SAND2018-7490C

Effects of Heavy Ion Radiation on the Classification Accuracy After Training of a TaO_X Hardware Based Neural Network (July 2018)

R. B. Jacobs- Gedrim, D. R. Hughart, S. Agarwal, G. Vizkelethy, E. S. Bielejec, B. L. Vaandrager, S. E. Swanson, K. E. Knisely, J. L. Taggart, H. J. Barnaby and M. J. Marinella

***Abstract*— The image classification accuracy of a TaO_x ReRAM based neuromorphic computing accelerator is evaluated after training with devices that have while intentionally induced displacement damage up to a fluence of 10^{14} 2.5 MeV Si ions / cm^2. Results are consistent with the radiation induced oxygen vacancy production mechanism. When the device is in the High Resistance State (HRS) during heavy ion radiation, the device resistance, linearity, and accuracy after training are only affected by high fluence levels. The findings in this study are in accordance with results of previous studies on TaO_x based digital resistive random-access memory. When the device is in the Low Resistance State (LRS) during irradiation, no resistance change was detected, but devices with a 4k in line resistor did show a reduction in accuracy after training at 10^{14} 2.5 MeV Si ions / cm^2, indicating that changes in resistance can only be somewhat correlated with changes to a devices' analog properties. This study demonstrates that TaO_x devices are radiation tolerant not only for high radiation environment digital memory applications, but also when operated in an analog mode suitable for neuromorphic computation and training on new datasets.**

***Index Terms*— Analog, Deep Learning, Neuromorphic Accelerator, Radiation Hardened, ReRAM, Tantalum Oxide**

I. Introduction

Resistive Random Access Memory (ReRAM) is one of the leading candidates of beyond CMOS non-volatile memory technologies. A typical ReRAM structure consists of two metal terminals sandwiching a substoichiometric metal oxide layer which is dielectric in the as-deposited state. The current voltage (I/V) characteristics of a typical Sandia TaO_x device are shown in Fig. 1. A diagram of the device cell is shown in the inset for Fig. 1. The I/V characteristics of the virgin device is shown on the left. The last step of device fabrication is a process known as FORMING shown in the middle of Fig. 1. When positive bias is applied to the top electrode of the

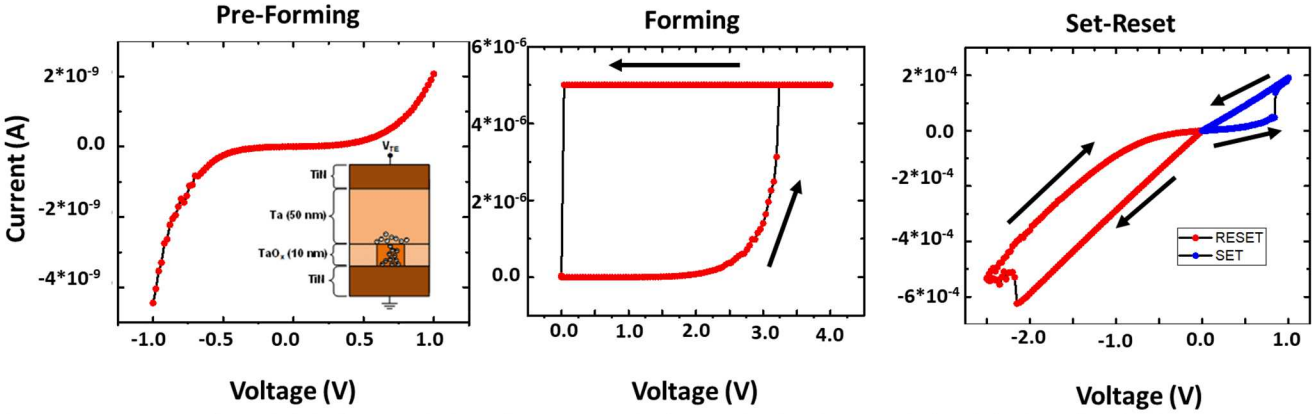

Figure 1. Current Voltage (I/V) characteristics of a ReRAM device. Left: I/V characteristic of device after fabrication and (Inset) diagram of device stack. Middle: I/V characteristics of soft breakdown process known as Forming. Right: Switching the device from high to low resistance and back again using Set and Reset operations. Elements of this figure are reproduced from [1].

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.M.

Robin Jacobs-Gedrim, Sandia National Laboratories, PO Box 5800, Mail Stop 1084, Albuquerque, NM 87185-1084, USA; phone: 505-284-8941, email: rbjaco@sandia.gov

D. R. Hughart, G. Vizkelethy, E. S. Bielejec, B. L. Vaandrager, S. E. Swanson and M. J. Marinella are with Sandia National Laboratories, Albuquerque, NM 87185

