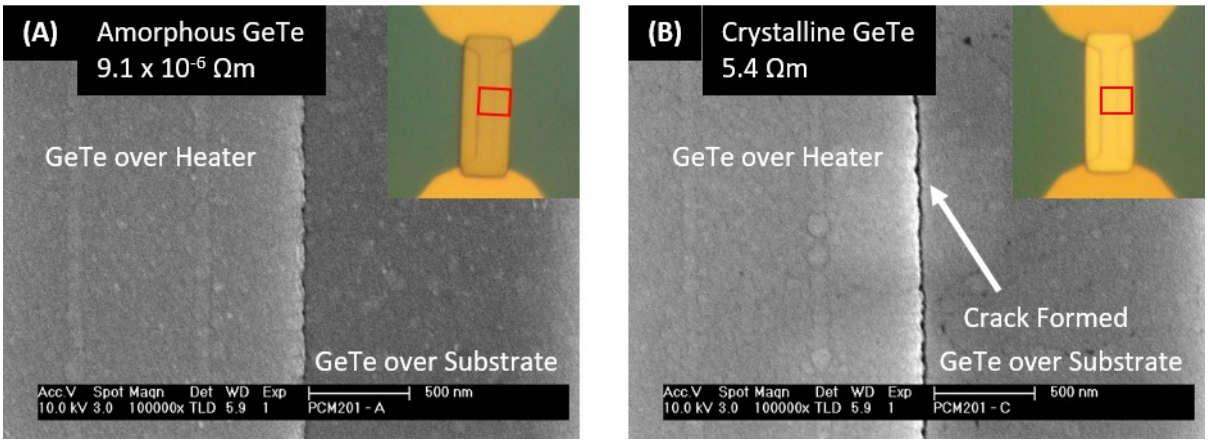

Figure 45. Conversion of patterned GeTe from (a) the amorphous state to (b) the crystalline state. The GeTe resistivity decreases by almost six orders of magnitude. Optical Microscope images show GeTe changes from opaque to reflective. SEM images show a crack forms at a step as GeTe contracts.

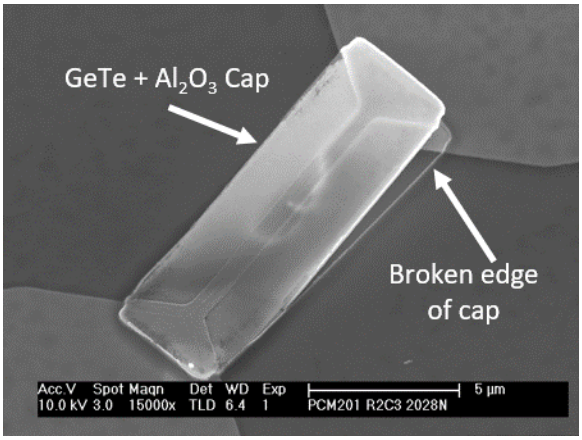

Figure 46. Actuator with delamination of the GeTe layer after actuation. GeTe was deposited as amorphous and converted to crystalline after patterning.

3.3.4 Etching GeTe

Depositing GeTe in the crystalline state prevents patterning by liftoff due to the high deposition temperature (400 °C). Patterning of GeTe with chlorine gas mixtures has previously been studied [92]. Due to the hazardous nature of the tellurium etching had to be performed in a Plasmatherm Versaline ICP RIE system. Etches with different compositions of chlorine and argon were tried to determine the best

etch recipe. Figure 47 compares the sidewall roughness for devices etched with and without chlorine present. While the chlorine greatly increases the etch rate, it creates a “sponge-like” sidewall. This extreme roughness would not be compatible with subsequent cap layers. An argon physical etch leaves a smooth sidewall, and is relatively selective to the Al_2O_3 seed layer.

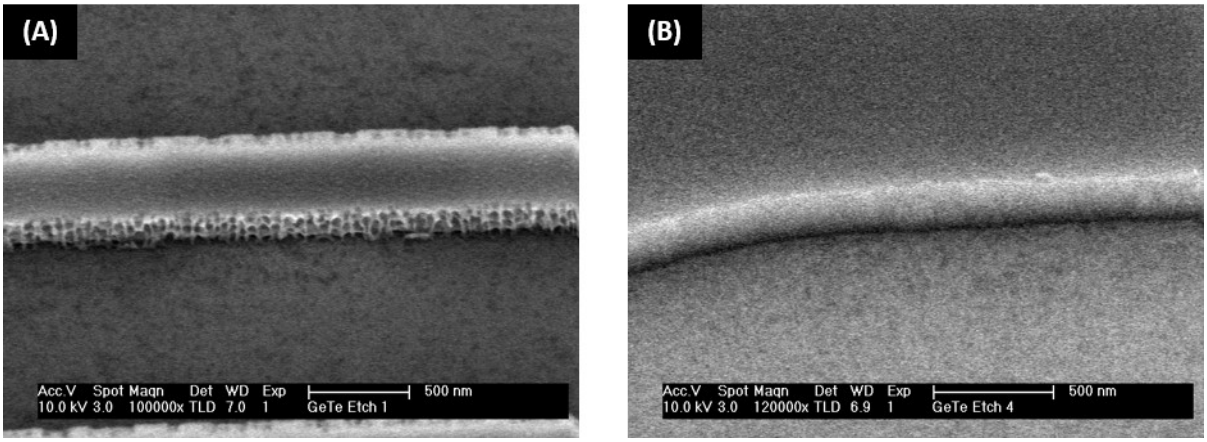

Figure 47. Sidewall comparison of GeTe patterned by RIE with (a) a mixture of Chlorine and Argon and (b) pure Argon.

3.4 Cap Layer

The cap layer has three purposes: prevent reflow of the PCM when melted, electrically isolate the metallic contact from the PCM, and protect the PCM from the release etch. Al_2O_3 is used as the cap because it can be conformally deposited by ALD. Different deposition temperatures and layer thicknesses were tested for survival in the vapor HF release etch and are shown in Figure 48. A 30 nm layer of Al_2O_3 deposited at 250 °C was found to be the thinnest layer that prevents vapor HF damage to the GeTe. Additionally the 250 °C Al_2O_3 deposition yields smoother and more defect free films as seen in Figure 49.

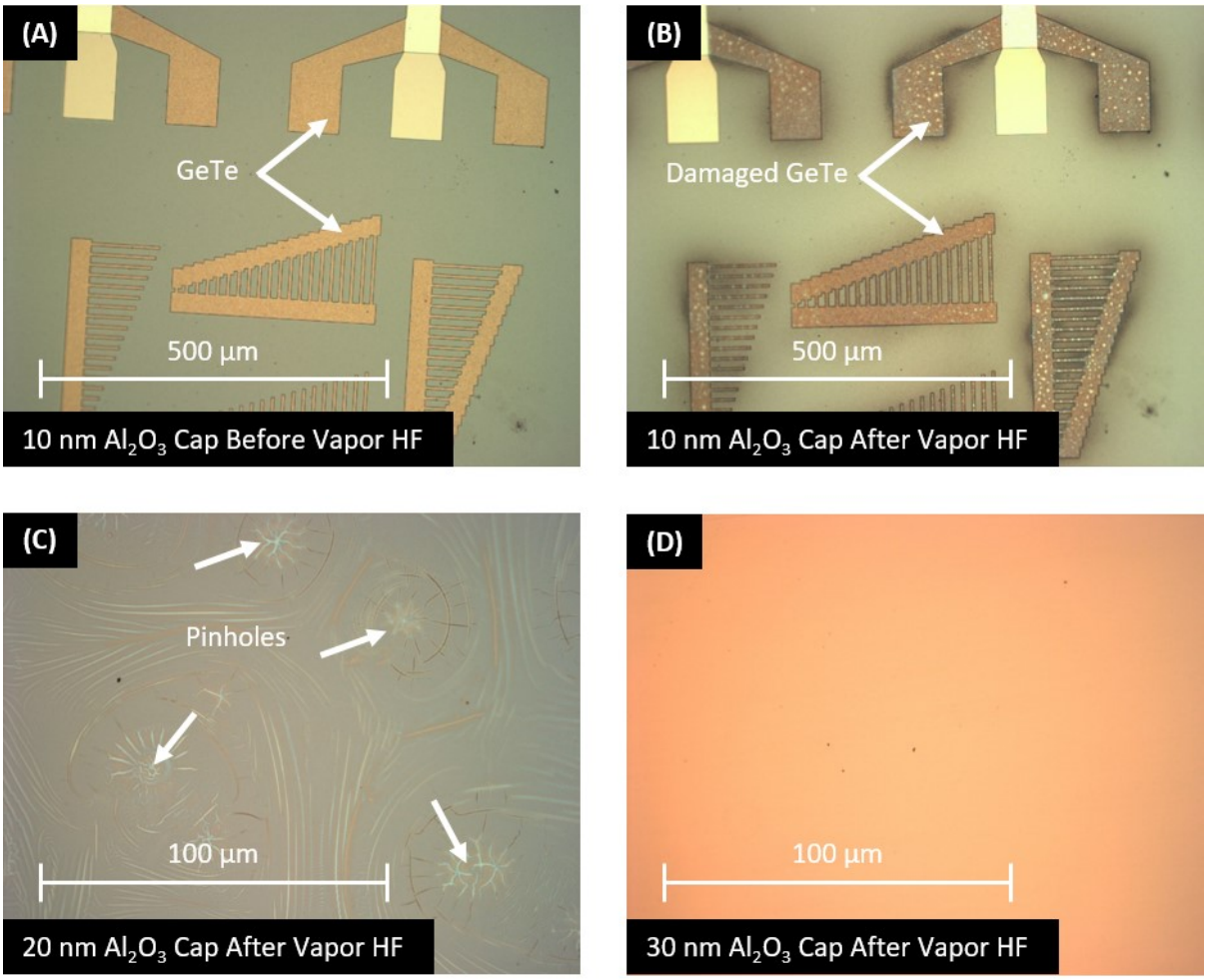

Figure 48. Optical microscope images of (a) devices fabricated with a 10 nm Al_2O_3 cap deposited by ALD at 150 °C before exposure to vapor HF, (b) devices with GeTe damaged by vapor HF, (c) a test layer of GeTe with a 20 nm Al_2O_3 cap deposited by ALD at 250 °C with GeTe damage from vapor HF stemming from pinhole defects in the cap, and (d) a test layer of GeTe perfectly protected by a 30 nm Al_2O_3 cap deposited by ALD at 250 °C.

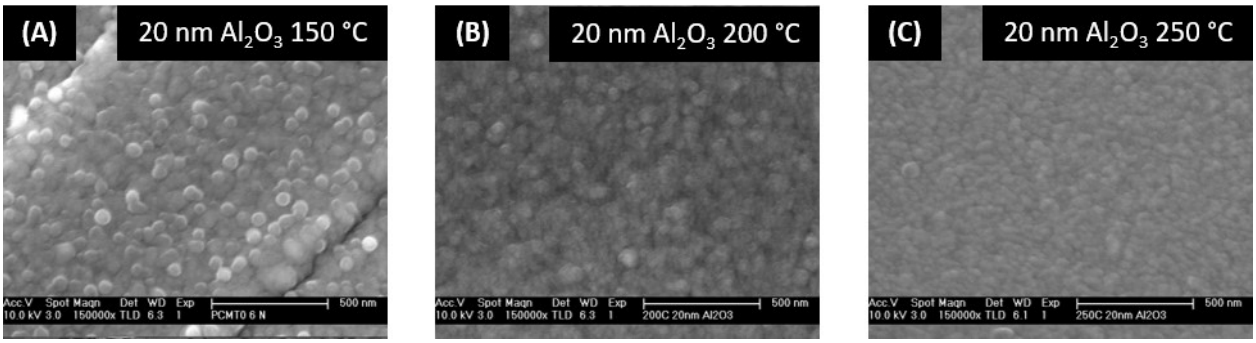

Figure 49. SEM images comparing roughness of a 20 nm film of Al_2O_3 deposited by ALD at temperatures of (a) 150 °C, (b) 200 °C, and (c) 250 °C.

3.5 Drain and Source

Multiple materials and patterning methods were tried to create the drain and source. The drain and source needs to be thick enough to prevent deformation from residual stress gradient, necessitating a 1 µm thick film. Initially W films were deposited by sputtering and patterned by SF_6 reactive ion etch. Figure 50 shows a comparison between a test and a fabricated device with the same drain and source process. The test shows cleanly patterned features with highly vertical sidewalls, while the final device shows severe undercut where the drain and source cover the step from the actuator. Vertical sidewalls are caused by a lower etch rate at grain boundaries. Undercut is likely caused by differences in grain growth at the actuator edge step.

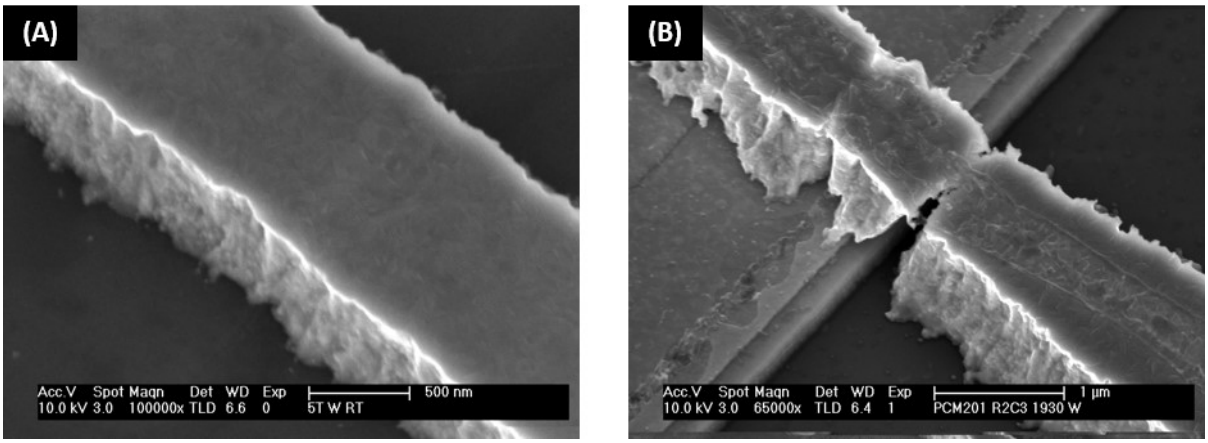

Figure 50. SEM images comparing wall roughness of patterned tungsten on (a) a test wafer with a step etched in the Si wafer, and (b) a fabricated device.

Liftoff using negative photoresist was next investigated to avoid the undercut created by etching the thick tungsten and allows for a wide range of materials to be explored for the drain and source. Unlike positive photoresist which can leave "fences" at the pattern edges, negative photoresist creates overhanging sidewalls that mask the pattern edges. Figure 51 shows the effects of this overhang on drain and source pattern quality. Too much undercut in the photoresist reduces pattern resolution and can lead to thin layers of material bridging sections of the pattern. Properly developed sidewalls will lead to sloped edges of the pattern. Additionally, narrow patterns relative to photoresist thickness can lead to reduced film thickness. Figure 52 shows 1 µm, 1.5 µm, and 2 µm wide drain and source patterns from the same deposition. The 1 µm wide drain and source is approximately half the thickness of the 2 µm wide drain and source, while the 1.5 µm wide drain and source is similar in thickness to the 2 µm wide drain and source.

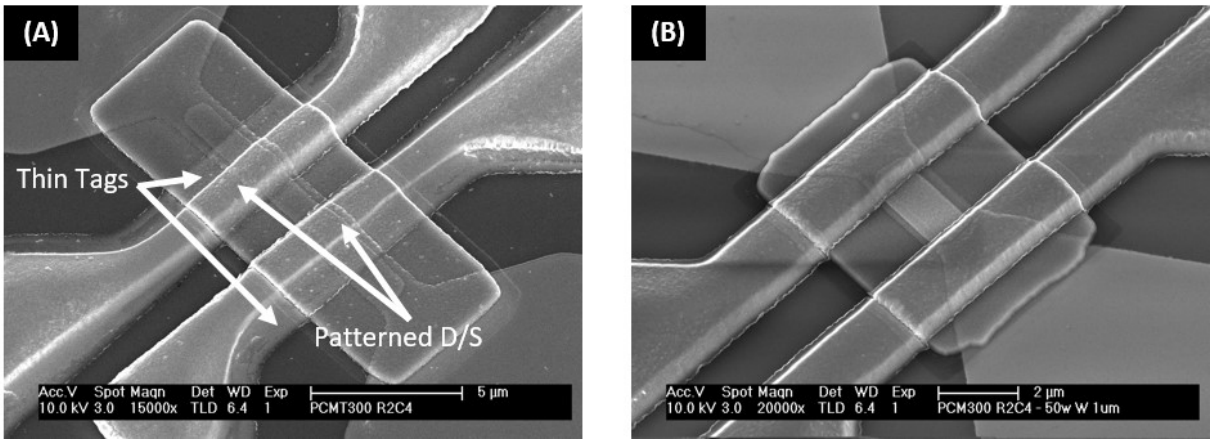

Figure 51. Comparison of drain and source liftoff with photoresist that has (a) too much undercut and (b) adequate undercut to prevent fences.

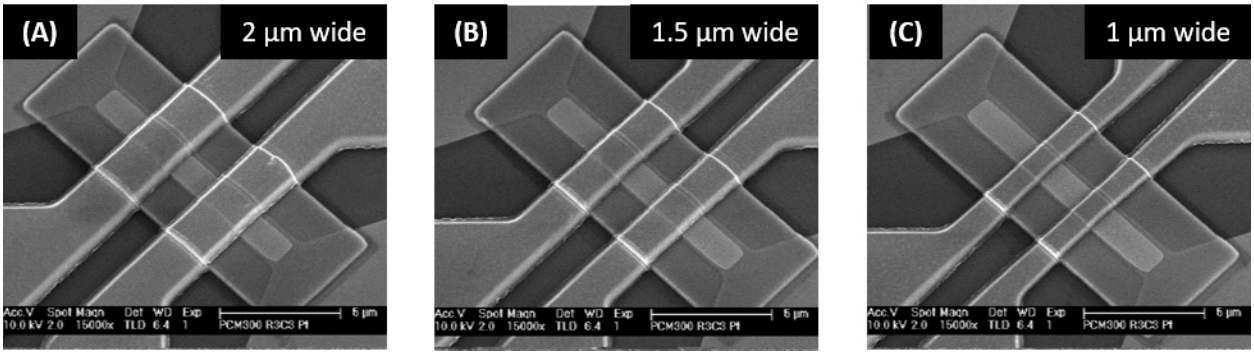

Figure 52. SEM images comparing the drain and source with widths of (a) 2 µm, (b) 1.5 µm, and (c) 1 µm.