J. L. Taggart and H. J. Barnaby, School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706

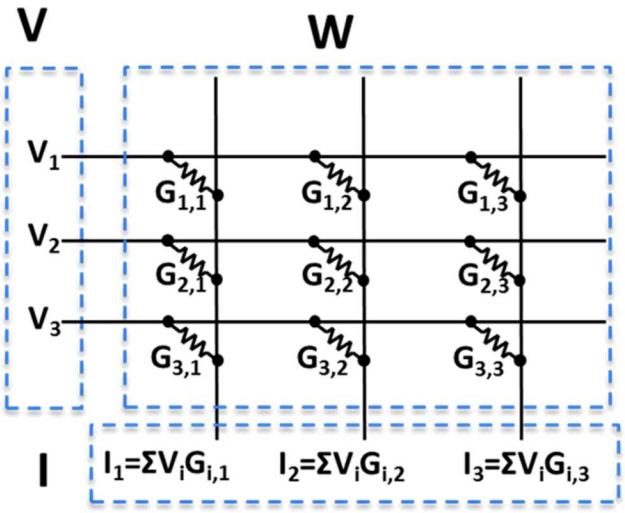

Figure 2: A vector matrix multiply may be implemented on a crossbar with ReRAM analog electronic elements. Figure reproduced from [1].

device inducing a soft breakdown in the TaO_x and producing a Ta rich filament across the oxide. Formed device operation is shown in the right figure. A negative bias applied to the top electrode can increase the resistance of the device, in a process known as RESET, which places the device in the High Resistance State (HRS). A subsequent positive bias can decrease the resistance of the device again, in a process known as SET which places the device in the Low Resistance State (LRS). The current-voltage (I/V) characteristics of these three processes are shown in Fig. 1.

In addition to computer memory applications, it has been suggested that ReRAM may be used for training deep networks [2, 3]. Deep learning is a pattern recognition algorithm that is capable of outperforming traditional machine learning techniques in image recognition, autonomous vehicles, and data science applications. Training deep networks is computationally intensive, and difficult to implement into portable, embedded systems (e.g. spacecraft, rovers, satellites etc.). Recent analysis has demonstrated that a special-purpose accelerator, application-specific integrated circuit (ASIC), based on analog ReRAM crossbars, can improve the performance per watt by about two orders of magnitude over a digital system – potentially enabling real-time training on embedded systems [3]. The performance increase is obtained by carrying out the two most computationally intensive steps in deep network training, a vector matrix multiply and weight update in one parallel step [4]. An analog electronic vector matrix multiply is diagramed in Fig. 2. By applying bias to the left side of the crossbar, Kirchhoff's laws give the current on the bottom of each column as sum of the weights in the column of the crossbar, multiplied by the applied voltage amplitude. Using the programmable conductance of ReRAM devices to represent weights in the matrix can allow the accelerator to not only carry out recognition tasks, but also train on new datasets with high energy efficiency by applying updates to the array using an outer product update. While there are significant energy advantages to using analog ReRAM devices as weights, over a conventional digital memory, the accuracy of the accelerator will negatively affected by nonlinear device characteristics [4].

It is necessary to study the impact of radiation on analog device characteristics of TaO_x devices and what impact this would have on a neuromorphic computing accelerator capable of training. Radiation studies for digital ReRAM based on tantalum [5, 6] titanium [7, 8] , and hafnium [9-11] oxide devices have demonstrated promising radiation tolerance results. TaO_x ReRAM showed gradual resistance degradation only at high fluences of Si and Ta ions due to additional oxygen vacancies introduced by ion displacement damage to the switching region [5, 6] [12-14].

While these studies have all shown high tolerance to radiation for ReRAM operated in a digital mode, these results do not necessarily imply that ReRAM devices operated in an analog mode for training in neuromorphic computing will have equivalent radiation tolerance. In digital operation of a ReRAM device, small variations in conductance do not matter if the device consistently switches between LRS and HRS. In the same device, instead operated in an analog mode, small changes in conductance trend matter as small changes in device linearity can affect accuracy after training on a ReRAM based neuromorphic computing accelerator [15]. An ideal analog device would have a completely linear conductance response, where each update results in an equal change in conductance, no matter what the starting conductance of the device before the update. TaO_x devices are far from perfectly linear in their conductance response to SET or RESET pulses. It is conceivable that a radiation dose far lower than the threshold required to make a digital TaO_x device lose its state, could effect a significant negative change in the analog conductance trend of the same device. J. L. Taggart et. al. showed a change to the analog conductance response in Conductive Bridge Random Access Memory (CBRAM) devices under ionizing radiation of a total dose greater than 1 Mrad [16]. An open question remains regarding the impact (if any) changes in the analog conductance response of a device under radiation will have on the ability of an analog neuromorphic computing accelerator