While liftoff avoids the severe undercut seen from etching, it does not solve the grain growth issue. Platinum was also explored as an alternative to the tungsten drain and source, however the step coverage still exhibits similar defects. Figure 53 shows fabricated devices with tungsten and platinum drain and source layers. Step coverage defects on the 1 µm thick layers do not appear to negatively affect device performance, however future work will need to be done to eliminate the crack formed at the actuator step.

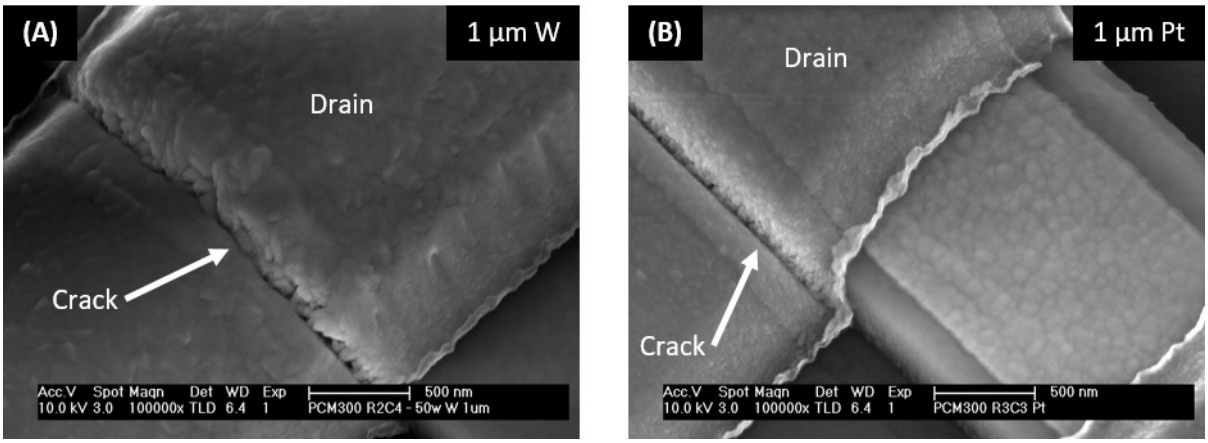

Figure 53. SEM images of step coverage of the drain and source fabricated from (a) 1 µm of tungsten and (b) 1 µm of platinum.

3.6 Release

The release etch uses vapor hydrofluoric acid (VHF) to remove the SiO_2 sacrificial material. SiO_2 and VHF were chosen as the sacrificial material because it can be readily deposited by ALD which ensures conformal coatings. VHF is able to selectively etch the SiO_2 and release narrow, high-aspect ratio gaps. Figure 54 shows before and after images of a release etch test where a 300 nm layer of W was deposited on 20 nm of ALD SiO_2. The etch front can be seen underneath the pads. A single 10 minute cycle is able to remove 15 µm of SiO_2, significantly more than the maximum 1 µm required etch distance.

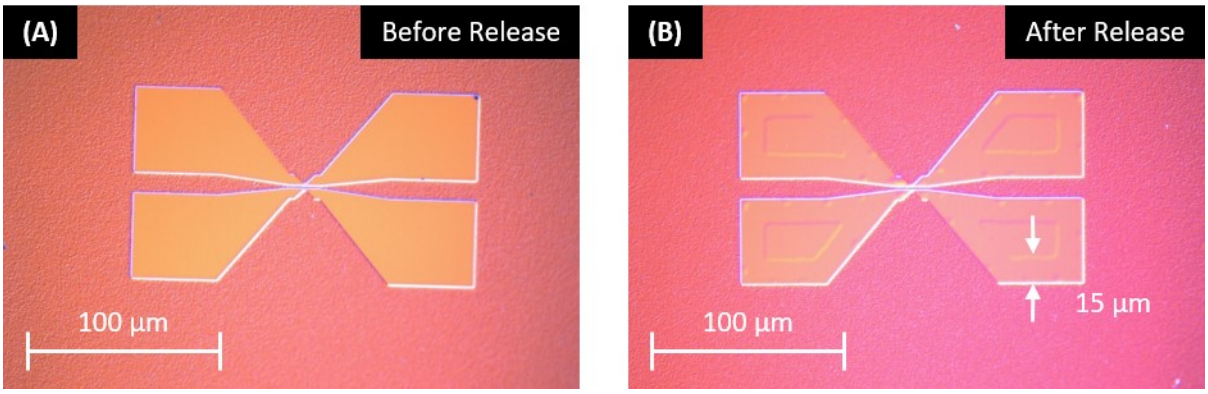

Figure 54. Optical microscope images of a tungsten drain and source test structure on a 20 nm SiO_2 film deposited by ALD (a) before and (b) after release with vapor HF.

3.7 Summary

In this chapter, the process flow developed to fabricate PCNR devices is described. A partial fabrication yields phase change actuators, which are used to calibrate expansion and set the sacrificial material thickness (air-gap). GeTe processes were characterized to create smooth films with good adhesion. GeTe requires: proper GeTe composition (50-50 ratio of Ge to Te), a pinhole free Al_2O_3 seed layer, and deposition in the crystalline state. Additionally, the Al_2O_3 cap must be pinhole free and the GeTe must have smooth sidewalls for the cap to protect the GeTe from damage during the vapor HF release etch. A lift-off process using negative photoresist was developed to pattern the drain and source to avoid significant lateral undercut around steps created when etching a very thick film. Vapor HF etch rates were measured for the 20 nm sacrificial SiO_2 used to form the narrow air-gap, and are adequate to remove all SiO_2 under the drain and source.

4 PCNR Actuator Testing

4.1 Abstract

The main goals for testing the fabricated actuators is to measure deformations and verify phase change occurs. Although not quantitatively compared with models, actuator testing was used to set the sacrificial layer (air-gap) thickness and determine the viability of the actuator fabrication process. Actuators tested were fabricated in an older process using thinner layers of Al_2O_3 for the isolation and cap layers. Of the three main property changes (electrical, optical, and mechanical), only volume change and appearance change are measured as the phase change material is fully encapsulated, preventing electrical contact with the phase change material. Actuator expansion was measured to be 13 % (26 nm) initially with 7 – 8 % (14 – 16 nm) in subsequent contractions and expansions.

4.2 Optical Characterization of Phase Change Actuation

Optical microscope images are used to verify the conversion of GeTe between crystalline and amorphous states. When switching between the crystalline and amorphous states, the phase change material changes from reflective to opaque. Optical microscope images of a single actuator in different actuation states are show in Figure 55. An initially fabricated device has a uniform appearance throughout the whole actuator as the GeTe is deposited in the crystalline state. After actuation, a section of the GeTe converts to amorphous and turns opaque. This is seen as a dark section appearing over the heater. As predicted from the COMSOL simulation, the heat is confined to a section of the GeTe just above the heater.

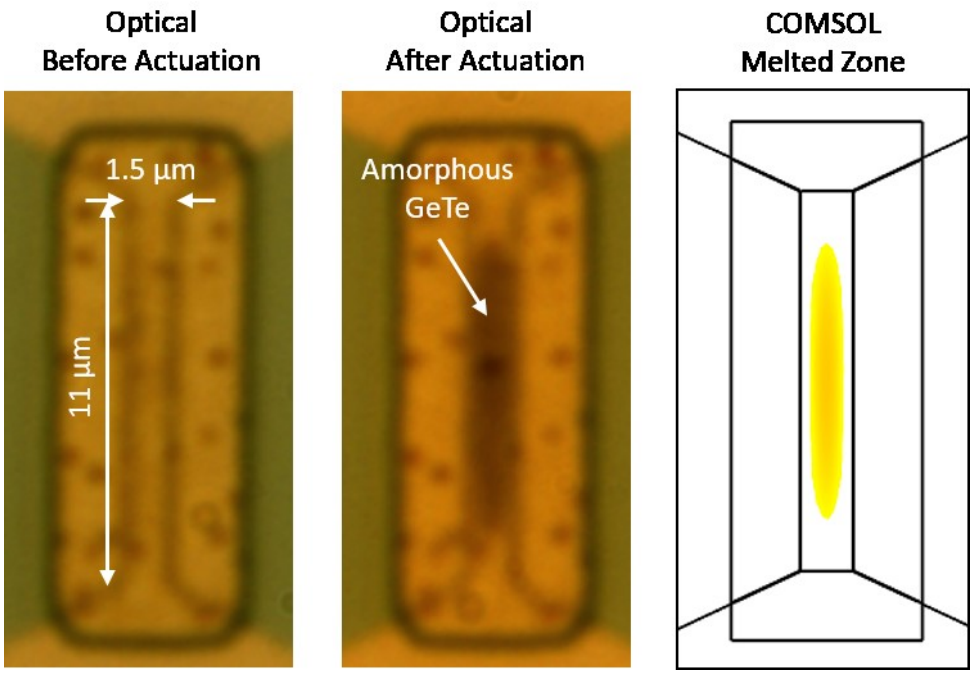

Figure 55. Optical microscope images highlighting the transformation of GeTe from crystalline to amorphous. The amorphous zone created matches the melted zone from the COMSOL model.

Additionally, optical images can be used to determine the operating space for pulse amplitude and time. Figure 56 shows a set of actuators with the same heater dimensions (1.5 µm x 9 µm) actuated with different combinations of pulse voltage and time. Overlap of the phase change material differs between the devices, but does not affect the melted zone as heat is confined to just above the heater. Actuations are characterized as one of four levels: no actuation, partially actuated, fully actuated, and damaged. Pulses that cause no significant change in color above the heater are defined as no actuation, pulses that cause a small section of phase change material to darken are defined as partially actuated, pulses that cause complete conversion of PCM above the heater are defined as fully actuated, and pulses that cause damage to the heater of PCM are defined as damaged. A tradeoff between actuation voltage and time can be seen in the band of fully actuated devices. Faster actuation is possible at the expense of higher amplitude pulses. Recrystallization pulses can be characterized using the same optical method as

the actuation pulses. Figure 57 shows the same devices before and after varied recrystallization pulses are applied. Figure 58 characterizes the effects of the applied recrystallization pulses. As expected, full recrystallization occurs at lower voltages than actuation. Figure 59 shows the combined actuation and recrystallization pulses. Distinct boundaries show the shift between no actuation, recrystallization pulses, actuation pulses, and damaging pulses. Although these pulses show the full range of effects, the range of tested amplitude and times is not exhaustive. Longer actuation times, especially for recrystallization may provide better reliability of the actuator as this ensures all amorphous PCM recrystallizes.

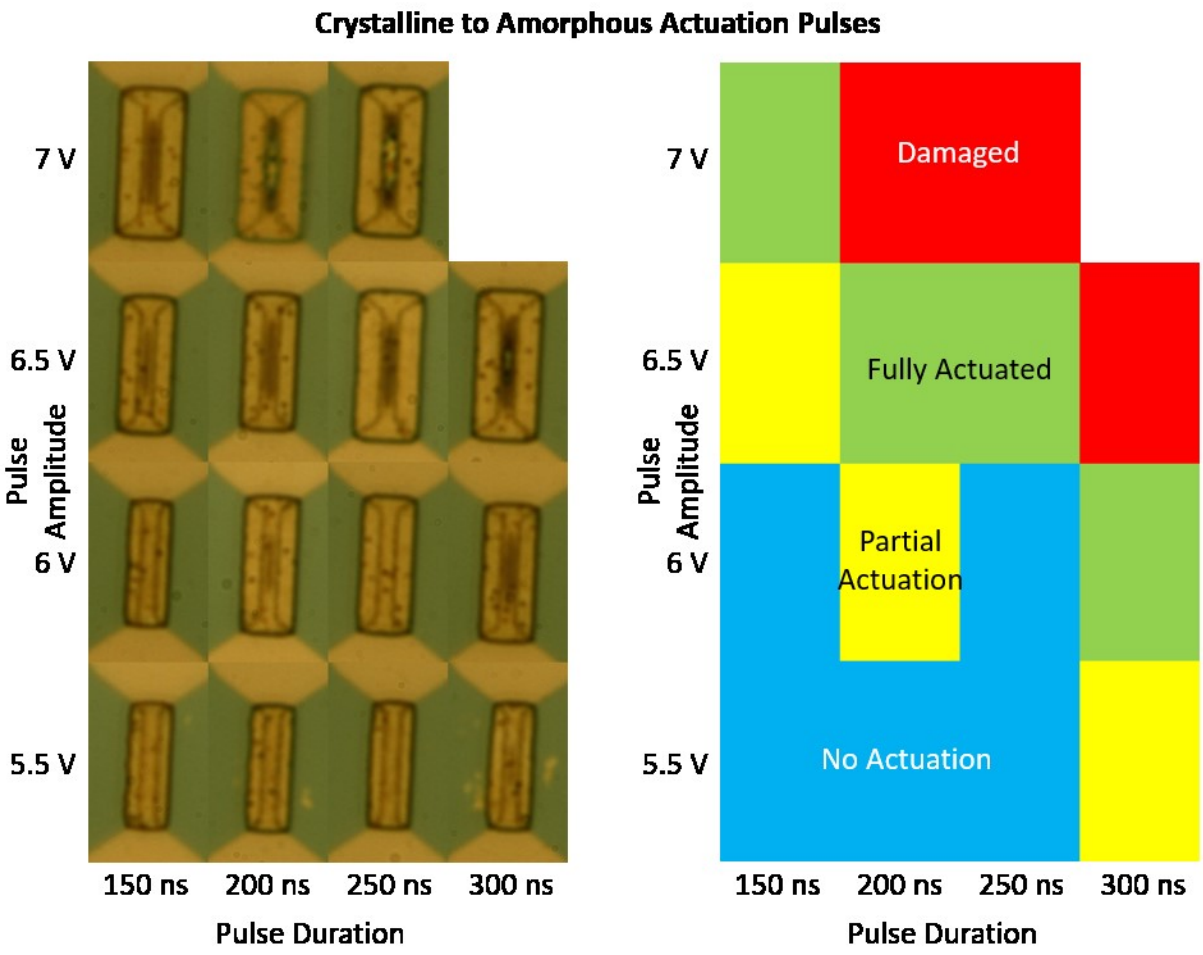

Figure 56. Optical microscope images of actuators after varied actuation pulses are applied. The effect of each pulse is characterized to have one of four effects: no actuation, partial actuation, fully actuated, or damaged.

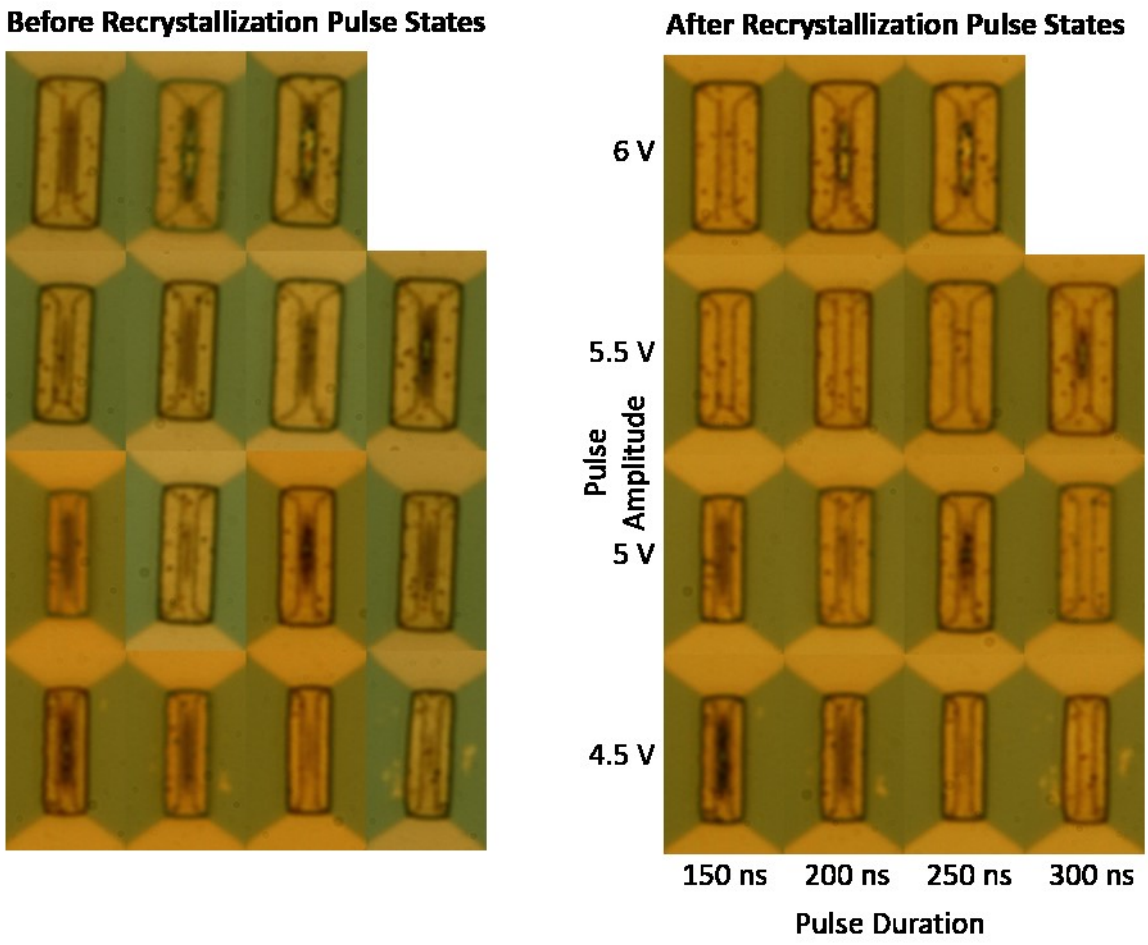

Figure 57. Optical microscope images of actuators before and after varied recrystallization pulses are applied. Recrystallization can be observed by the disappearance of dark areas in the actuator.

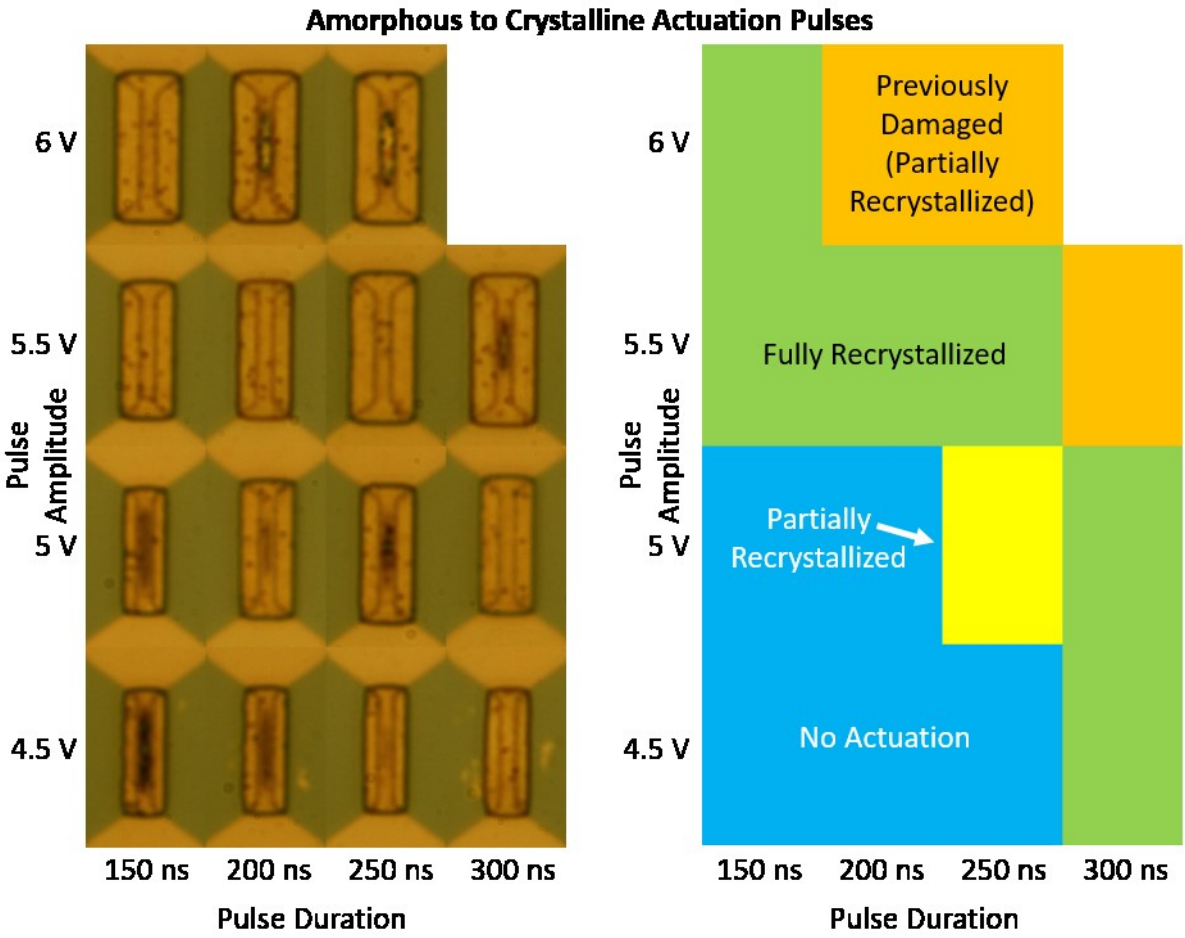

Figure 58. Optical microscope images of actuators after varied recrystallization pulses are applied. Pulse effects are categorized as: no actuation, partially recrystallized, fully recrystallized, and previously damaged (partially recrystallized).

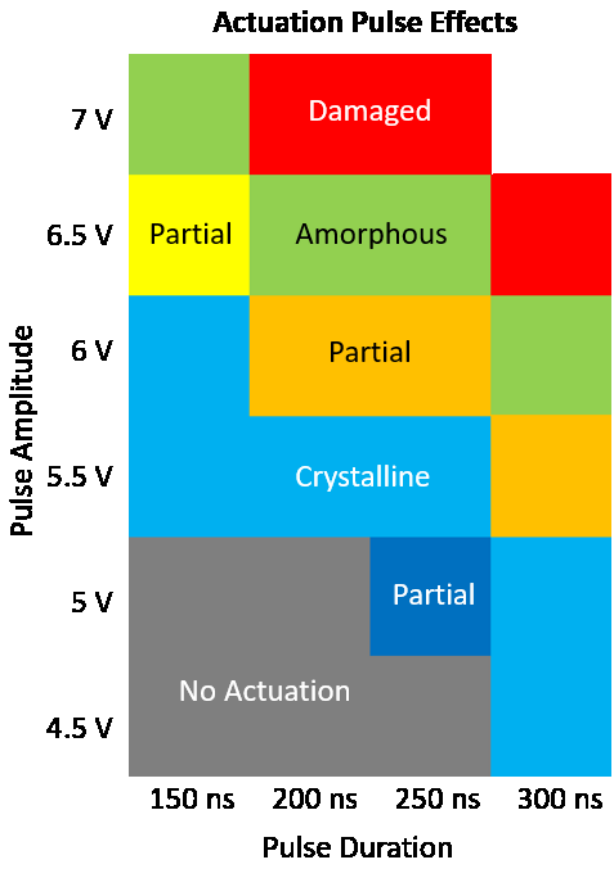

Figure 59. Actuation pulse effect matrix. The desired areas of operation are amorphous and crystalline. Pulses labeled as amorphous are used to expand the actuator, and pulses labeled as crystalline are used to contract the actuator.

4.3 Phase Change Material Reflow

Actuators fabricated without the Al_2O_3 top cap layer are also tested. During the actuation pulse GeTe is melted and held in the liquid state for a short period of time (100 – 200 ns). Reflow of the GeTe can be seen in Figure 60**Error! Reference source not found.**. Liquid GeTe tends to reform into balls as the GeTe does not wet well to the Al_2O_3 surface. Areas with reflow further corroborate that the heater is able to melt the GeTe and that the areas of GeTe above the melting temperature are limited to GeTe directly above the heater.

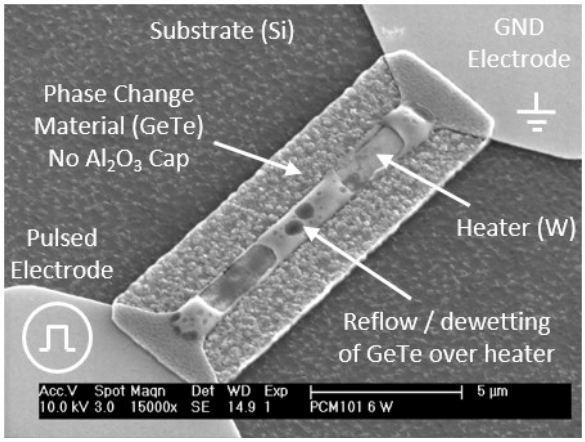

Figure 60. SEM image of an actuated device exhibiting reflow. This device does not have an Al_2O_3 cap, allowing GeTe to reflow and dewet from the heater.

4.4 Expansion Characterization via Atomic Force Microscopy

Mechanical expansion of the actuator was measured by atomic force microscopy (AFM) before and after various actuation pulses. All AFM measurements are referenced to the heater electrodes. A height cross-section is taken down the length of the heater for each actuation and is compared to determine the relative expansion or contraction of the actuator. Figure 61 shows an AFM measurement of an unactuated actuator with a 1.5 µm wide and 11 µm long heater. The patterned GeTe rises 200nm above the heater. The GeTe film shows high roughness in the form of raised bumps. This roughness is due to the older fabrication process which used a 10 nm layer of Al_2O_3 deposited at 150 °C by ALD as the isolation layer between the heater and GeTe.

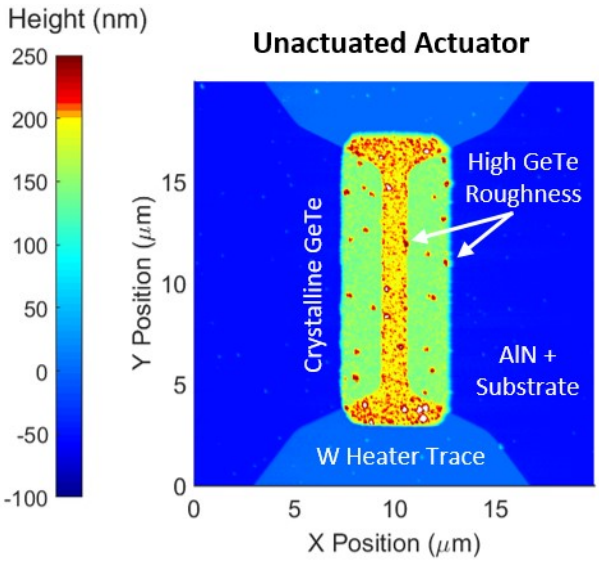

Figure 61. AFM measurement of an unactuated (as fabricated) actuator. A nonlinear color scale is used to highlight height variation of GeTe at the surface of the actuator. Height is referenced to the heater trace.

Figure 62 shows the AFM profile measurements for a device as fabricated and after a single actuation. Profile A is the as-fabricated actuator. All PCM is in the crystalline state, leading to a flat cross section profile. A 7 V 200 ns pulse is applied to the device, resulting in the shift from profile A to profile B. A section of PCM over the heater rises, matching the profile of amorphous material seen in the example optical measurements of Figure 55. The cross section profiles of A and B are compared, showing the height difference between the two states. An average height increase of 26 nm was measured over the actuated area. This increase in height is larger than 10 % of the film thickness, indicating some reflow of the GeTe while briefly in the liquid state.

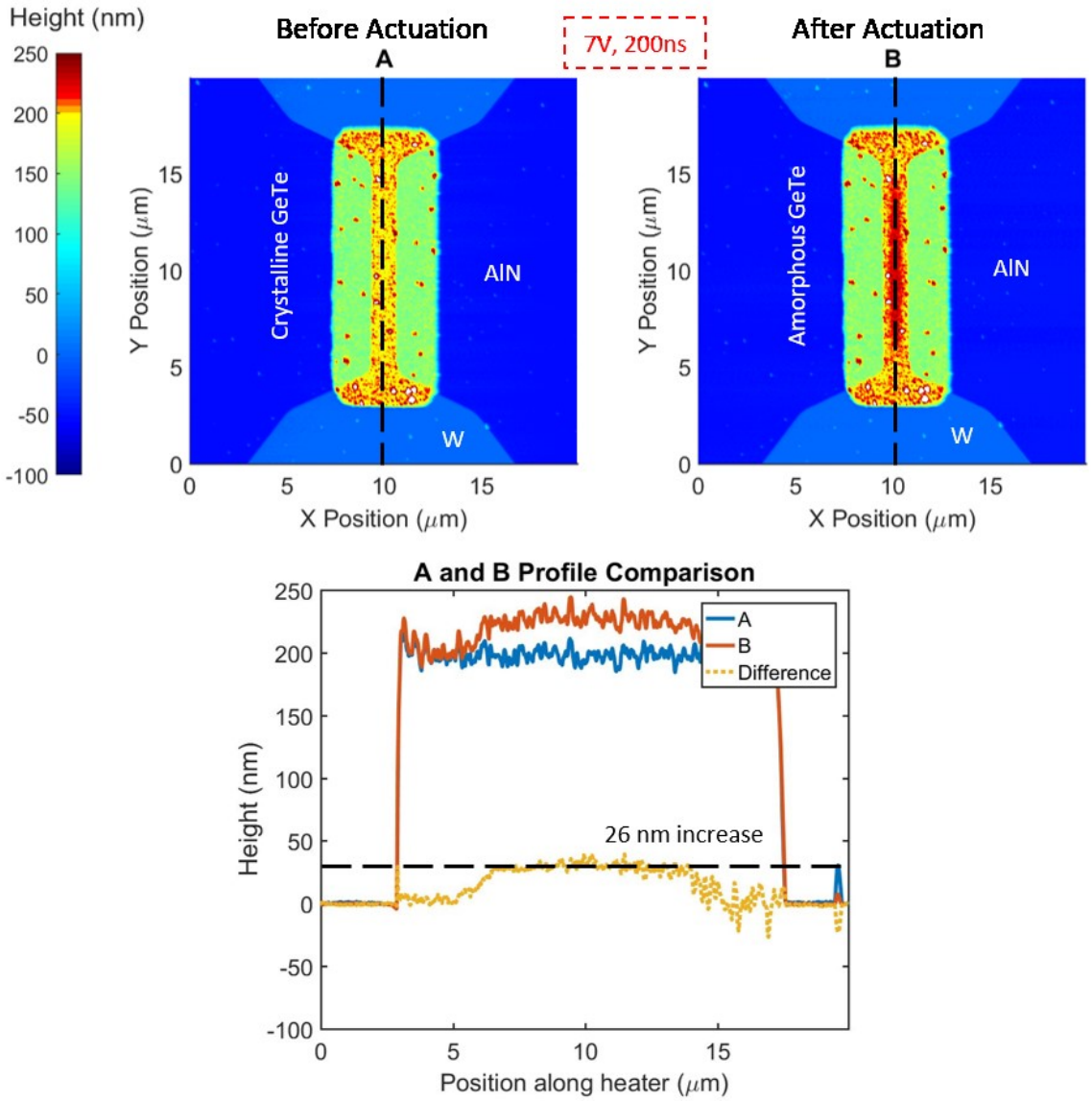

Figure 62. AFM measurements of a device before and after an actuation pulse. AFM profile A corresponds to an as fabricated unactuated device. AFM profile B corresponds to the same device after actuation. Profile comparison shows the increase in actuator thickness after actuation (profile B – profile A).

Next, a 6 V 200 ns pulse was applied to the device, resulting in the shift from profile B to profile C (Figure 63). The amorphous section of PCM contracts. The cross section profiles of B and C are compared, showing the height difference between the amorphous and re-crystalized PCM. An average height decrease of 14 nm is measured over the actuated area – 7 % of the fabricated PCM thickness. The center

of the actuator still remains partially expanded compared to the initially fabricated actuator (profile A). This hysteresis is attributed to two sources: reflow of GeTe and incomplete recrystallization. The initial expansion of the actuator is greater than expected, requiring some reflow of the GeTe, whereas the contraction cannot reflow since GeTe remains solid throughout the recrystallization process. Incomplete recrystallization is also a possible cause of the smaller contraction. Later testing of the fully fabricated switches shows a significantly longer (2 μs) recrystallization pulse improves the total number of cycles before failure.

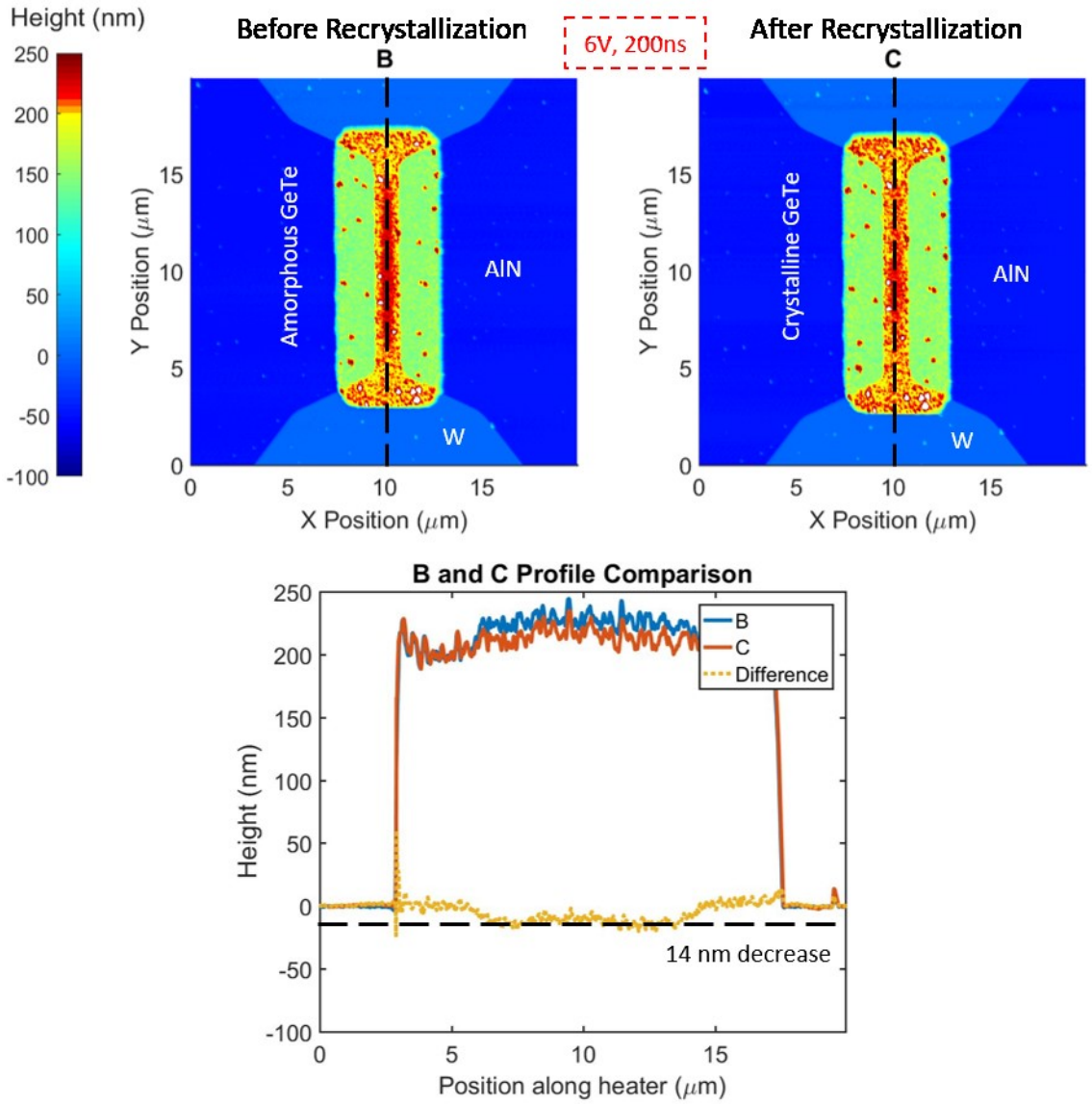

Figure 63. AFM measurements of a device before and after a recrystallization pulse. AFM profile B corresponds to the previously actuated device. AFM profile C corresponds to the same device after the recrystallization pulse. Profile comparison shows the decrease in actuator thickness after recrystallization (profile C – profile B).