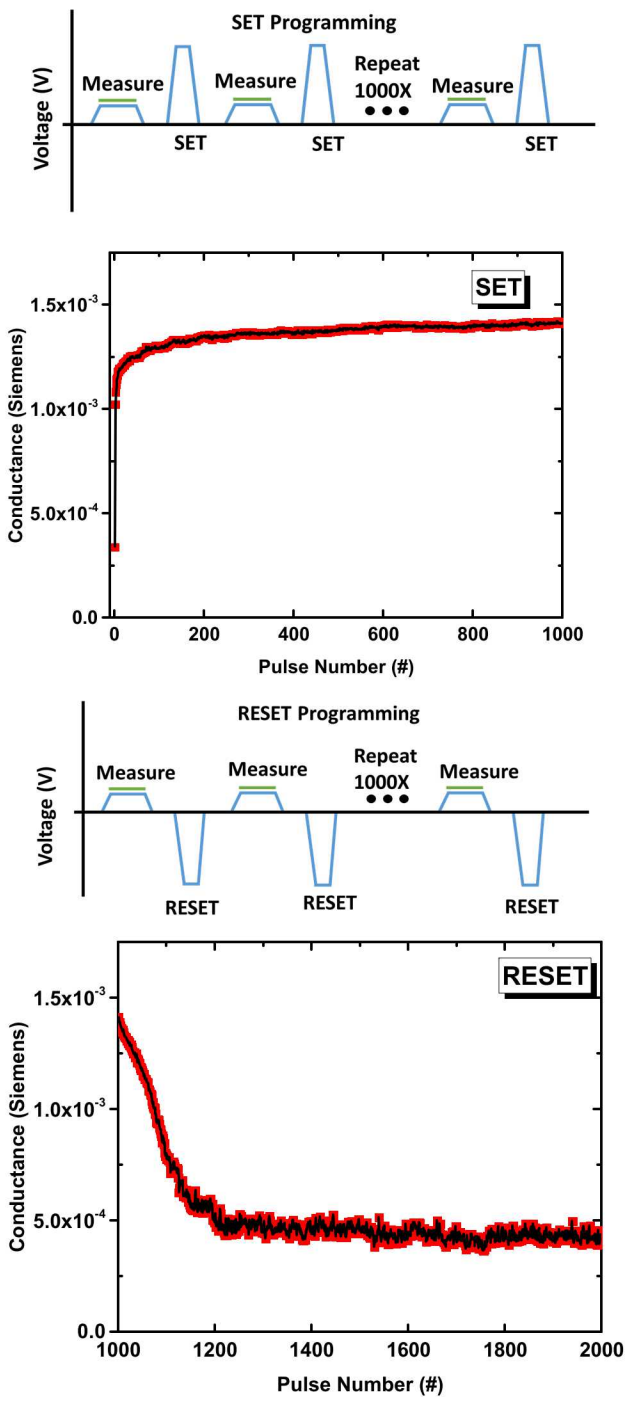

Figure 3: Example of the analog programming pulses used to tune the conductance of the device. First, a measurement pulse is applied to the device, with a measurement during the peak of the pulse. Then a voltage update is applied to the device, which changes the device conductance. This programming pulse and measure cycle is applied 1000 times for SET and 1000 times for RESET with opposite polarity. Note the conductance change is dependent on the state of the device.

to train on new data as well as its ability to perform image recognition tasks after training.

In the following, we investigate the effects of radiation on the classification accuracy of an analog ReRAM based neuromorphic computing accelerator capable of training for the first time. Results show that neuromorphic computing accelerators based on TaOx ReRAM devices could potentially operate in extremely high radiation environments.

II. Experimental Details

TaO_x based ReRAM devices were fabricated in the Sandia National Laboratories' Silicon Microfab facility on 6" wafers. The switching stack composed of TiN-TaO_x-Ta-TiN layers, is deposited using reactive sputtering, which employs a feedback technique described in [17]. The reduced TaO_x layer was 15 nm thick and the Ta layer was 15 nm thick, the active region had 1.0 µm X 1.0 µm lateral dimensions. All measurements were made on a Cascade Microtech manual probe station with an Agilent B1500A Semiconductor Parameter Analyzer mainframe equipped with a B1530 Waveform Generator Fast Measurement Unit. GGB Industries Picoprobe model 40A-GS-250 ground signal probes were used to reducing ringing during SET and RESET pulses. Devices were formed using a standard I/V sweep, as seen in Fig. 1, but with a maximum voltage of 3 V and a 5 uA compliance. Fig. 3 shows the analog programming scheme used to set the device conductance to a value such that it corresponds to a matrix weight. All updates must be the same, as we cannot know the device state prior to update without doing a serial read of all devices in the matrix. A series of

identical 1000 measure and update pulses are applied to the device, and the conductance response is plotted. For devices without a 4 k ohm resistor in line, SET was +1 V and RESET was -1.1 V. For devices with a 4k ohm resistor in-line SET was +1.5 V and RESET was -2.17 V. All measurement pulses were +100 mV and current through the device was measured at the top of the pulse using the B1500 RSU to extract resistance and conductance. The devices were irradiated using the nuclear microbeam on Sandia National Laboratories' Tandem particle accelerator. A 2.5 MeV Si ion broad beam was directed at the device array and measurements were made ex-situ. Results were collected in a lookup table for processing in CrossSim, the Sandia National Laboratories' open-sourced platform for determining a device's applicability to neuromorphic computing [18]. CrossSim takes a statistical device conductance dataset of multiple SET and RESET cycles and converts it into a lookup table. It uses this data to simulate training a neural network with weights based on analog ReRAM cells. It is assumed that either the write pulse length or write voltage can be adjusted to scale the amount written. The MNIST small images dataset was used as a standard for the purposes of this paper.