Finally, a 7 V 200 ns pulse is subsequently applied to the device, resulting in the shift from profile C to profile D (Figure 64). The previously actuated section of PCM expands again. The cross section profiles of C and D are compared, showing the height difference between the previously re-crystalized and the

amorphous PCM. An average height increase of 16 nm is measured over the actuated area – 8 % of the fabricated PCM thickness. This increase in thickness is similar in magnitude to the recrystallization contraction indicating a new operating point of the actuator. Actuators appear to expand the most during the first conversion to the amorphous state (reflow) and then expand and contract a smaller amount around a new operating point. Multiple different actuators exhibited the same behavior, so the hysteretic expansion is not unique to the particular device being shown. This hysteretic expansion is actually an advantageous feature for the creation of a relay. The larger initial expansion is able to bridge a larger fabricated gap allowing for a wider tolerance in stress control for the drain and source contacts. Subsequent contractions and expansions are able to break and reconnect the contacts. The secondary expansion is somewhat larger than the contraction (2 nm or 1 %), likely indicating the incomplete recrystallization of the actuator

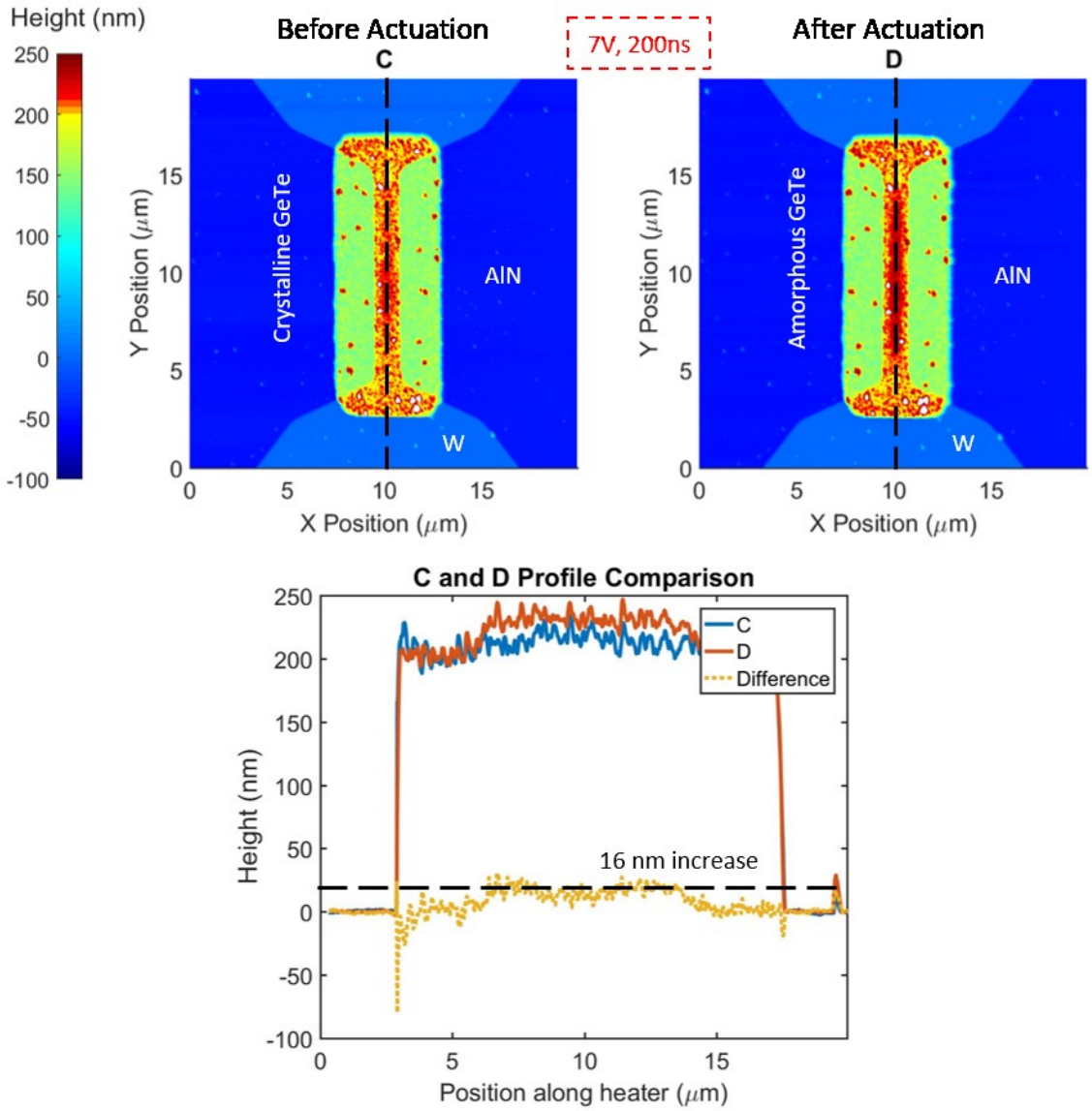

Figure 64. AFM measurements of a device before and after a second actuation pulse. AFM profile C corresponds to the previously recrystallized device. AFM profile D corresponds to the same device after the actuation pulse. Profile comparison shows the increase in actuator thickness after recrystallization (profile D – profile C).

4.5 Summary

In this chapter, tests on PCNR actuators is described. PCNR actuators were fabricated and tested to verify conversion from amorphous to crystalline GeTe, determine appropriate actuation pulses, measure

the expansion and contraction of the actuator, and determine the required sacrificial SiO_2 layer thickness for complete relay fabrication. Optical microscope measurements before and after actuation show GeTe is converted to the amorphous state in an area directly above the heater strip. A range of actuation pulse amplitudes (4.5 V – 7 V) and widths (150 ns to 300 ns) were characterized by optical microscope to determine the most appropriate actuation and recrystallization pulse. AFM measurements of multiple actuations show initially fabricated devices expanding and contracting for actuation and re-crystallization pulses respectively. Differences between expanded and contracted states show 7 – 8 % changes in thickness with an initial expansion of 13 %. These expansions show a 20 nm sacrificial layer is appropriate for a 200 nm thick GeTe actuator.

5 PCNR Relay Testing

5.1 Abstract

Completed PCNR devices are tested to determine key performance metrics: on and off switching times, contact resistance, leakage, non-volatility, reliability, and actuation voltage. Devices are cycled between on and off states until failure while measuring heater voltage and drain / source current. Ideal actuation pulses are designed to switch the PCNR from off to on as close to the end of the actuation pulse as possible. For 300 ns actuation pulses, a PCNR begins to turn on at 300 ns. Recrystallization pulses are significantly longer to ensure complete recrystallization of the actuator. When switching from on to off a PCNR begins to turn off after 0.6 µs.

I-V curves are used to measure contact resistance and leakage. Contact resistance is measured to be as low as 260 Ω for a PCNR or approximately 130 Ω per contact, but tends to be higher in the range of 1000's kΩ for "perfectly" actuated devices. Off state leakage is measured, but exhibits a breakdown event consistent with residue left in the air-gap after release. The residue is confirmed by TEM cross sections of

a released PCNR, where fluorinated alumina is found to fill the air-gap. It is suspected that contact is made only on the released edges of the source and drain.

Non-volatility is demonstrated by measuring contact resistance over a 24 hour period. No change in resistance is observed, although longer testing at elevated temperature is required to determine the limits of non-volatility. Reliability of the PCNR is determined by cycles before failure. Although low, an ideally actuated device is shown to switch up to 36 cycles before failure.

Actuation voltages range from 3.5 V to sub 1 V as heater dimensions shift from narrow and long (1 µm by 11 µm) to wide and short (1.5 µm by 3 µm). A comparison between experimental actuation voltages and those predicted by COMSOL shows good agreement for a range of heater lengths and widths, which verifies the accuracy of the model.

5.2 Switching Time

Actuator testing showed GeTe can be expanded and contracted with enough displacement to open and close the contact of a switch, however only a complete PCNR device can verify the phase change material can be used to actuated a switch. Measurement of switching time is the primary method to determine functionality of the switch as well as evaluate performance. Switching time requires simultaneous measurement of the actuation pulse and resistance between the drain and source contacts.

5.2.1 Test Setup

The fabricated devices are tested using the setup shown in Figure 65. The actuation pulses are generated by an Agilent B1110A pulse generator with a 50 Ω internal resistance. Rise and fall time is variable as well as pulse width and length. Voltage drop over 50 Ω internal resistance of the pulse generator allows for back calculation of the actuation current. One contact (the drain) is biased with a

1 V DC power supply. An Agilent N1025A 1 GHz differential probe senses voltage drop over a series 100 kΩ resistor, thereby measuring current through the drain and source.

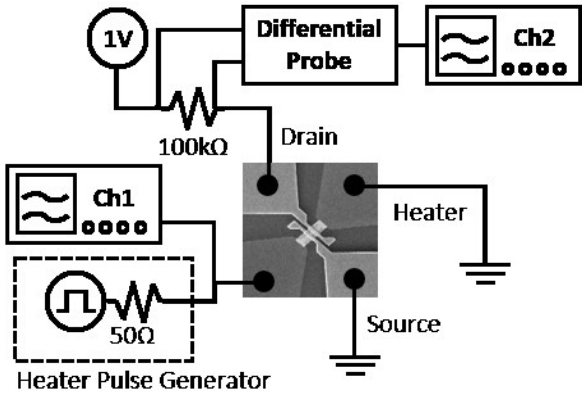

Figure 65. Relay actuation test setup. A pulse generator with 50 Ω internal impedance is used to drive the PCNR heater. A 1 V potential with a 100 kΩ resistor is used to sense contact between the source and drain.

Actuation pulse voltage is determined experimentally for each device with a unique heater combination. PCNR devices fabricated without the drain / source are actuated with pulses of increasing amplitude until the characteristic dark spot associated with amorphous phase change material is observed by optical microscope. Testing is conducted in both air and vacuum with minimal change in actuation voltage, although devices tested in vacuum exhibit larger numbers of cycles before failure. Devices are actuated through on and off cycles until failure. Failure is defined as when the tested device stops changing states with applied pulses. On and off time measurements are measured by back to back actuations of the same device.

5.2.2 On Time

Figure 66 shows oscilloscope traces of five sequential actuations of a device being switched on. This device has a heater with a 1.5 µm wide and 9 µm long heater and is actuated by a 300 ns wide pulse with 2 ns rise and fall times. The oscilloscope for the test setup directly records device actuation voltage

drop (voltage drop over probes, pads, and heater) and voltage drop over the 100 kΩ resistor in series with the drain and source. The device actuation voltage drop does not accurately reflect the voltage required to actuate the PCNR as the pads and probes are similar in resistance to the heater. The pads and probes have a total resistance around 30 Ω, which is comparable to the heater which has a resistance that ranges from about 15 Ω to 35 Ω as it changes temperature. Figure 67 shows heater voltage, heater current density, and drain / source current for multiple actuations of the same device. Heater voltage is significantly lower than input voltage as the traces are comparable in resistance to the heater. Heater voltage is not constant as the heater resistance increases with temperature. The beginning and end of the pulses also exhibit ripples due to reflections from the imperfect impedance matching between the pulse generator, oscilloscope, and device. Drain and source current becomes negative during the pulse. Electrical coupling with the heater is unlikely as the high amplitude reflection pulses are not visible in the drain and source current measurement. Coupling between the pulse generator and differential probe is also unlikely as this drop is only seen when connected to the drain and source. The cause of this negative current is unknown. Current flow through the drain and source begins to rise towards the end of the actuation pulse. This is not inherent to device operation, but rather the amplitude of the applied actuation pulse. The actuation pulse is the minimum required to actuate the device, and therefore by definition turns the device on at the end of the pulse. The rise time of the current is not necessarily indicative of the rate of change of contact resistance. The capacitance of the cable used to probe the drain and source must be discharged through the contact resistance, so rise time is limited by the time constant of the probing setup.

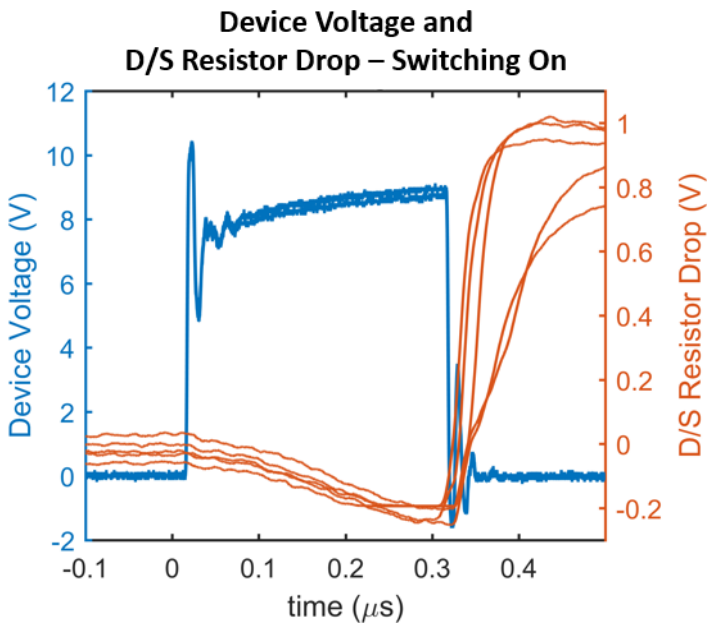

Figure 66. Unprocessed oscilloscope traces of five consecutive actuations of a device being turned on. Device voltage includes voltage drop over the probes and traces. D/S resistor drop is proportional to current flow through the drain and source.

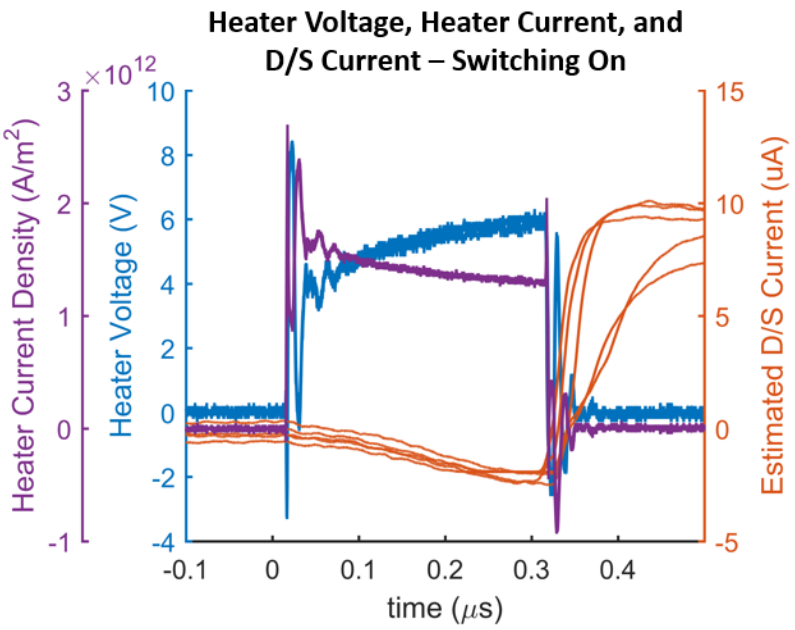

Figure 67. Extracted heater voltage, heater current, and D/S current of five consecutive actuations of a device being turned on. Heater voltage is only the voltage drop over the heater. Heater current density is the current normalized to the 1.5 µm by 50 nm heater cross-section.

5.2.3 Off Time

Figure 68 shows the oscilloscope traces and Figure 69 shows the extracted heater voltage, heater current density, and drain / source current for five sequential actuations of a device being switched off. This is the same device (1.5 µm by 9 µm heater) actuated in Figure 66 and Figure 67. Recrystallization is performed by a 2 µs wide pulse with a 100 ns rise time and 1 µs fall time. A slow fall time is used in favor of a fast one (2 ns) to prevent the formation of amorphous phase change material in the event some PCM is melted by the heater. The slow rise time ensures temperature falls slowly in the heater allowing any phase change material to recrystallize. Reflections are not visible as the rise time is slow (100 ns vs 2 ns) and does not excite large ripples in the actuation voltage. Voltage on the heater does not raise significantly during the pulse because the rise time of the actuation pulse is similar to the rise time of the heater. Drain / source current begins to fall around 0.6 ns after the beginning of the recrystallization pulse. Once

again the fall time of the current is limited by the RC time constant of the probing setup. This time constant is significantly longer than the switching on time constant because the resistance is higher. In this case the capacitance of the cable must be charged through the 100 kΩ rather than the resistance of the switch (around 200 Ω).

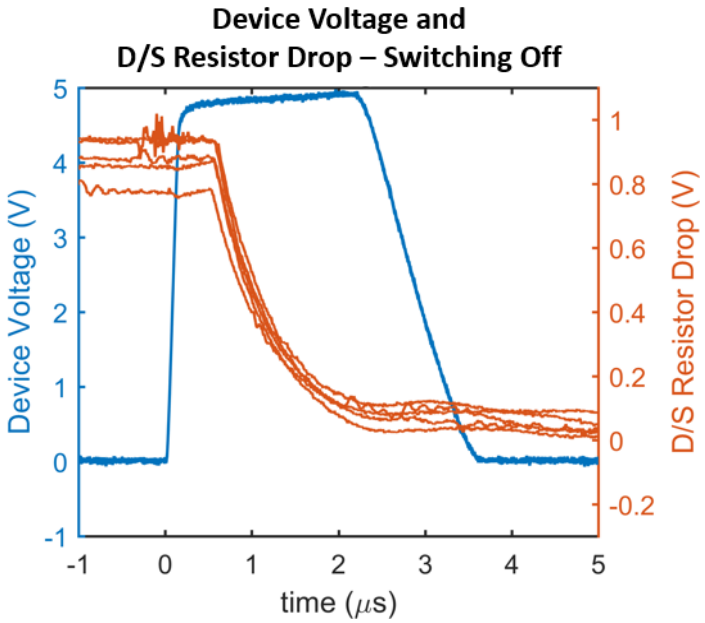

Figure 68. Unprocessed oscilloscope traces of five consecutive actuations of a device being turned off. Device voltage includes voltage drop over the probes and traces. D/S resistor drop is proportional to current flow through the drain and source.

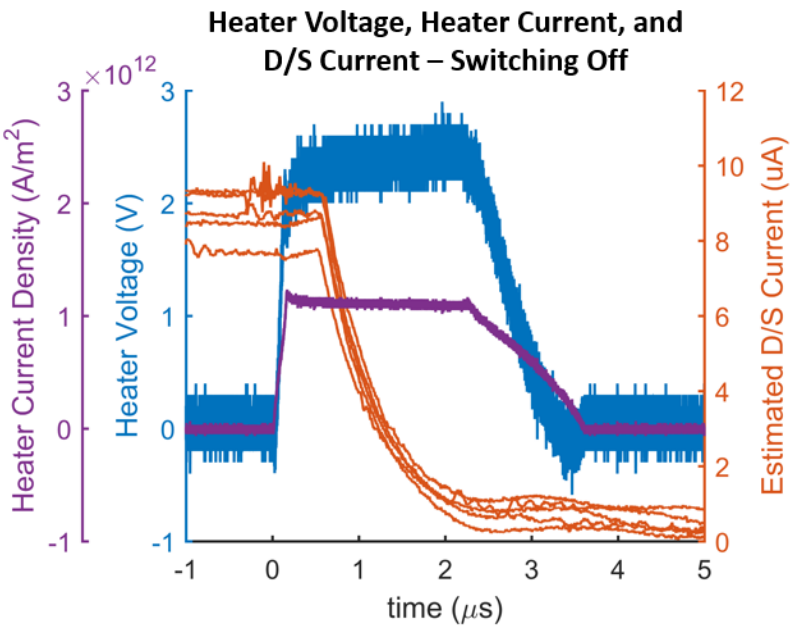

Figure 69. Extracted heater voltage, heater current, and D/S current of five consecutive actuations of a device being turned off. Heater voltage is only the voltage drop over the heater. Heater current density is the current normalized to the 1.5 µm by 50 nm heater.