III. RESULTS

To obtain a set of data which will simulate training of devices in analog mode, 20 SET and RESET cycles were performed as shown in Fig. 4. The device was then irradiated with 2.5 MeV Si ions to a fluence of $1x10^{10}$ ions/cm^2, followed by an I/V sweep and a repetition of the cycling pulses. Another shot was then performed at an order of magnitude increased fluence. Note that there is a subtle change in the analog cycling behavior of the device after the shot with fluence of 10^{14} ions / cm^2. Device resistance was extracted using Ohm's law from 100 mV

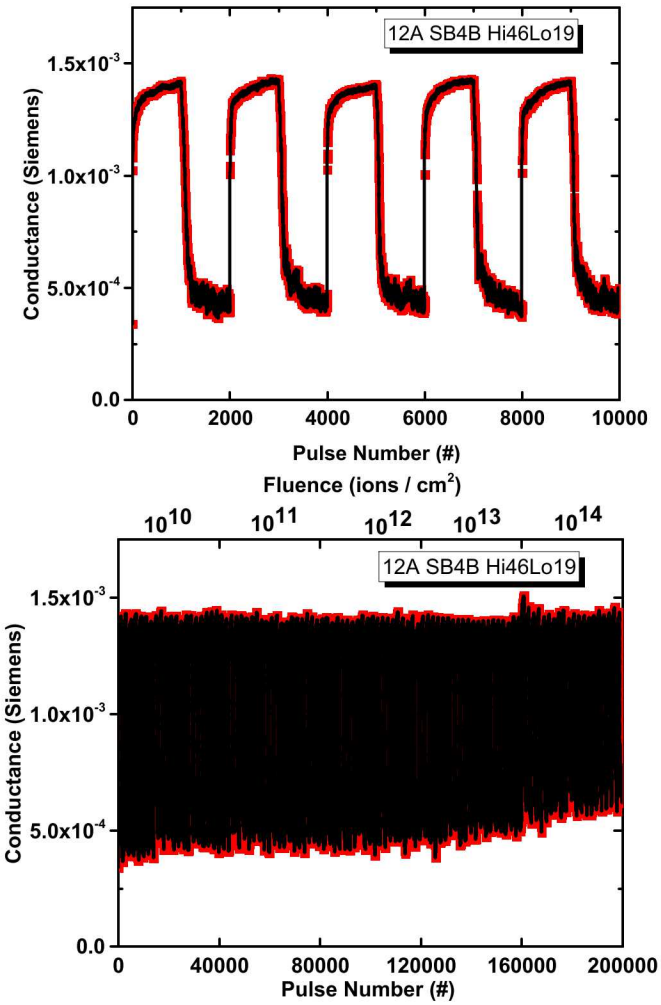

Figure 4: Operation of a TaO_x ReRAM device to simulate neural network training. Top: Example of cycled conductance response for one of the devices in the study during analog operation. Bottom: Operation of the same device with ion shots every 40,000 updates. Note that there is a subtle change in the analog conductivity response after the 10^{14} fluence.

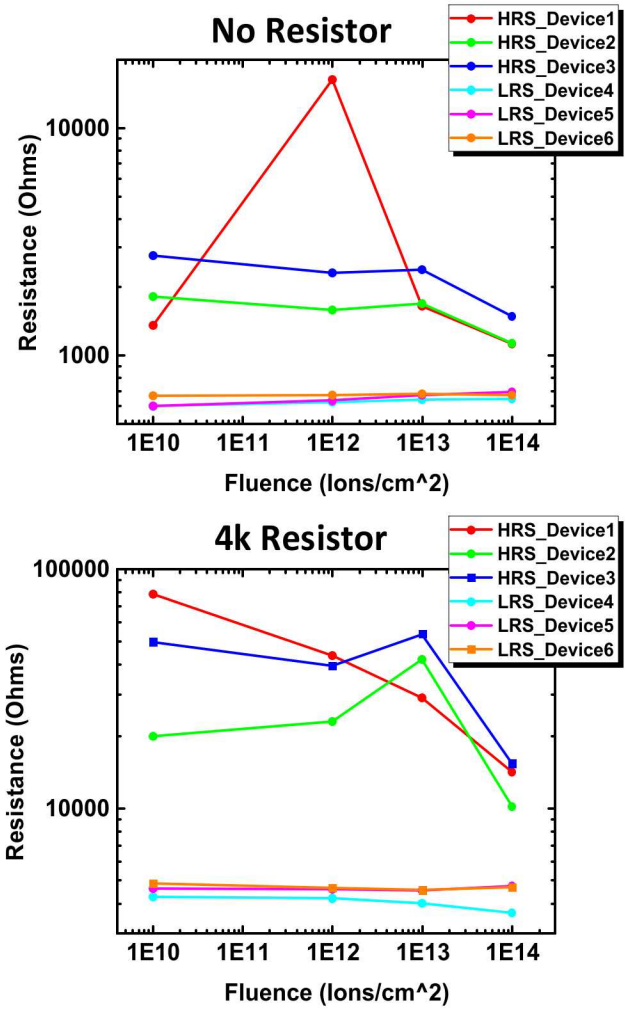

Figure 5: Device resistance immediately following irradiation, extracted by Ohm's law from 100 mV I/V sweeps. HRS indicates that the device was in the high resistance state during the heavy ion exposure, and LRS indicates that the device was in the low resistance state

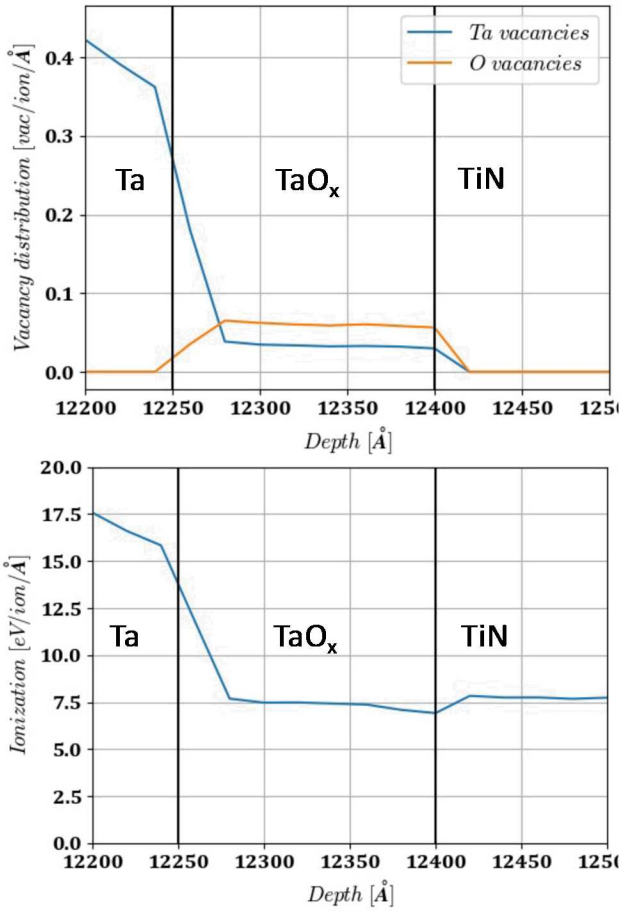

Figure 6: SRIM calculations for displacement and ionization for the TaO_x devices under 2.5 MeV Si ion irradiation. Top: vacancy production vs. depth. Total vacancy production in the TaO_x film is $5.2x10^{20}$ tantalum vacancies and $5.7x10^{20}$ oxygen vacancies. Bottom: Ionization vs. depth. The average ionization in the TaO_x layer given by SRIM calculations, is 7.5 eV/ion/Å. This corresponds to 208 MRad for $1x10^{14}$ ions / cm^2.

I/V sweeps. Results show a significant drop in resistance after the 10^{14} shot for all six HRS devices in the study.

Fig. 5 shows the device resistance extracted from 100 mV I/V sweeps performed immediately after each radiation shot. After exposure to 10^{14} ions/cm^2 all devices which were in the HRS show a resistance drop. Devices in the LRS state show a significant drop in resistance state at 10^{14} ions/cm^2 fluence. Some noise in the data and an outlying datapoint may be observed in the HRS devices which is typical of filamentary ReRAM devices. Devices in the graph on the top do not have a 4k in line resistor whereas devices on the bottom do. It was initially hypothesized that increasing the line resistance of the top TiN line to 4K ohms would increase radiation tolerance. Fig. 5 shows that there was not a significant difference between the devices with the resistor and those without the resistor.

The resistance with respect to fluence results in Fig. 5 are comparable to our earlier result on binary ReRAM devices where the onset of resistance change occurred near similar thresholds of calculated oxygen vacancy concentrations [19]. The devices in this study have a thicker set of metals, both above and below the switching region, than those devices tested in our previous digital studies. Shots conducted while the devices were in the full high resistance state (HRS) showed consistent degradation only at extremely high fluence (10^{14} ions / cm^2), whereas the devices in the low resistance state (LRS) showed no degradation in resistance up to the highest ion fluence tested. This is, again, consistent with previous studies on digital devices which have indicated that these devices are highly radiation tolerant [5, 6] [12-14]. Our results are also consistent with the mechanism in which the result of heavy ions is oxygen vacancy production. In the LRS state of the device, an oxygen vacancy rich filament creates a conductive channel through the device. As heavy ion irradiation creates additional oxygen vacancies, the device, which already has a low resistance channel, does not change resistance. In the high resistance state, however, a critical region in the filament has become largely oxidized, and is no longer conductive. When heavy ion irradiation creates additional vacancies within this critical region, the device resistance drops. It is worth noting that, in addition to the high activation energy required to break an O-Ta bond and displace the oxygen atom enough such that the bond does not immediately reform, the tiny scale of the critical region may well contribute to the radiation tolerance of these devices.