The recrystallization pulse is significantly longer than the time to switch the device off (0.6 µs) as well as longer than the recrystallization pulse characterized during optical characterization of the actuator (200 ns). This extended pulse time ensures all amorphous material is recrystallized and requires lower voltage than the previously characterized pulse. Although the contact opens early in the pulse, the remainder of the pulse is needed to recrystallize all of the material and ensure the largest reopening of the air gap.

5.3 Contact Characterization

Device contacts are measured with an Agilent B1500A Semiconductor Device Parameter Analyzer to obtain I-V curves of the contact. The I-V curve measurements are obtained by sweeping varied voltage up and down while simultaneously measuring the current through the probe. Testing is performed in vacuum to avoid possible effects of air. I-V curves were measured with two probes, so resistance of the

probes and probe contact is included in the measurements. Measuring in a four point probe setup could remove this resistance by applying current and measuring voltage directly at the device, but the probe and contact resistance is negligible relative to the PCNR contact resistance or leakage. A two point probe setup is adequate to measure the contact resistance and leakage.

5.3.1 On State Contact Resistance

Contact resistance can vary depending on the applied actuation pulse. In the switching on time measurements of Figure 66 and Figure 67 devices are "perfectly" actuated, meaning the device turns on as the pulse ends. A longer or higher amplitude pulse would yield an "over" actuated device, meaning more energy was applied to the device than required for switching. "Over" actuated devices have the lowest contact resistance likely caused by the most complete expansion of the actuator (highest contact area and forces). An example measurement of on-state resistance is shown in Figure 70. As expected current is proportional to input voltage (ohmic) and it does not exhibit hysteretic behavior (the same profile as voltage is increased or decreased). Contact resistance is approximately 260 Ω, or 130 Ω per contact (under the assumption that each contact behaves in the same way), since the metallic contact of the switch must contact both the drain and source for current to flow.

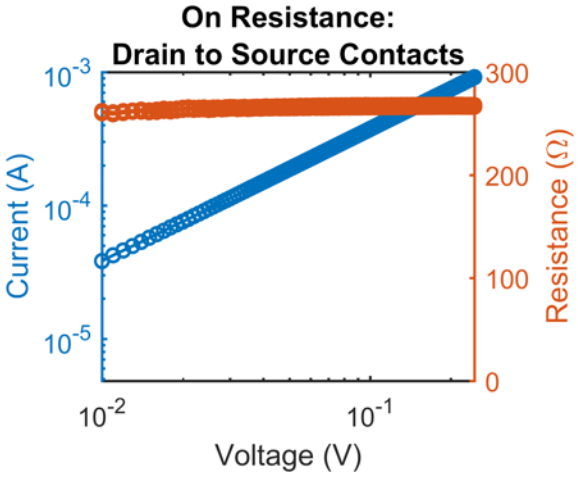

Figure 70. I-V curve measurement of a PCNR device in the on state.

Limited testing of non-volatility was performed. A device was actuated and then measured for 24 hours. A constant voltage of 10 mV was applied and current was measured once per minute. Figure 71 shows the evolution of the current as a function of time. As expected, current does not change throughout the 24 hour test. Contact resistance is measured to be 2.3 kΩ, which is higher than the previously measured on resistance. This higher on resistance is characteristic of a "perfectly" actuated device, while the 260 Ω on resistance is characteristic of an "over" actuated device. The 24 hour test demonstrates non-volatility of the on-state, but does not prove long term stability. Accelerated testing at elevated temperatures is required to measure the room temperature non-volatility of the PCNR.

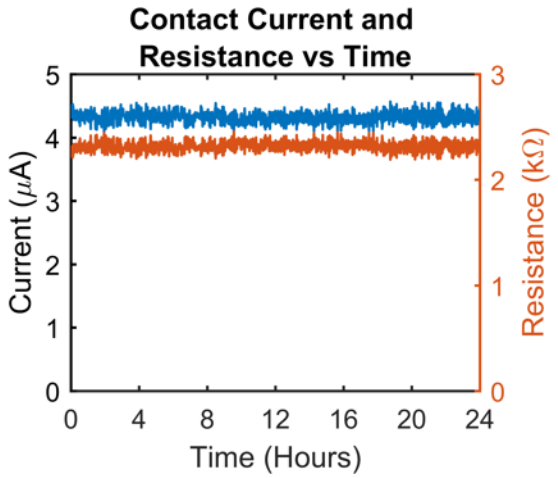

Figure 71. Current and resistance measurements over 24 hours. The device being tested for non-volatility is at room temperature.

On and off time measurements were performed by cycling devices until failure. The largest number of cycles before failure of a measured PCNR is 36. Figure 72 shows the evolution of contact resistance as a function of cycle count. The high cycle count can be attributed to "perfectly" actuating the device rather than "over" actuating the device. A single amplitude pulse is used until the vertical line, where amplitude is slightly increased to actuate the device more strongly and reduce contact resistance. The device eventually fails in the closed state, which is typical of other relays with high cycle counts.

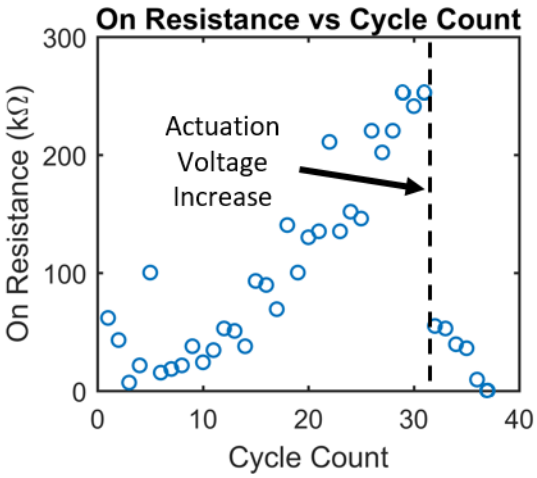

Figure 72. On resistance versus cycle count for a device cycling to failure. A small increase in actuation voltage is used to reduce on resistance after 31 cycles. Failure occurs after 36 complete cycles.

5.3.2 Off State Leakage

Off state leakage is important to characterize as low leakage is one of the defining characteristics of a NEMS relay. Leakage can come from other sources aside from the contacts. For example, the large overlap of the probing pads with the substrate can be a source of leakage. Devices with a dummy drain and source, where only the drain and source probing pads are fabricated, are used to characterize leakage to the substrate. Figure 73 shows the leakage of the dummy pads for a device. Leakage is below the noise floor at low voltage and shows no breakdown even over 60 V. The noise floor for this leakage test is higher than the other I-V curve measurements because a different sensing range was set on the parameter analyzer.

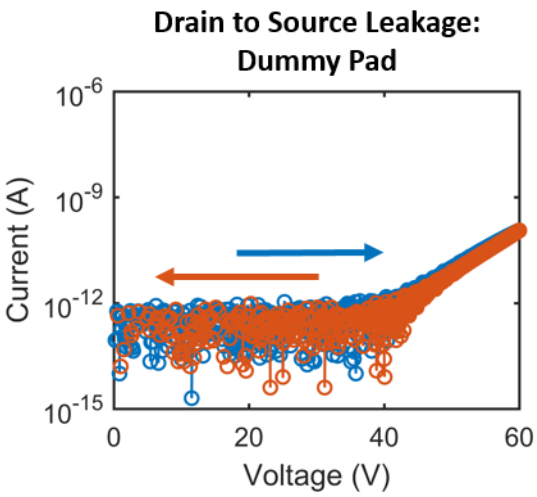

Figure 73. I-V curve measurement of a PCNR device with dummy drain and source pads. Current is below the noise floor for this configuration until an applied bias of 40 V.

A set of PCNR devices were measured before and after the vapor HF release etch to illustrate the difference between leakage through the sacrificial oxide and the air-gap. Figure 74 shows the leakage of a device before and after release. Before release the device exhibits around 10 fA of leakage current and does not exhibit any breakdown up to 15 V of applied bias. No hysteresis is visible in the current as the voltage increases and decreases. After release the device exhibits a breakdown starting at 8 V. Current saturates and follows a new trend. Large hysteresis is seen as the applied voltage increases and decreases. Subsequent I-V curve measurements (Figure 75) of the device follow the same path and do not exhibit hysteresis, indicating the breakdown of a film. This film is likely a residue left from the release process rather than an incomplete release since the unreleased device shows significantly lower leakage.

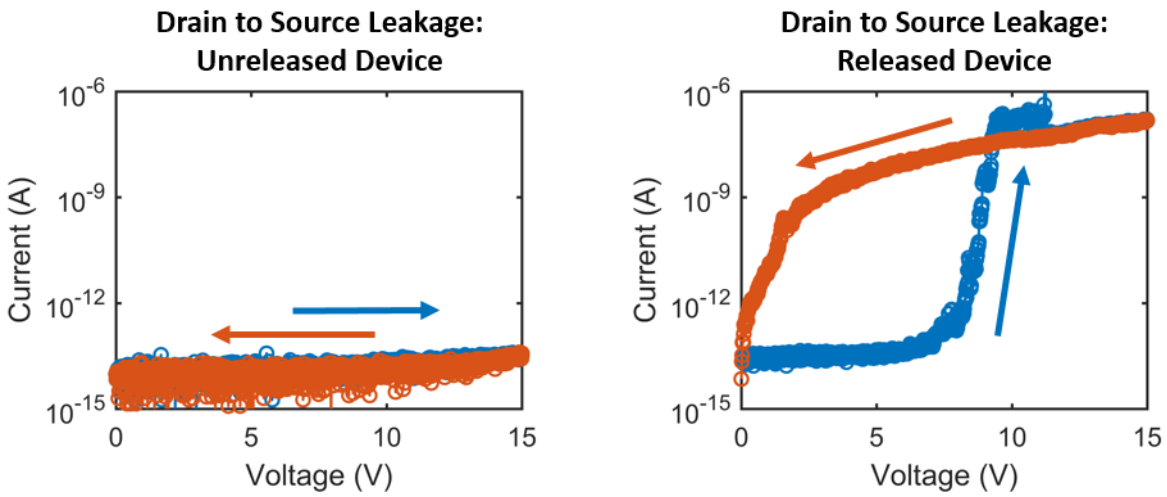

Figure 74. I-V curve measurements of an unreleased and released PCNR device. Breakdown can be seen in the hysteresis of the released device.

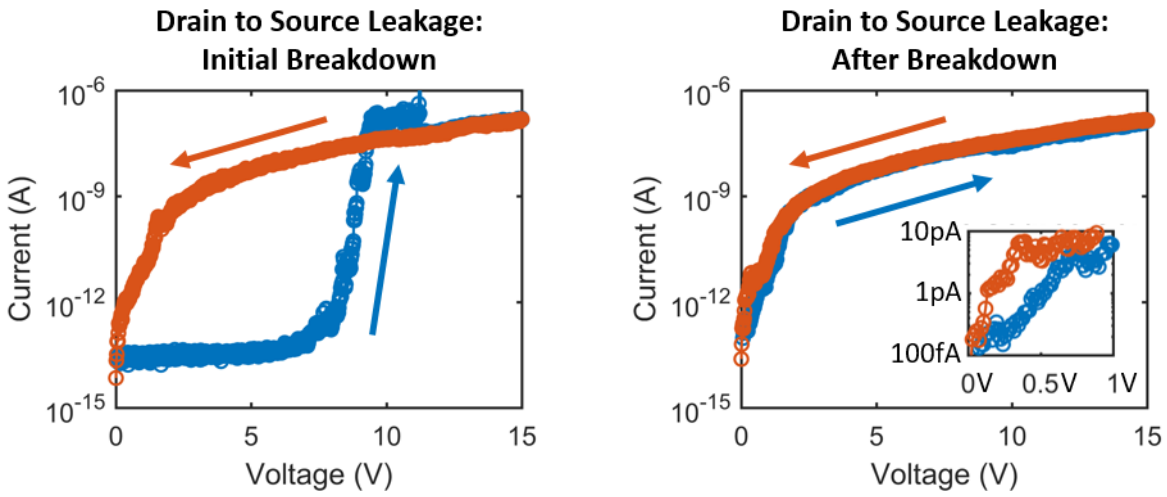

Figure 75. I-V curve measurements of the same released device at initial breakdown and after breakdown. The after breakdown I-V curve is consistent for subsequent I-V curves.

5.3.3 Release Contamination

The I-V curves in the off state indicate the possibility of residual sacrificial material in the air-gap. TEM images were taken of cross-section cuts to determine the presence of residual material in the air-gap. Figure 76 shows a schematic cross section and labeled TEM cross section of a released, but untested

device. The TEM sample is cut perpendicular to the heater, directly through the drain / source. The cross-section shows adequate conformal coverage of each layer. GeTe is fully encapsulated in the Al_2O_3 cap, the heater and metallic contact are fully isolated from the phase change material, and the metallic contact is fully separated from the drain / source contact. The drain / source shows some cracks at the step coverage of the actuator, however conductive paths remain and does not prevent function of the device. Although the air-gap appears to be properly formed it is contaminated with a residual from the release process. Figure 77 shows a TEM image looking directly at the metallic contact and air-gap.

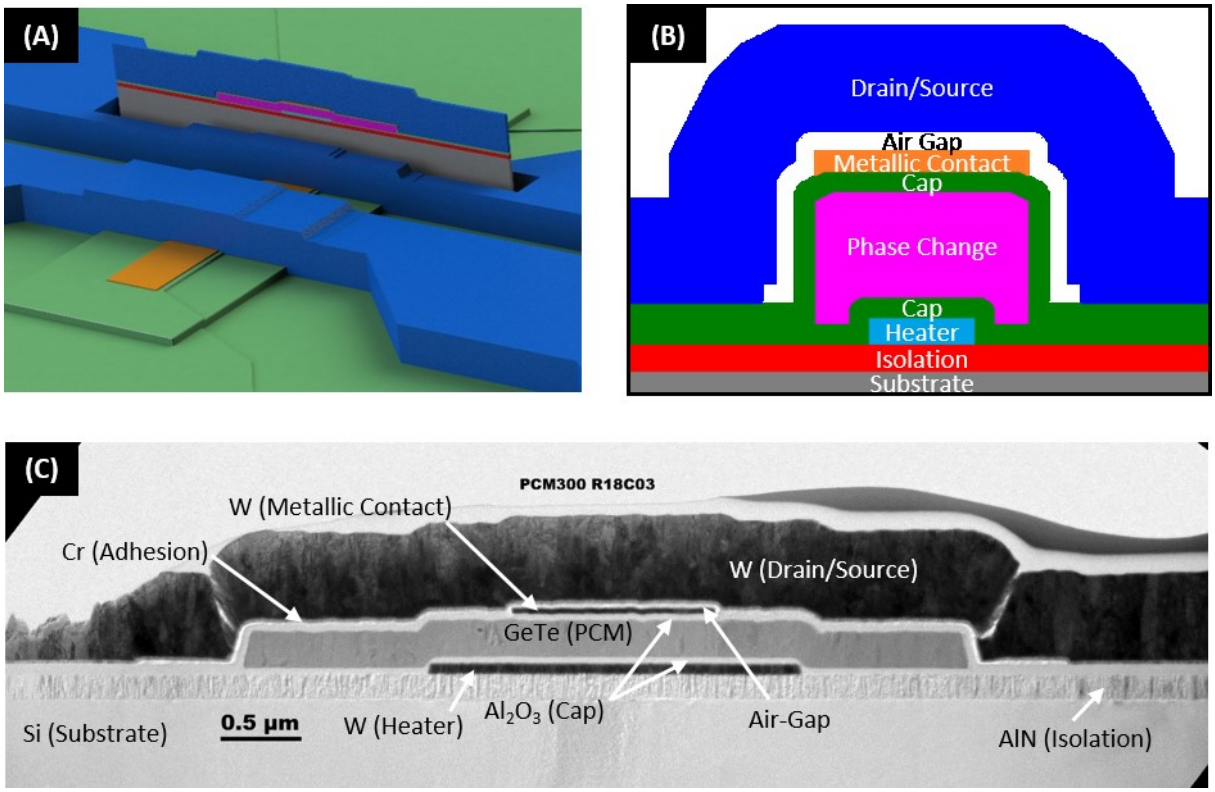

Figure 76. Illustrated schematics of the (a) TEM cross-section sample cut location and (b) the PCNR cross-section and (c) TEM image. (TEM courtesy of IBM Almaden Research)

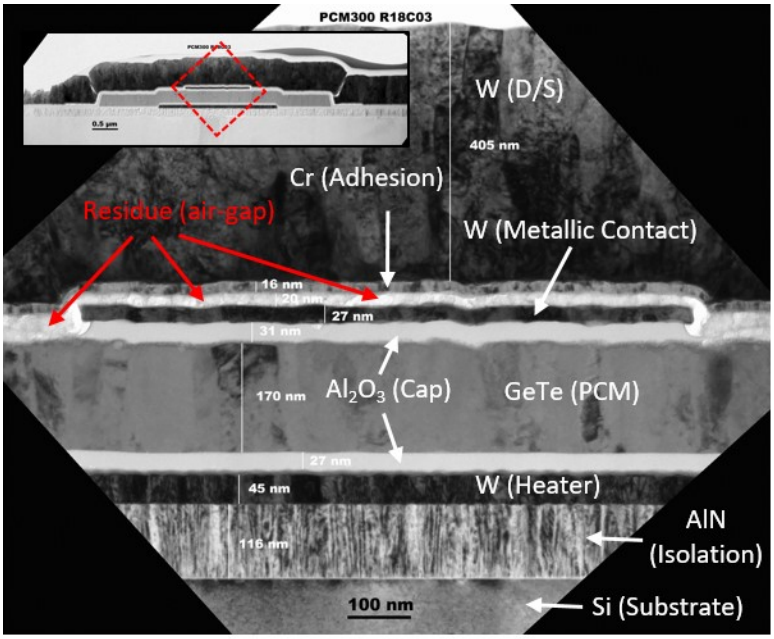

Figure 77. TEM cross-section focused on the metallic contact and air gap of a released unactuated PCNR. Residue can be seen in the air gap, likely preventing contact between the metallic contact and drain / source contacts. (TEM courtesy of IBM Almaden Research)

To get a better understanding of the contamination, elements of the cross-section are mapped using electron energy loss spectroscopy (EELS). Figure 78 shows the EELS map of a section of the PCNR centered on the metallic contact and air-gap. Four of the most prevalent elements in the device are shown: aluminum, oxygen, fluorine, tungsten, chrome, tellurium, and germanium. Silicon is not shown as little to no levels were detected. The "air-gap" shows no remaining silicon so the release etch can be deemed complete, however a combination of aluminum, oxygen, and fluorine can be observed in the air gap. Fluorinated alumina is present in the air gap. The vapor HF used to etch the SiO_2 sacrificial layer reacts with the Al_2O_3 cap layer and leaves residue even though previous tests of Al_2O_3 showed adequate protection of the GeTe PCM.