Calculations using the Stopping and Range of Ions in Matter (SRIM) application estimate the vacancy production in the active region of the device to be $5.2x10^{20}$ tantalum vacancies and $5.7x10^{20}$ oxygen vacancies [20]. A plot of the vacancy distribution with respect to depth is given in Fig. 6. These calculations are only an order of magnitude estimate, due to potential device stack thickness variation which is not represented properly in the simulation. Compound densities were calculated based on stoichiometry, and SRIM default displacement threshold energies were used. The average ionization in the TaO_x layer given by SRIM calculations, was 7.5 eV/ion/Å. This corresponds to 208 MRad for $1x10^{14}$ ions / cm^2. A plot of the average ionization with respect to depth is also given in Fig. 6. The exceptionally high radiation tolerance is likely due to the small volume of the critical area, which can be disrupted within the TaO_x layer. While there are some unknown aspects to the conductance switching mechanism in these devices, it has been demonstrated that conductance switching occurs within a nanoscale tantalum rich filament which is created in the FORMING process [13]. The switch from LRS to

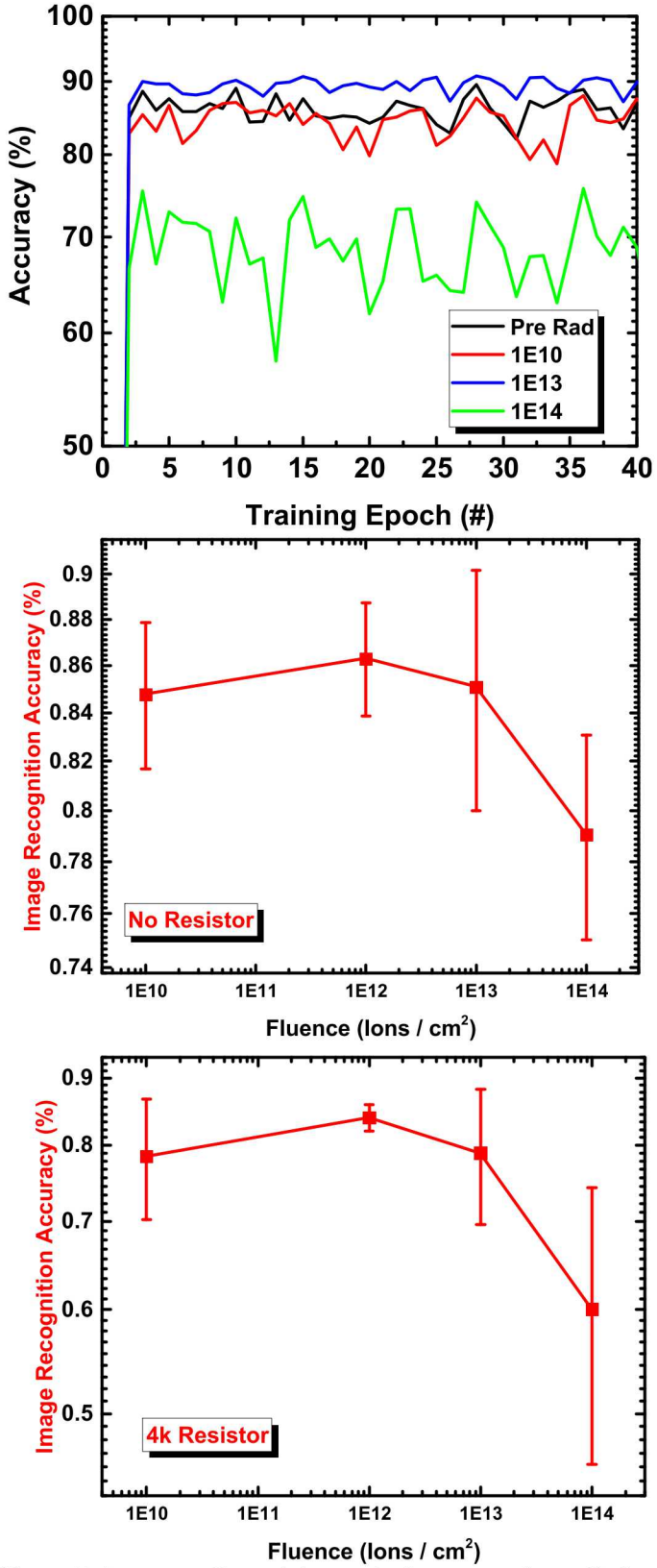

Figure 7: Accuracy after training on devices exposed to radiation in the high resistance state (HRS). Top: Image recognition accuracy after each training update for an example device. Middle: Accuracy after training for three devices taken from the maximum classification accuracy of all epochs for devices without the in-line 4k resistor. Bottom: Accuracy after training for three devices taken from the maximum classification accuracy of all epochs for devices with an in-line 4k resistor.