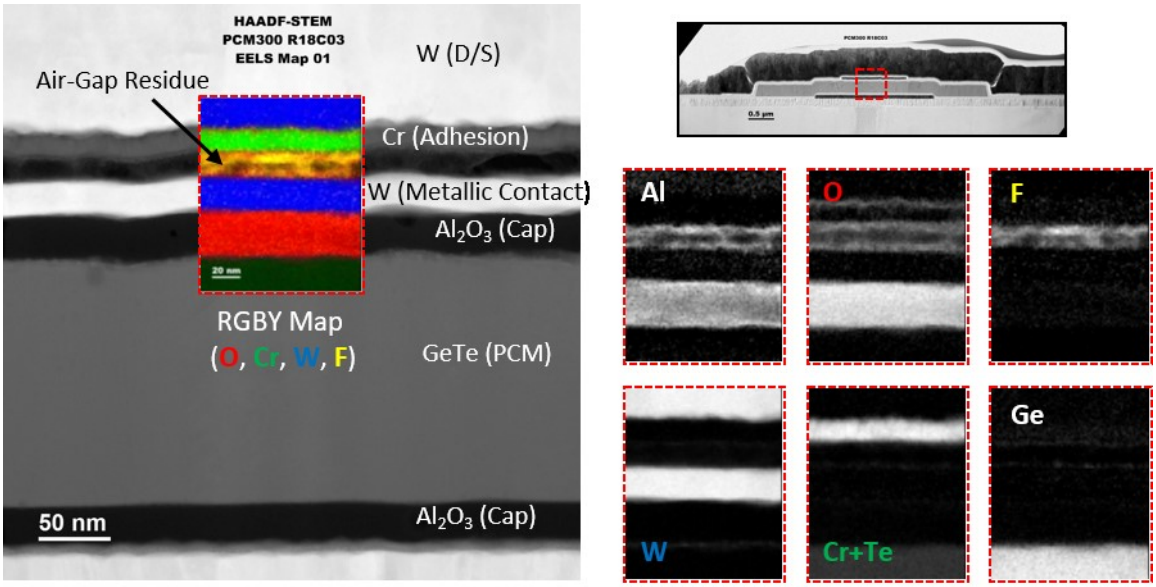

Figure 78. Electron energy loss spectroscopy (EELS) elemental mapping of materials in and around the air-gap. Fluorinated alumina can be seen as a residual material inside the air-gap. (TEM courtesy of IBM Almaden Research)

Residual fluorinated alumina contaminates the air gap and likely prevents contact between the metallic contact and the drain / source contacts. The contact points when the device is switched on are assumed to be at the edge of the drain / source. These contact points are highlighted in the illustration in Figure 79. Since the PCM is melted during actuation, reflow is able to deform the metallic contact to conform to the residue and drain / source to make contact. TEM cross-section images of an actuated PCNR along the length of the heater are required to confirm this theory.

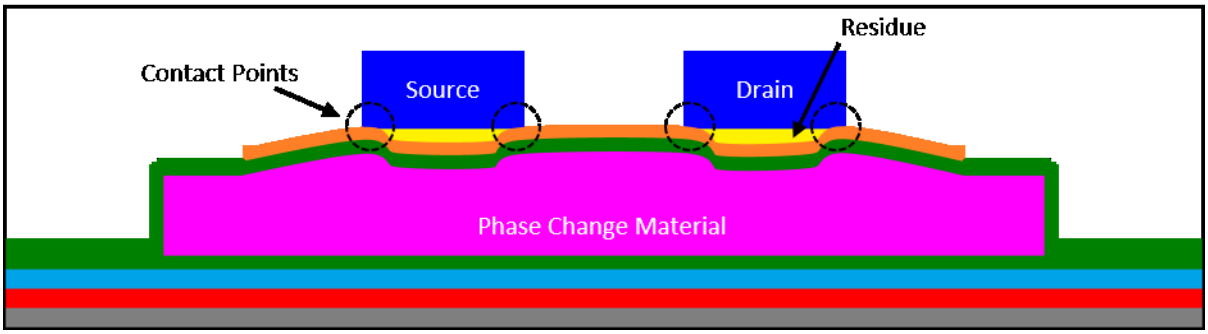

Figure 79. Schematic of contact points assuming fluorinated alumina residue blocks contact at the center of the drain and source.

5.4 Modeling Comparison

A range of PCNR devices with heaters of varied length (3 µm – 11 µm) and width (1 µm and 1.5 µm) were tested to determine the minimum actuation voltage. The minimum actuation voltage is the smallest amplitude pulse that is capable of switching the PCNR to the off state. Minimum actuation voltage is measured by testing for a switching event with progressively increasing actuation voltage. Recrystallization pulses must be applied after each actuation pulse to ensure a fresh actuator, as a partial actuation may occur at lower voltages that does not expand the actuator enough to make contact with the drain and source. Heater voltage is extracted from the applied pulse.

Minimum starting actuation voltage over the heater is used to compare fabricated devices with the COMSOL modeling. When the actuation pulse is applied, the heater increases temperature leading to an increase in temperature. Since the heater is in a voltage divider with the traces and pulse generator internal 50 Ω impedance, the heater voltage increases as temperature increases. The starting actuation voltage is the voltage drop over the heater before it has time to heat up. Since the applied pulse rise time is faster than the thermal time constant, this voltage is initially seen by the heater. The starting heater voltage is not however measured from the recorded pulse as reflections in the cable lead to inaccurate voltage measurements. The starting heater voltage is instead calculated from the input pulse and the

known heater resistance at room temperature. Figure 80 shows resistance measurements for 1 µm and 1.5 µm wide heaters of varied length (3 µm – 11 µm). Fitting to the resistance measurements for each heater length gives the heater resistivity and trace resistance. The trace resistance should be approximately the same for all devices, and is the Y intercept for the fit (30.2 Ω and 28.7 Ω for 1 µm and 1.5 µm wide heaters respectively). The heater resistance is 2.84 Ω/µm for 1 µm wide heaters and 1.72 Ω/µm for 1.5 µm wide heaters.

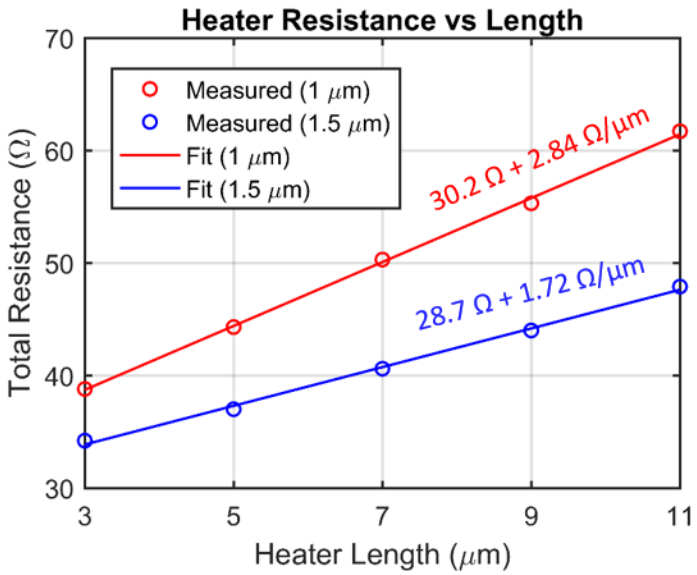

Figure 80. Heater resistance measurements for 1 µm and 1.5 µm wide heaters. Measurements include probe and pad resistance of the device. Fit slopes extract true heater resistance, while offsets extract true pad and probe resistance.

The same range of devices is simulated in COMSOL. Switching is defined as when the top surface of GeTe reaches 1050 K. The minimum starting actuation voltage is extracted from the simulation following the same process used for the fabricated devices. Figure 81 compares minimum starting actuation voltages for the COMSOL model and fabricated devices. Good agreement is seen between the measured PCNR devices and the COMSOL model. Both exhibit the expected linear relationship between actuation voltage and heater length, as devices are heated by current rather than voltage. A heater that is three

times longer requires three times the voltage on the heater. The strong agreement between the measured and simulated devices validates the modeling and gives credibility to the scaling predictions based on the COMSOL model.

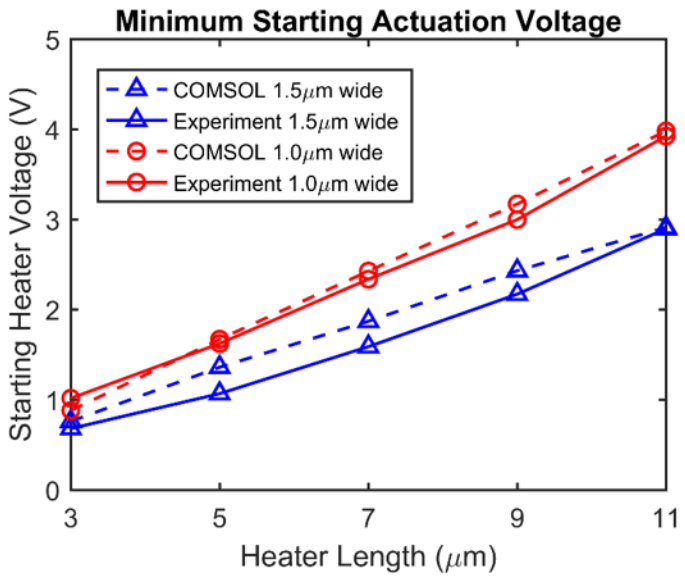

Figure 81. Comparison of minimum starting actuation voltages between devices simulated in COMSOL and measured fabricated devices. 1 µm and 1.5 µm wide heaters with lengths of 3 µm – 11 µm are measured.

5.5 Summary

In this chapter the PCNR device is tested to determine each of its properties: on and off switching times, reliability, non-volatility, contact resistance, leakage, and actuation voltage. Devices are cycled between on and off states until failure while measuring heater voltage and drain / source current. A typical failure leaves the relay shorted. Ideal actuation pulses are designed to switch the PCNR from off to on as close to the end of the actuation pulse as possible. For 300 ns actuation pulses, a PCNR begins to turn on at 300 ns. Recrystallization pulses are significantly longer to ensure complete recrystallization of the actuator. When switching from on to off a PCNR begins to turn off after 0.6 µs. Contact resistance is measured to be as low as 260 Ω for a PCNR or 130 Ω per contact, but tends to be higher in the range of

1000's kΩ for "perfectly" actuated devices. Non-volatility is demonstrated by measuring an invariable contact resistance over a 24 hour period. Although low, an ideally actuated device has been shown to switch up to 36 cycles before failing in the shorted state. Off state leakage is measured, but exhibits a breakdown event consistent with residue left in the air-gap after release. The residue is confirmed by TEM cross sections of a released PCNR, where fluorinated alumina is found to fill the air-gap. Actuation voltages range from 3.5 V down to sub 1 V as heater dimensions shift from narrow and long (1 µm by 11 µm) to wide and short (1.5 µm by 3 µm). Comparing experimental actuation voltages with those predicted by COMSOL show good agreement for a range of heater lengths and widths, verifying accuracy of the modeling.

6 Conclusion / Future Work

6.1 Conclusions

This thesis presented the development of a novel non-volatile NEMS relay that addresses the key scalability issues associated with traditional MEMS relays. The concept, design, fabrication, testing, and validation of the phase change NEMS relay (PCNR) is discussed in detail. This thesis presented both the first ever use of GeTe in a mechanical actuator as well as the first ever use of a GeTe actuator in a mechanical relay. Modeling was developed to estimate the expansion of a GeTe phase change actuator. Scaling analysis showed a path towards CMOS comparable device sizes (5 nm by 20 nm heater) as well as lower than CMOS actuation voltage (0.52 V). Actuation energy was predicted to be as low as 1.8 pJ with a 1.8 ns actuation time. A novel fabrication flow was developed to create the phase change actuator and form a narrow 20 nm air gap required for the relay.

Testing of the actuators showed initial expansions greater than 10 % of the fabricated film thickness. This large expansion was attributed to the reflow of phase change material when melted during switching. Repeatable contraction was also demonstrated, although smaller in magnitude than the initial

expansion due to the phase change material remaining solid during the recrystallization process. Testing of the full PCNR relay showed switching up to 36 cycles with on and off times of 300 ns and 600 ns, respectively. On state resistance was measured to be as low was 260 Ω (130 Ω per contact). Non-volatility was verified by monitoring of contact resistance over a 24 hour time period. Off state leakage was found to range from 10^{-14} A to 10^{-11} A at 1 V depending on the presence of a breakdown event. The breakdown event was attributed to the presence of residues in the air-gap. Measured heater actuation voltage was compared with COMSOL heater actuation voltage to validate the modeling. The model and measured data was found to match well for a range of heater widths and lengths.

6.2 Future Work

Initial testing of the PCNR has shown promise for a new highly-scalable relay technology, however additional work will be required to fully validate the relay technology before it can be eventually transferred to industry. The two key points to address are reliability and scalability. Reliability will primarily be addressed through improved materials and processing techniques. Scalability will need to be addressed through geometry changes in the PCNR design as well as testing of devices at reduced dimensions.

6.2.1 Improved Materials

One key problem with the currently fabricated PCNR is the residue left after release. This residue is most likely caused by the HF used in the release process reacting with the Al_2O_3 used for the encapsulation of the actuator. Switching to an alternate release process such as XeF_2 may mitigate the residue, however compatibility with other materials used in the relay is known to be poor. XeF_2 will readily etch the tungsten used for the contacts. Using XeF_2 for the release process also requires exchanging the sacrificial material. An alternate method to solving the residue issue is replacing the cap layer with another ALD compatible material. AlN has been shown to be deposited by ALD and is not attacked in vapor HF.

The significantly higher thermal conductivity of the AlN may change device performance and require modifications to the geometry to maintain scalability.

Although the relay only demonstrated up to 36 switching cycles, contact reliability is likely to become a bottleneck in achieving high cycle counts [93]. Tungsten and chrome are known to readily oxidize, and oxidation on the contacts is readily visible by the TEM cross-section elemental maps. Research has shown high quality platinum can greatly improve contact reliability [94]. One key difference between the PCNR and other MEMS relays is the high actuation temperature. The metallic contact will reach the melting point of the phase change material during actuation and may change the performance of the alloys. Alternatives to platinum like conductive metal oxides have also showed promise as they are likely to be resistant to the high temperatures of the switching process [41]. A thorough investigation of contact materials will be needed to address the contact reliability concerns for the PCNR.

6.2.2 Scaling and Alternate "Fin" Geometry

Further development of the PCNR is required before commercialization. The PCNR is presented as a highly scalable device, however the smallest demonstrated relay used a heater of 0.5 µm by 3 µm. Scaling analysis through COMSOL shows heater dimensions of 5 nm by 20 nm are achievable, however the influence of phase change material overlap and the encapsulation limit power scaling. An alternate geometry for these CMOS scale relays is desirable to further improve switching voltage and energy. One possible alternative geometry is a "fin" based PCNR. The fin based design, illustrated in Figure 82, differs from the currently fabricated PCNR in that it uses lateral expansion of the phase change material rather than vertical expansion. Preliminary analysis has shown lower actuation voltage, and lower actuation energy compared to the existing PCNR (Figure 83). Fabrication will need to change to incorporate more conformal coatings. Some research has shown metals like tungsten can be deposited by ALD and likely will not be a challenge for adoption of the fin PCNR [95].

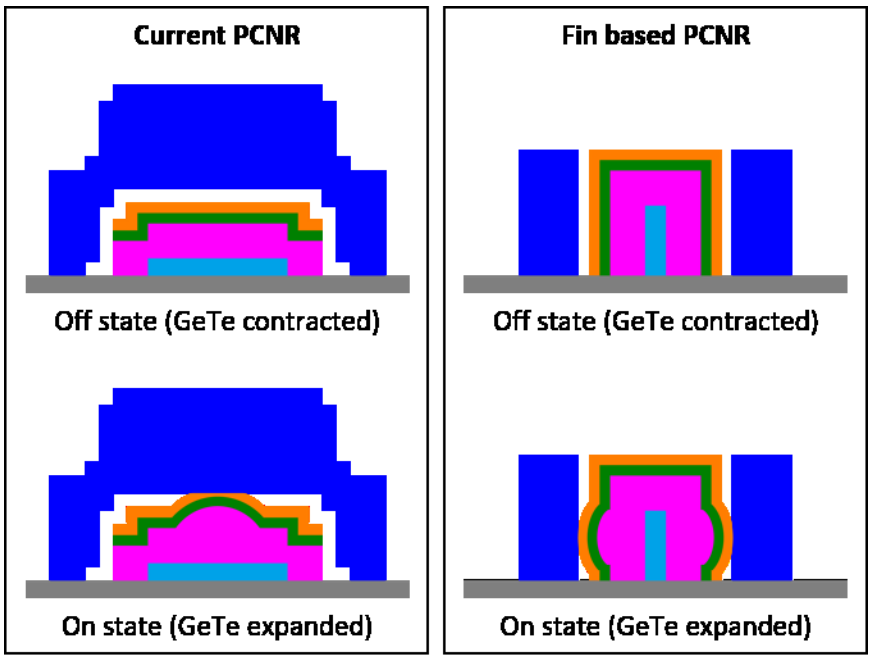

Figure 82. Schematics of the current PCNR geometry and the proposed fin PCNR geometry. The current PCNR expands vertically, while the fin PCNR expands laterally.

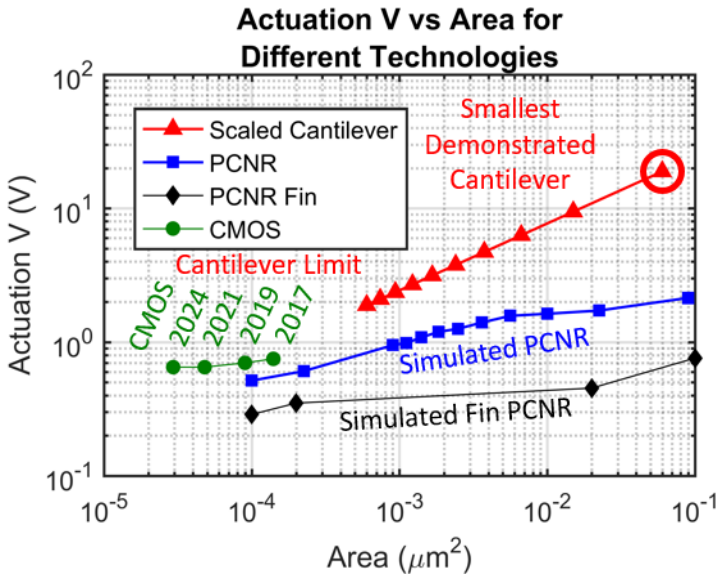

Figure 83. Comparison of actuation voltage and area between the "fin" PCNR, the highly scaled current PCNR geometry, the scaled electrostatic cantilever MEMS relay, and state of the art CMOS. CMOS data is pulled from [1] *and* [2].

References

[1] "The International Roadmap for Devices and Systems (IRDS)," 2017.

[2] "The International Roadmap for Devices and Systems (IRDS)," 2018.

[3] J. Xue, T. Li, Y. Deng, and Z. Yu, "Full-chip leakage analysis for 65 nm CMOS technology and beyond," *Integration*, vol. 43, no. 4, pp. 353–364, Sep. 2010.

[4] E. Morifuji, T. Yoshida, M. Kanda, S. Matsuda, S. Yamada, and F. Matsuoka, "Supply and threshold-Voltage trends for scaled logic and SRAM MOSFETs," *IEEE Trans. Electron Devices*, vol. 53, no. 6, pp. 1427–1432, Jun. 2006.

[5] Y. Taur, "CMOS design near the limit of scaling," *IBM J. Res. Dev.*, vol. 46, no. 2.3, pp. 213–222, Mar. 2002.

[6] S. Lai, "Non-volatile memory technologies: The quest for ever lower cost," in *2008 IEEE International Electron Devices Meeting*, 2008, vol. 2007, no. Figure 2, pp. 1–6.

[7] B. C. Lee *et al.*, "Phase-Change Technology and the Future of Main Memory," *IEEE Micro*, vol. 30, no. 1, pp. 143–143, Jan. 2010.

[8] L. Zhu, J. Zhou, Z. Guo, and Z. Sun, "An overview of materials issues in resistive random access memory," *J. Mater.*, vol. 1, no. 4, pp. 285–295, Dec. 2015.

[9] H. Akinaga and H. Shima, "Resistive Random Access Memory (ReRAM) Based on Metal Oxides," *Proc. IEEE*, vol. 98, no. 12, pp. 2237–2251, Dec. 2010.

[10] H. Wang and X. Yan, "Overview of Resistive Random Access Memory (RRAM): Materials, Filament Mechanisms, Performance Optimization, and Prospects," *Phys. status solidi – Rapid Res. Lett.*, vol. 13, no. 9, p. 1900073, Sep. 2019.

[11] P. Chi *et al.*, “PRIME: A Novel Processing-in-Memory Architecture for Neural Network Computation in ReRAM-Based Main Memory,” in *2016 ACM/IEEE 43rd Annual International Symposium on Computer Architecture (ISCA)*, 2016, pp. 27–39.

[12] J.-W. Jang, S. Park, Y.-H. Jeong, and H. Hwang, “ReRAM-based synaptic device for neuromorphic computing,” in *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2014, pp. 1054–1057.

[13] D. Apalkov, B. Dieny, and J. M. Slaughter, “Magnetoresistive Random Access Memory,” *Proc. IEEE*, vol. 104, no. 10, pp. 1796–1830, Oct. 2016.

[14] S. Tehrani, “Status and Outlook of MRAM Memory Technology (Invited),” in *2006 International Electron Devices Meeting*, 2006, pp. 1–4.

[15] T. M. Maffitt *et al.*, “Design considerations for MRAM,” *IBM J. Res. Dev.*, vol. 50, no. 1, pp. 25–39, Jan. 2006.

[16] G. W. Burr *et al.*, “Phase change memory technology,” *J. Vac. Sci. Technol. B, Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.*, vol. 28, no. 2, pp. 223–262, Mar. 2010.

[17] S. Ahn, “Phase Change Memory Reliability,” in *2006 IEEE International Integrated Reliability Workshop Final Report*, 2006, pp. 216–216.

[18] S. C. Lai *et al.*, “A scalable volume-confined phase change memory using physical vapor deposition,” in *2013 Symposium on VLSI Technology (VLSIT)*, 2013, pp. T132–T133.

[19] G. Betti Beneventi *et al.*, “Carbon-doped GeTe Phase-Change Memory featuring remarkable RESET current reduction,” in *2010 Proceedings of the European Solid State Device Research Conference*, 2010, pp. 313–316.

[20] D.-S. Chao *et al.*, "Impact of incomplete set programing on the performance of phase change memory cell," *Appl. Phys. Lett.*, vol. 92, no. 6, p. 062108, Feb. 2008.

[21] H. Mulaosmanovic *et al.*, "Switching Kinetics in Nanoscale Hafnium Oxide Based Ferroelectric Field-Effect Transistors," *ACS Appl. Mater. Interfaces*, vol. 9, no. 4, pp. 3792–3798, Feb. 2017.

[22] D. Reis *et al.*, "Design and analysis of an ultra-dense, low-leakage and fast FeFET-based random access memory array," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, pp. 1–1, 2019.

[23] K. Akarvardar and H.-S. P. Wong, "Nanoelectromechanical Logic and Memory Devices," *ECS Trans.*, vol. 19, no. 1, pp. 49–59, 2009.

[24] Hei Kam, Tsu-Jae King-Liu, E. Alon, and M. Horowitz, "Circuit-level requirements for MOSFET-replacement devices," in *2008 IEEE International Electron Devices Meeting*, 2008, vol. 11, no. 6, pp. 1–1.

[25] Daesung Lee *et al.*, "Combinational Logic Design Using Six-Terminal NEM Relays," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 32, no. 5, pp. 653–666, May 2013.

[26] K. Amponsah, N. Yoshimizu, S. Ardanuc, and A. Lal, "Near-kT switching-energy lateral NEMS switch," in *2010 IEEE 5th International Conference on Nano/Micro Engineered and Molecular Systems*, 2010, pp. 985–988.

[27] M. Spencer *et al.*, "Demonstration of Integrated Micro-Electro-Mechanical Relay Circuits for VLSI Applications," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 308–320, Jan. 2011.

[28] C. Chen *et al.*, "Efficient FPGAs using nanoelectromechanical relays," in *Proceedings of the 18th annual ACM/SIGDA international symposium on Field programmable gate arrays - FPGA '10*, 2010, p. 273.

[29] V. Rochus *et al.*, “Design of SiGe Nano-Electromechanical relays for logic applications,” in *2013 14th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)*, 2013, pp. 1–7.

[30] M. Ramezani, S. Severi, A. Moussa, H. Osman, H. A. C. Tilmans, and K. De Meyer, “Contact reliability improvement of a poly-SiGe based nano-relay with titanium nitride coating,” in *2015 Transducers - 2015 18th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, 2015, vol. 1, pp. 576–579.

[31] D. Grogg *et al.*, “Curved in-plane electromechanical relay for low power logic applications,” *J. Micromechanics Microengineering*, vol. 23, no. 2, p. 025024, Feb. 2013.

[32] T. He *et al.*, “Silicon carbide (SiC) nanoelectromechanical switches and logic gates with long cycles and robust performance in ambient air and at high temperature,” in *2013 IEEE International Electron Devices Meeting*, 2013, no. d, pp. 4.6.1-4.6.4.

[33] J. Rubin, R. Sundararaman, M. K. Kim, and S. Tiwari, “A single lithography vertical NEMS switch,” in *2011 IEEE 24th International Conference on Micro Electro Mechanical Systems*, 2011, pp. 95–98.

[34] X. L. Feng, M. H. Matheny, C. A. Zorman, M. Mehregany, and M. L. Roukes, “Low Voltage Nanoelectromechanical Switches Based on Silicon Carbide Nanowires,” *Nano Lett.*, vol. 10, no. 8, pp. 2891–2896, Aug. 2010.

[35] C. L. Ayala *et al.*, “A 6.7 MHz nanoelectromechanical ring oscillator using curved cantilever switches coated with amorphous carbon,” *Eur. Solid-State Device Res. Conf.*, pp. 66–69, 2014.

[36] W. W. Jang *et al.*, “NEMS switch with 30nm-thick beam and 20nm-thick air-gap for high density non-volatile memory applications,” *Solid. State. Electron.*, vol. 52, no. 10, pp. 1578–1583, Oct.

2008.

[37] W. W. Jang *et al.*, “Fabrication and characterization of a nanoelectromechanical switch with 15-nm-thick suspension air gap,” *Appl. Phys. Lett.*, vol. 92, no. 10, p. 103110, Mar. 2008.

[38] R. Nathanael, V. Pott, H. Kam, J. Jeon, and T.-J. K. Liu, “4-terminal relay technology for complementary logic,” in *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1–4.

[39] H. Kam, V. Pott, R. Nathanael, Jaeseok Jeon, Elad Alon, and Tsu-Jae King Liu, “Design and reliability of a micro-relay technology for zero-standby-power digital logic applications,” in *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1–4.

[40] T.-J. K. Liu, N. Xu, I.-R. Chen, C. Qian, and J. Fujiki, “NEM relay design for compact, ultra-low-power digital logic circuits,” in *2014 IEEE International Electron Devices Meeting*, 2014, vol. 2015-Febru, no. February, pp. 13.1.1-13.1.4.

[41] M. P. de Boer, D. A. Czaplewski, M. S. Baker, S. L. Wolfley, and J. A. Ohlhausen, “Design, fabrication, performance and reliability of Pt- and RuO 2 -coated microrelays tested in ultra-high purity gas environments,” *J. Micromechanics Microengineering*, vol. 22, no. 10, p. 105027, Oct. 2012.

[42] U. Zaghloul and G. Piazza, “Synthesis and characterization of 10 nm thick piezoelectric AlN films with high c-axis orientation for miniaturized nanoelectromechanical devices,” *Appl. Phys. Lett.*, vol. 104, no. 25, p. 253101, Jun. 2014.

[43] U. Zaghloul and G. Piazza, “10-25 NM piezoelectric nano-actuators and NEMS switches for millivolt computational logic,” in *2013 IEEE 26th International Conference on Micro Electro Mechanical Systems (MEMS)*, 2013, pp. 233–236.

[44] U. Zaghloul and G. Piazza, “Sub-1-volt Piezoelectric Nanoelectromechanical Relays With Millivolt Switching Capability,” *IEEE Electron Device Lett.*, vol. 35, no. 6, pp. 669–671, Jun. 2014.

[45] U. Zaghloul and G. Piazza, “Highly Scalable NEMS Relays With Stress-Tuned Switching Voltage Using Piezoelectric Buckling Actuators,” *IEEE Trans. Electron Devices*, vol. 61, no. 10, pp. 3520–3528, Oct. 2014.

[46] U. Zaghloul and G. Piazza, “Piezoelectric buckling-based NEMS relays for millivolt mechanical logic,” in *2014 IEEE 27th International Conference on Micro Electro Mechanical Systems (MEMS)*, 2014, pp. 1099–1102.

[47] W. Y. Choi, H. Kam, D. Lee, J. Lai, and T.-J. K. Liu, “Compact Nano-Electro-Mechanical Non-Volatile Memory (NEMory) for 3D Integration,” in *2007 IEEE International Electron Devices Meeting*, 2007, pp. 603–606.

[48] R. Parsa *et al.*, “Laterally Actuated Platinum-Coated Polysilicon NEM Relays,” *Microelectromechanical Syst. J.*, vol. 22, no. 3, pp. 768–778, May 2013.

[49] B. W. Soon, N. Singh, J. M. Tsai, and C. Lee, “A bi-stable silicon nanofin nanoelectromechanical switch based on van der Waals force for non-volatile memory applications,” in *The 8th Annual IEEE International Conference on Nano/Micro Engineered and Molecular Systems*, 2013, vol. 1, pp. 562–565.

[50] V. X. H. Leong *et al.*, “Vertical silicon nano-pillar for non-volatile memory,” in *2011 16th International Solid-State Sensors, Actuators and Microsystems Conference*, 2011, no. 4, pp. 649–652.

[51] H. F. Dadgour, M. M. Hussain, A. Cassell, N. Singh, and K. Banerjee, “Impact of scaling on the performance and reliability degradation of metal-contacts in NEMS devices,” in *2011*

International Reliability Physics Symposium, 2011, pp. 3D.3.1-3D.3.10.

[52] J. Yang, Y. Qi, H. D. Kim, and A. M. Rappe, "Mechanism of Benzene Tribopolymerization on the RuO2 (110) Surface," *Phys. Rev. Appl.*, vol. 9, no. 4, p. 044038, Apr. 2018.

[53] C. Oh, C. B. Stovall, W. Dhaouadi, R. W. Carpick, and M. P. de Boer, "The strong effect on MEMS switch reliability of film deposition conditions and electrode geometry," *Microelectron. Reliab.*, vol. 98, no. April, pp. 131–143, Jul. 2019.

[54] F. W. DelRio, M. P. de Boer, J. A. Knapp, E. David Reedy, P. J. Clews, and M. L. Dunn, "The role of van der Waals forces in adhesion of micromachined surfaces," *Nat. Mater.*, vol. 4, no. 8, pp. 629–634, Aug. 2005.

[55] J. A. Knapp and M. P. de Boer, "Mechanics of microcantilever beams subject to combined electrostatic and adhesive forces," *J. Microelectromechanical Syst.*, vol. 11, no. 6, pp. 754–764, Dec. 2002.

[56] J. Yaung, L. Hutin, J. Jeon, and T.-J. K. Liu, "Adhesive Force Characterization for MEM Logic Relays With Sub-Micron Contacting Regions," *J. Microelectromechanical Syst.*, vol. 23, no. 1, pp. 198–203, Feb. 2014.

[57] D. Lee, V. Pott, H. Kam, R. Nathanael, and T.-J. K. Liu, "AFM characterization of adhesion force in micro-relays," in *2010 IEEE 23rd International Conference on Micro Electro Mechanical Systems (MEMS)*, 2010, no. 1, pp. 232–235.

[58] W. Fang and J. A. Wickert, "Determining mean and gradient residual stresses in thin films using micromachined cantilevers," *J. Micromechanics Microengineering*, vol. 6, no. 3, pp. 301–309, Sep. 1996.

[59] R. Gaddi, C. Schepens, C. Smith, C. Zambelli, A. Chimenton, and P. Olivo, "Reliability and

performance characterization of a mems-based non-volatile switch," in *2011 International Reliability Physics Symposium*, 2011, pp. 2G.2.1-2G.2.6.

[60] C. Zambelli, P. Olivo, R. Gaddi, C. Schepens, and C. Smith, "Characterization of a MEMS-Based Embedded Non Volatile Memory Array for Extreme Environments," in *2011 3rd IEEE International Memory Workshop (IMW)*, 2011, pp. 1–4.

[61] V. Pott, G. L. Chua, R. Vaddi, J. M.-L. Tsai, and T. T. Kim, "The Shuttle Nanoelectromechanical Nonvolatile Memory," *IEEE Trans. Electron Devices*, vol. 59, no. 4, pp. 1137–1143, Apr. 2012.

[62] R. Vaddi, T. T. Kim, V. Pott, and J. T. M. Lin, "Design and analysis of anchorless shuttle nano-electro-mechanical non-volatile memory for high temperature applications," in *2012 IEEE International Reliability Physics Symposium (IRPS)*, 2012, p. ME.3.1-ME.3.6.

[63] V. Pott, R. Vaddi, G. L. Chua, J. T. M. Lin, and T. T. Kim, "Design Optimization of Pulsed-Mode Electromechanical Nonvolatile Memory," *IEEE Electron Device Lett.*, vol. 33, no. 8, pp. 1207–1209, Aug. 2012.

[64] R. Vaddi, V. Pott, G. L. Chua, J. T. M. Lin, and T. T. Kim, "Design and Scalability of a Memory Array Utilizing Anchor-Free Nanoelectromechanical Nonvolatile Memory Device," *IEEE Electron Device Lett.*, vol. 33, no. 9, pp. 1315–1317, Sep. 2012.

[65] J. T. Best and G. Piazza, "The pulse-activated piezo-NEMS shuttle relay," in *2017 IEEE 17th International Conference on Nanotechnology (IEEE-NANO)*, 2017, pp. 349–350.

[66] J. Best and G. Piazza, "Electrostatic actuation of the pulse-activated Piezo-NEMS shuttle relay," in *2018 IEEE Micro Electro Mechanical Systems (MEMS)*, 2018, vol. 2018-Janua, no. January, pp. 638–641.

[67] E. Wilhelm, C. Richter, and B. E. Rapp, "Phase change materials in microactuators: Basics,

applications and perspectives," *Sensors Actuators A Phys.*, vol. 271, pp. 303–347, Mar. 2018.

[68] K. Liu, C. Cheng, Z. Cheng, K. Wang, R. Ramesh, and J. Wu, "Giant-Amplitude, High-Work Density Microactuators with Phase Transition Activated Nanolayer Bimorphs," *Nano Lett.*, vol. 12, no. 12, pp. 6302–6308, Dec. 2012.

[69] Y. NAKAMURA, S. NAKAMURA, L. BUCHAILLOT, M. ATAKA, and H. FUJITA, "Two thin film shape memory alloy microactuators," *IEEJ Trans. Sensors Micromachines*, vol. 117, no. 11, pp. 554–559, 1997.

[70] A. V Kolobov and J. Tominaga, "Chalcogenide glasses in optical recording: recent progress," *J. Optoelectron. Adv. Mater*, vol. 4, no. 3, pp. 679–686, 2002.

[71] S. Senkader and C. D. Wright, "Models for phase-change of Ge2Sb2Te5 in optical and electrical memory devices," *J. Appl. Phys.*, vol. 95, no. 2, pp. 504–511, Jan. 2004.