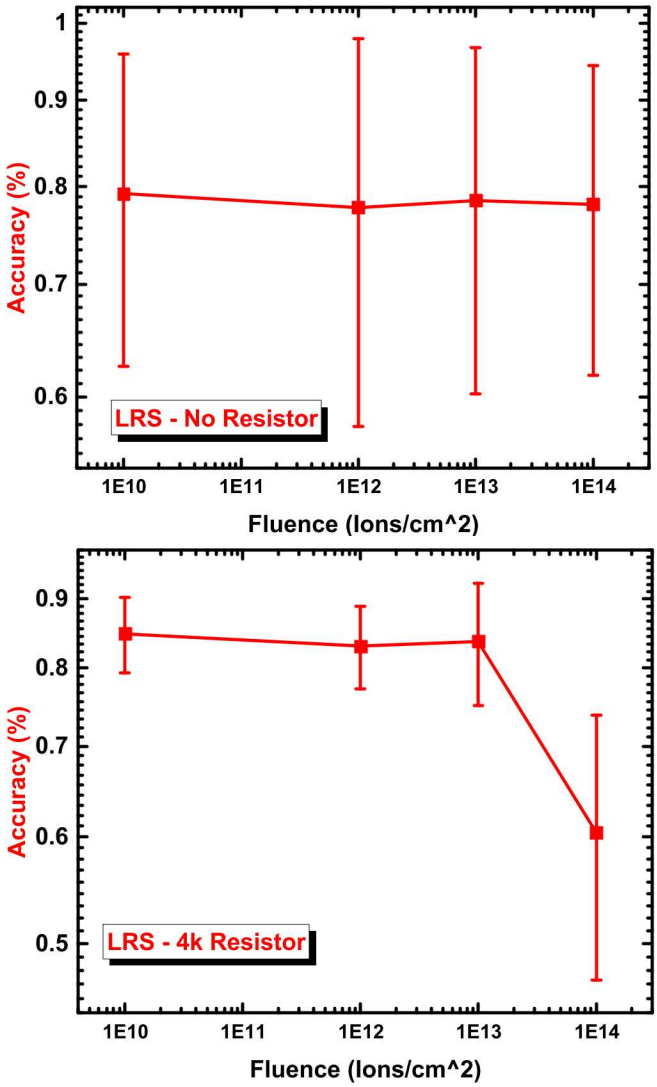

Figure 8: Accuracy after training on devices exposed to radiation in the Low Resistance State (LRS). The devices without a in-line resistor showed no significant change at any of the tested radiation exposure levels. The devices with an in-line 4k resistor tested in LRS state showed a significant decrease in accuracy at 10^{14} ions/cm^2 fluence.

HRS by recombination of some of the Ta and oxygen ions is hypothesized to not occur across the entire length of this filament during RESET [13]. The resistance of the device will only change if damage occurs within a small critical portion of the filament, which is already nanoscale in all three dimensions [13].

Sandia's CrossSim code was used to evaluate the effect of radiation on training an analog ReRAM crossbar [17], following the procedure in [3]. A standardized dataset of handwritten digit images, known as MNIST, was used to train the network [21]. CrossSim created a dataset from the cycles following each shot and was used to simulate a plot of classification accuracy (correctly identifying the handwritten digit) versus the neural network iteration (epoch) for a sample device (Fig. 7). While the device

largely continues to switch after the shot with 10^{14} ions / cm^2 fluence, subtle changes to the analog characteristics result in a lower average accuracy after training than in the non-irradiated devices. Fig. 8 shows the devices which were irradiated in the LRS state. While the resistance plots in Fig. 5 would seem to indicate that no significant change occurred to the devices in LRS the accuracy after training on devices with an inline 4k resistor did show an average accuracy drop after being irradiated at a fluence of 10^{14} ions / cm^2. This indicates that resistance changes are not perfectly correlated with devices' analog/neuromorphic performance after irradiation damage. In all cases, the radiation level required to show an effect is very high. The very high fluence levels required to impact the accuracy after training of a neuromorphic accelerator based on TaO_x RRAM devices implies applicability of this technology to embedded systems in high radiation environments.

IV. Conclusions

The effect of displacement damage on neural network accuracy after training on an analog-crossbar based on TaO_x ReRAM array is investigated. It is found that a neural network training accelerator based on this technology is robust to ion displacement damage, requiring a shot of 2.5 MeV Si ions with fluence of 10^{14} ions / cm^2 to show a significant effect on the classification accuracy after training. At this dose, the device conductance is found to shift. The devices were still functionally switching, and although subtle changes to their analog conductance response were observed, neural training was still possible. SRIM calculations confirm that displacement damage is occurring in the oxide. This study provides a promising direction for embedded pattern recognition systems for high radiation environments.

References

[1] R. B. Jacobs-Gedrim, S. Agarwal, K. E. Knisely, J. E. Stevens, M. S. v. Heukelom, D. R. Hughart, *et al.*, "Impact of Linearity and Write Noise of Analog Resistive Memory Devices in a Neural Algorithm Accelerator," in *2017 IEEE International Conference on Rebooting Computing (ICRC)*, 2017, pp. 1-10.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature,* vol. 521, p. 436, 2015.

[3] M. J. Marinella, S. Agarwal, A. Hsia, I. Richter, R. Jacobs-Gedrim, J. Niroula, *et al.*, "Multiscale Co-Design Analysis of Energy, Latency, Area, and Accuracy of a ReRAM Analog Neural Training Accelerator," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* 2018.