[72] Y. Ding *et al.*, "A new Ge2Sb2Te5 (GST) liner stressor featuring stress enhancement due to amorphous-crystalline phase change for sub-20 nm p-channel FinFETs," in *2011 International Electron Devices Meeting*, 2011, vol. 5, pp. 35.4.1-35.4.4.

[73] R. Cheng *et al.*, "A new liner stressor (GeTe) featuring stress enhancement due to very large phase-change induced volume contraction for p-channel FinFETs," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, vol. C, pp. 93–94.

[74] T. P. L. Pedersen, J. Kalb, W. K. Njoroge, D. Wamwangi, M. Wuttig, and F. Spaepen, "Mechanical stresses upon crystallization in phase change materials," *Appl. Phys. Lett.*, vol. 79, no. 22, pp. 3597–3599, Nov. 2001.

[75] A. Fantini *et al.*, "Comparative Assessment of GST and GeTe Materials for Application to Embedded Phase-Change Memory Devices," in *2009 IEEE International Memory Workshop*, 2009,

pp. 1–2.

[76] L. Perniola *et al.*, “Electrical Behavior of Phase-Change Memory Cells Based on GeTe,” *IEEE Electron Device Lett.*, vol. 31, no. 5, pp. 488–490, May 2010.

[77] G. Bruns *et al.*, “Nanosecond switching in GeTe phase change memory cells,” *Appl. Phys. Lett.*, vol. 95, no. 4, p. 043108, Jul. 2009.

[78] W. J. Wang *et al.*, “Fast phase transitions induced by picosecond electrical pulses on phase change memory cells,” *Appl. Phys. Lett.*, vol. 93, no. 4, p. 043121, Jul. 2008.

[79] M. H. R. Lankhorst, B. W. S. M. M. Ketelaars, and R. A. M. Wolters, “Low-cost and nanoscale non-volatile memory concept for future silicon chips,” *Nat. Mater.*, vol. 4, no. 4, pp. 347–352, Apr. 2005.

[80] Y. C. Chen *et al.*, “Ultra-Thin Phase-Change Bridge Memory Device Using GeSb,” in *2006 International Electron Devices Meeting*, 2006, pp. 1–4.

[81] H. Zhang, Y. Zhang, Y. Yin, and S. Hosaka, “The dependence of crystallization on temperature in the nanosecond timescale for GeTe-based fast phase-change resistor,” *Chem. Phys. Lett.*, vol. 650, pp. 102–106, Apr. 2016.

[82] H. A. Jones, “A Temperature Scale for Tungsten,” *Phys. Rev.*, vol. 28, no. 1, pp. 202–207, Jul. 1926.

[83] S. R. Choi, D. Kim, S. H. Choa, S. H. Lee, and J. K. Kim, “Thermal conductivity of AlN and SiC thin films,” *Int. J. Thermophys.*, vol. 27, no. 3, pp. 896–905, 2006.

[84] E. Bozorg-Grayeli, J. P. Reifenberg, M. A. Panzer, J. A. Rowlette, and K. E. Goodson, “Temperature-Dependent Thermal Properties of Phase-Change Memory Electrode Materials,”

IEEE Electron Device Lett., vol. 32, no. 9, pp. 1281–1283, Sep. 2011.

[85] J. P. Reifenberg *et al.*, "Thermal Boundary Resistance Measurements for Phase-Change Memory Devices," *IEEE Electron Device Lett.*, vol. 31, no. 1, pp. 56–58, Jan. 2010.

[86] J. L. Bosse, M. Timofeeva, P. D. Tovee, B. J. Robinson, B. D. Huey, and O. V. Kolosov, "Nanothermal characterization of amorphous and crystalline phases in chalcogenide thin films with scanning thermal microscopy," *J. Appl. Phys.*, vol. 116, no. 13, p. 134904, Oct. 2014.

[87] P. Nath and K. L. Chopra, "Thermal conductivity of amorphous and crystalline Ge and GeTe films," *Phys. Rev. B*, vol. 10, no. 8, pp. 3412–3418, Oct. 1974.

[88] E. A. Scott, J. T. Gaskins, S. W. King, and P. E. Hopkins, "Thermal conductivity and thermal boundary resistance of atomic layer deposited high- k dielectric aluminum oxide, hafnium oxide, and titanium oxide thin films on silicon," *APL Mater.*, vol. 6, no. 5, p. 058302, May 2018.

[89] J. Martan, N. Semmar, C. Boulmer-Leborgne, P. Plantin, and E. Le Menn, "Thermal Characterization of Tungsten Thin Films by Pulsed Photothermal Radiometry," *Nanoscale Microscale Thermophys. Eng.*, vol. 10, no. 4, pp. 333–344, Dec. 2006.

[90] J. Liu, X. Xu, L. Brush, and M. P. Anantram, "A multi-scale analysis of the crystallization of amorphous germanium telluride using ab initio simulations and classical crystallization theory," *J. Appl. Phys.*, vol. 115, no. 2, p. 023513, Jan. 2014.

[91] A. Schlieper, Y. Feutelais, S. G. Fries, B. Legendre, and R. Blachnik, "Thermodynamic evaluation of the Germanium — Tellurium system," *Calphad*, vol. 23, no. 1, pp. 1–18, Mar. 1999.

[92] Y. Xia *et al.*, "Etching characteristics of phase change material GeTe in inductively coupled BCl3/Ar plasma for phase change memory," *Microelectron. Eng.*, vol. 161, pp. 69–73, Aug. 2016.

[93] C. Oh, F. Streller, R. W. Carpick, and M. P. de Boer, “Effectiveness of oxygen plasma versus UHV bakeout in cleaning MEMS switch surfaces,” in *2015 IEEE 61st Holm Conference on Electrical Contacts (Holm)*, 2015, vol. 2015-Decem, no. Cmmi, pp. 358–362.

[94] C. Oh and M. de Boer, “Effects of Hot Switching and Contamination on Contact Reliability of Pt-coated Microswitches,” *IEEE Trans. Components, Packag. Manuf. Technol.*, vol. PP, no. c, pp. 1–1, 2019.

[95] B. D. Davidson, D. Seghete, S. M. George, and V. M. Bright, “ALD tungsten NEMS switches and tunneling devices,” *Sensors Actuators A Phys.*, vol. 166, no. 2, pp. 269–276, Apr. 2011.

Appendix

Standard Photo-Resist Recipes

HMDS (vapor prime oven)

1. Place wafer(s) in glass wafer boat
2. Place in HMDS oven (set to 150 C)
3. Run recipe
    a. Pump purge time – 2.0
    b. Bake time – 10
    c. Pump time – 2.0
    d. HMDS time – 3.0
    e. Post purge time – 0.5
    f. Vent time – 2.5
    g. Pre purge cycles – 1
    h. Post purge cycles – 1
4. Remove from oven and let cool on table while still in glass wafer boat

AZ4110

1. Clean wafer with nitrogen gun for ~10 seconds (removes possible particles)
2. **Apply AZ4110 using dynamic spin**

   a. **<u>Use spin recipe: 00_4000</u>**
      i. 500 rpm spread step for 6 seconds
         1. 1000 rpm / sec ramp
      ii. 4000 rpm spin step for 2minutes
   b. Apply resist directly at center of wafer during spread step
      i. Squeeze dropper fully (including some of neck) to fill with maximum resist
      ii. DO NOT TILT DROPPER BACKWARDS
      iii. Drop a few droplets of resist on bench before application of resist (eliminates air bubbles)
      iv. Use 80% of dropper – make sure there is some resist left in dropper to avoid possible air bubbles
3. Soft bake resist
   a. Recipe: **<u>115C for 60 seconds, contact mode</u>** (don't use vacuum or proximity)
   b. Cool wafer on bench
4. Expose resist
   a. **<u>Expose at 200mj/cm^2 dose and 0.1um focus offset</u>**
5. Post exposure bake resist
   a. Recipe: **<u>115C for 60 seconds, contact mode</u>** (don't use vacuum or proximity)
   b. Cool wafer on bench
6. Develop (All glassware should say MIC)
   a. **<u>AZ Developer in 1:1 mixture with DI Water</u>**
      i. 300ml AZ Developer, 300ml Di Water
   b. Develop in large rectangular glass tray (easy to remove wafer, keeps wafer away from sidewall of tray)
      i. Agitate by sliding wafer across tray from side to side using wafer tweezers
      ii. **<u>Develop for 2:00 minutes nominal</u>**, but certain layers may require different times
   c. Dunk wafer in water bath, gently twist for 10 seconds while holding with wafer tweezers
   d. Rinse GENTLY with DI water gun for 60 seconds
   e. Dry GENTLY with nitrogen gun

S1805

1. Clean wafer with nitrogen gun for ~10 seconds (removes possible particles)
2. **Apply S1805 using dynamic spin**
    a. **Use spin recipe: 00_4000**
        i. 500 rpm spread step for 6 seconds
            1. 1000 rpm / sec ramp
        ii. 4000 rpm spin step for 2minutes
    b. Apply resist directly at center of wafer during spread step
        i. Squeeze dropper fully (including some of neck) to fill with maximum resist
        ii. DO NOT TILT DROPPER BACKWARDS
        iii. Drop a few droplets of resist on bench before application of resist (eliminates air bubbles)
        iv. Use 80% of dropper – make sure there is some resist left in dropper to avoid possible air bubbles
3. Soft bake resist
    a. Recipe: **115C for 60 seconds, contact mode** (don't use vacuum or proximity)
    b. Cool wafer on bench
4. Expose resist
    a. **Expose at 115mj/cm^2 dose and 0.05um focus offset**
5. Post exposure bake resist
    a. Recipe: **115C for 60 seconds, contact mode** (don't use vacuum or proximity)
    b. Cool wafer on bench
6. Develop (all glassware should say MIC)
    a. **Microposit Developer Concentrate in 1:3 mixture with DI Water**
        i. 100ml Microposit Developer Concentrate, 300ml Di Water
            1. Use same beaker 4 times (1 time for developer, 3 times for DI Water)
    b. Develop in large rectangular glass tray (easy to remove wafer, keeps wafer away from sidewall of tray)
        i. Agitate by sliding wafer across tray from side to side using wafer tweezers

ii. **Develop for 2:10 minutes nominal**, but certain layers may require different times

c. Dunk wafer in water bath, gently twist for 10 seconds while holding with wafer tweezers

d. Rinse GENTLY with DI water gun for 60 seconds

e. Dry GENTLY with nitrogen gun

NR9-3000PY

1. Clean wafer with nitrogen gun for ~10 seconds (removes possible particles)
2. **Apply NR9-3000PY using dynamic spin**
    a. **Use spin recipe: 00_4000**
        i. 500 rpm spread step for 6 seconds
            1. 1000 rpm / sec ramp
        ii. 4000 rpm spin step for 2minutes
    b. Apply resist directly at center of wafer during spread step
        i. Squeeze dropper fully (including some of neck) to fill with maximum resist
        ii. DO NOT TILT DROPPER BACKWARDS
        iii. Drop a few droplets of resist on bench before application of resist (eliminates air bubbles)
        iv. Use 80% of dropper – make sure there is some resist left in dropper to avoid possible air bubbles
3. Soft bake resist
    a. Recipe: **120C for 60 seconds, contact mode** (don't use vacuum or proximity)
    b. Cool wafer on bench
4. Expose resist
    a. **Expose at 300mj/cm^2 dose and 0.1um focus offset**
5. Post exposure bake resist
    a. Recipe: **120C for 60 seconds, contact mode** (don't use vacuum or proximity)
    b. Cool wafer on bench
6. Develop (All glassware should say MIF)
    a. **CD-26 Developer (pure, no mixture)**
        i. 300ml CD-26

b. Develop in deep circular crystallization glass (requires less CD-26 to submerge the wafer)
    i. Agitate by sliding wafer across tray from side to side using wafer tweezers
    ii. **Develop for 30 seconds nominal**, but certain layers may require different times
c. Dunk wafer in water bath, gently twist for 10 seconds while holding with wafer tweezers
d. Rinse GENTLY with DI water gun for 60 seconds
e. Dispose of developer in labeled waste bottle when done (cannot pour down drain)

Dry GENTLY with nitrogen gun

Resist Descum

1. Oxygen barrel asher
    a. 3 minutes, 1T, 100W
2. (don't use ozone cleaner or it will hardbake the resist)

Resist Strip

1. Acetone soak (as long as needed, typically overnight)
    a. 10-30 minutes ultrasonic loop (as needed)
        i. Sweeps through all frequencies for 30 seconds each
    b. Finish removal in Ozone Cleaner (30 minutes to 1.5 hours)
        i. Full oxygen flow, ozone generator on, uv on, and heat on
        ii. Start at 100C for the first 5 minutes
        iii. Change to 220C for remaining time
        iv. Remove wafer and let cool on table

PCNR Process Flow

1. Alignment Marks
    a. Start with wafer (HR for device wafers)
    b. Lithography
        i. HMDS application by vapor prime oven
        ii. AZ4110 standard recipe
        iii. Standard descum recipe
    c. Etch
        i. Plasmatherm 790 recipe LCSIO2ET
            1. O2 – 26 sccm / CHF3 – 2.5 sccm, 100 mT, 100W, 300 seconds
    d. Strip PR
        i. standard acetone soak recipe
2. AlN isolation/etchstop layer
    a. Use tegal with JB_100nm_AlN_3KW recipe. (has etch and thin conditioning steps)
        i. Depostion step:

Step	Time (s)	Thickness (A)	AC (kw)	DC (kw)	RF (w)	Argon 100 (sccm)	Nitrogen 50 (sccm)
1	15				50	23	60
2	2		3		50	23	60
3		10	3			23	32
4	208		3	0.1		23	32
5	10						

    b. Run 2 dummy runs before actual device wafer to warm up the machine
3. Heaters
    a. W deposition
        i. Use AJA system with high-temperature chuck for Tungsten
        ii. Make sure washers are underneath the clamp fingers (prevents wafer being cracked by too much pressure)

iii. Run 50nm W deposition with 10 min soak at 800c

b. Lithography

i. S1805 standard recipe

ii. Standard Descum recipe

c. Etch

i. Change gas on plasmatherm 790 from CF4 to SF6

ii. Etch with Jbest_W1 recipe

1. SF6 – 20 sccm, 20 mT, 30W, 60 seconds

iii. Change gas back to CF4

d. Strip PR

i. Standard acetone soak recipe

4. Alumina isolation

a. Al_2O_3 deposition

i. Use ALD system

1. Set temperatures to 250c (wait until all 4 measurements reach 250c)

2. Load sample

3. Run Al2O3 Thermal 250c recipe

a. Use 300 cycles to get a 30nm thick film

5. GeTe

a. GeTe deposition

i. Use AJA system with chuck for crystalline GeTe

ii. Make sure washers are underneath the clamp fingers (prevents wafer being cracked by too much pressure)

iii. Run 1000S GeTe deposition at 400c for crystalline GeTe (look in log to find exact name of recipe used)

b. Lithography

i. HMDS application by vapor prime oven

ii. AZ4110 standard recipe

iii. Standard descum recipe

c. Etch

i. Clean edge of wafer (3mm from edge) of all photoresist to prevent wafer getting stuck in tool

ii. Run GeTe_JBest recipe (70 seconds etch with Ar)

1. Make sure plasma strikes properly

2. 70S, 5mT, Ar-30 sccm, Bias – 100W, ICP – 400W, electrode – 25C, liner – 70C, lid – 120C, Spool – 120C.

d. Strip PR

i. Standard acetone soak recipe

6. Alumina cap

a. Al_2O_3 deposition

i. Use ALD system

1. Set temperatures to 250c (wait until all 4 measurements reach 250c)

2. Load sample

3. Run Al2O3 Thermal 250c recipe

a. Use 300 cycles to get a 30nm thick film

7. Top Contact

a. W deposition

i. Use AJA system with high-temperature chuck

ii. Make sure washers are underneath the clamp fingers (prevents wafer being cracked by too much pressure)

iii. Run 30nm W deposition RT (not heated)

b. Lithography

i. HMDS application by vapor prime oven

ii. AZ4110 standard recipe

iii. Standard descum recipe

c. Etch

i. Change gas on plasmatherm 790 from CF4 to SF6

ii. Etch with plasmatherm 790 recipe Jbest_W1 recipe

1. SF6 – 20 sccm, 20 mT, 30W, 60 seconds

iii. Change gas back to CF4

d. Strip PR

i. Standard acetone soak recipe

8. Sacrificial

a. SiO_2 deposition

i. Use ALD system

1. Set temperatures to 150c (standard temperature)
2. Load sample
3. Run SiO_2 plasma 150c recipe
    a. Use 300 cycles to get a 20nm thick film

b. Lithography

i. HMDS application by vapor prime oven

ii. AZ4110 standard recipe

iii. Standard descum recipe

c. Etch

i. Plasmatherm 790 recipe JBest_O1

1. O2 – 26 sccm / CHF3 – 2.5 sccm, 100 mT, 100W, 60 seconds

d. Strip PR

i. Standard acetone soak recipe

9. Vias

a. Lithography

i. HMDS application by vapor prime oven

ii. AZ4110 standard recipe

1. Hard-bake in oven at 150c for 10 minutes

iii. Standard descum recipe

b. Etch

i. Buffer HF wet etch – previously 90 seconds for 60nm of alumina total.

1. After etch, pour used HF back into the same bottle (it can be used many times)

2. Check resistivity of pads to make sure you can make electrical contact (complete etch). Resistance may be much higher than expected (like 200 ohms between two ground pads)

c. Strip PR

i. Standard acetone soak recipe

10. D/S

a. Cr deposition

i. Use CVC sputtering system

ii. Use recipe 20nm chrome

b. Lithography

i. HMDS application by vapor prime oven

ii. NR9-300PY standard recipe

c. W deposition

i. Use 5 target system

ii. Deposit 1um of W

1. 2 hours, 50W DC at 6.7mT pressure

d. Lift-off

i. Standard acetone soak recipe

e. Chrome strip

i. Versaline

1. Run recipe Longchang_Cr_Etch

a. 40 seconds, 10 mT, Cl2 – 40 sccm, O2 – 10 sccm, 50W bias, 600W ICP, electrode – 25C, Liner – 70C, Lid – 120C, Spool – 120C.

11. Dicing

a. Photoresist application

i. Clean wafer with nitrogen gun

ii. Apply AZ4110 dynamically (spin recipe: 00_4000)

iii. Bake at 95c for 2 minutes in contact mode (thermal recipe: 95c 2 min novac)

iv. Place on table until cold

b. Send to cleanroom staff for dicing

12. Vapor HF release

    a. Use recipe 3, each cycle should be 600 seconds, use 1 cycle and check to see if released. Add more cycles as needed.