[4] S. Agarwal, T.-T. Quach, O. Parekh, A. H. Hsia, E. P. DeBenedictis, C. D. James, *et al.*, "Energy Scaling Advantages of Resistive Memory Crossbar Based Computation and Its Application to Sparse Coding," *Frontiers in Neuroscience,* vol. 9, 2016-January-06 2016.

[5] M. J. Marinella, S. M. Dalton, P. R. Mickel, P. E. D. Dodd, M. R. Shaneyfelt, E. Bielejec, *et al.*, "Initial Assessment of the Effects of Radiation on the Electrical Characteristics of TaO_x Memristive Memories," *IEEE Transactions on Nuclear Science,* vol. 59, pp. 2987-2994, 2012.

[6] D. R. Hughart, S. M. Dalton, P. R. Mickel, P. E. Dodd, M. R. Shaneyfelt, E. Bielejec, *et al.*, "Total ionizing dose and displacement damage effects on TaO_x memristive memories," in *2013 IEEE Aerospace Conference*, 2013, pp. 1-10.

[7] W. M. Tong, J. J. Yang, P. J. Kuekes, D. R. Stewart, R. S. Williams, E. DeIonno, *et al.*, "Radiation Hardness of TiO_2 Memristive Junctions," *IEEE Transactions on Nuclear Science,* vol. 57, pp. 1640-1643, 2010.

[8] E. DeIonno, M. D. Looper, J. V. Osborn, H. J. Barnaby, and W. M. Tong, "Radiation effects studies on thin film TiO_2 memristor devices," in *2013 IEEE Aerospace Conference*, 2013, pp. 1-8.

[9] Y. Wang, H. Lv, W. Wang, Q. Liu, S. Long, Q. Wang, *et al.*, "Highly Stable Radiation-Hardened Resistive-Switching Memory," *IEEE Electron Device Letters,* vol. 31, pp. 1470-1472, 2010.

[10] X. He, W. Wang, B. Butcher, S. Tanachutiwat, and R. E. Geer, "Superior TID Hardness in TiN/HfO_2/TiN ReRAMs After Proton Radiation," *IEEE Transactions on Nuclear Science,* vol. 59, pp. 2550-2555, 2012.

[11] J. S. Bi, Z. S. Han, E. X. Zhang, M. W. McCurdy, R. A. Reed, R. D. Schrimpf, *et al.*, "The Impact of X-Ray and Proton Irradiation on HfO_2 Based Bipolar Resistive Memories," *IEEE Transactions on Nuclear Science,* vol. 60, pp. 4540-4546, 2013.

[12] P. R. Mickel, A. J. Lohn, B. J. Choi, J. J. Yang, M.-X. Zhang, M. J. Marinella, *et al.*, "A physical model of switching dynamics in tantalum oxide memristive devices," *Applied Physics Letters,* vol. 102, p. 223502, 2013.

[13] M. Feng, S. J. Paul, Y. J. Joshua, Z. Min-Xian, G. Ilan, T. A. C., *et al.*, "Anatomy of a Nanoscale Conduction Channel Reveals the Mechanism of a High-Performance Memristor," *Advanced Materials,* vol. 23, pp. 5633-5640, 2011.

[14] J. J. Yang, D. B. Strukov, and D. R. Stewart, "Memristive devices for computing," *Nature Nanotechnology,* vol. 8, p. 13, 12/27/online 2012.

[15] S. Agarwal, S. J. Plimpton, D. R. Hughart, A. H. Hsia, I. Richter, J. A. Cox, *et al.*, "Resistive memory device requirements for a neural algorithm accelerator," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 929-938.

[16] J. L. Taggart, W. Chen, Y. Gonzalez-Velo, H. J. Barnaby, K. Holbert, and M. N. Kozicki, "Synaptic Programming of CBRAM in an Ionizing Radiation Environment," *IEEE Transactions on Nuclear Science,* vol. 65, pp. 192-199, 2018.

[17] J. E. Stevens, A. J. Lohn, S. A. Decker, B. L. Doyle, P. R. Mickel, and M. J. Marinella, "Reactive sputtering of substoichiometric Ta_2O_x for resistive memory applications," *Journal of Vacuum Science & Technology A,* vol. 32, p. 021501, 2014/03/01 2013.

[18] S. Agarwal, S. J. Plimpton, R. L. Schiek, I. Richter, A. H. Hsia, D. R. Hughart, *et al.* (2017). *CrossSim*. Available: http://cross-sim.sandia.gov

[19] D. R. Hughart, A. J. Lohn, P. R. Mickel, S. M. Dalton, P. E. Dodd, M. R. Shaneyfelt, *et al.*, "A Comparison of the Radiation Response of TaO_x and TiO_x Memristors," *IEEE Transactions on Nuclear Science,* vol. 60, pp. 4512-4519, 2013.

[20] J. F. Ziegler, "SRIM," ed, 2012.

[21] Y. LeCun. MNIST [Online]. Available: http://yann.lecun.com/exdb/mnist