Ira A. Fulton Schools of Engineering
Arizona State University


Hoyeol Bae


Wataru Tamaki


Benjamin Tung


Justin Weidmann


Zachary White


Dr. Marinella

Reducing accuracy loss from radiation in analog memristive chips for image recognition

Background

As artificial intelligence becomes increasingly common in commercial and industrial applications, the demand for energy-efficient and high-performance hardware has surged. This need for better hardware is due to the highly computational nature of AI algorithms, oftentimes needing to perform billions, trillions, or more operations per second. ASU's Sol Supercomputer, the one used to train the models in this project, can process up to 2.272 petaflops per second! [1] The current state-of-the-art in computational hardware uses digital circuits, which suffer from (relatively) high power consumption and (relatively) slow computation time

Memristive chips offer a promising solution by performing computations directly in memory, drastically reducing power usage and execution time. These chips use trapped electrons in SONOS cells to store values, which can then be "multiplied" with an applied voltage.

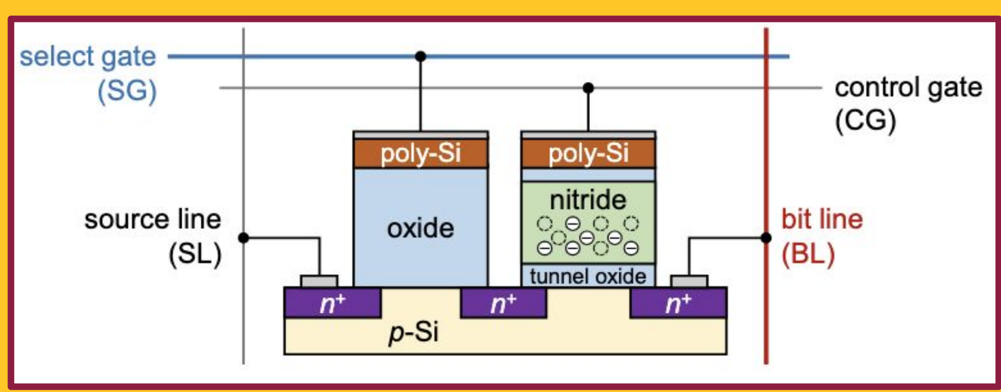


Figure 1: SONOS Cell Illustration [2]

Problem Statement

- The amount of charge in the cell directly represents a stored value in memory, and any small changes to this charge will be read as a different stored value. If this value changes, any multiplication done using that memory data will be inaccurate.
- SONOS cells are vulnerable to radiation, as electrons can be ejected and more will be lost over time.
- In space, radiation exposure on the scale of 0.1 mrad/s to 10 mrad/s [3]
- This charge loss, which we call "Conductance Drift", leads to inaccurate computations, directly decreasing the performance of machine learning models, such as image recognition.
- Addressing these issues is essential for deploying memristive technology in critical, power-sensitive applications like autonomous vehicles, aerospace systems, and other advanced AI-driven platforms.

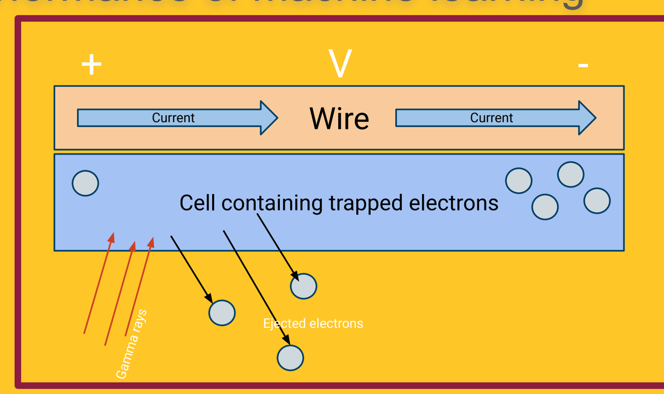


Figure 2: Radiation effects on charge cell (oversimplified)

Method 1 - Alpha Correction Factor

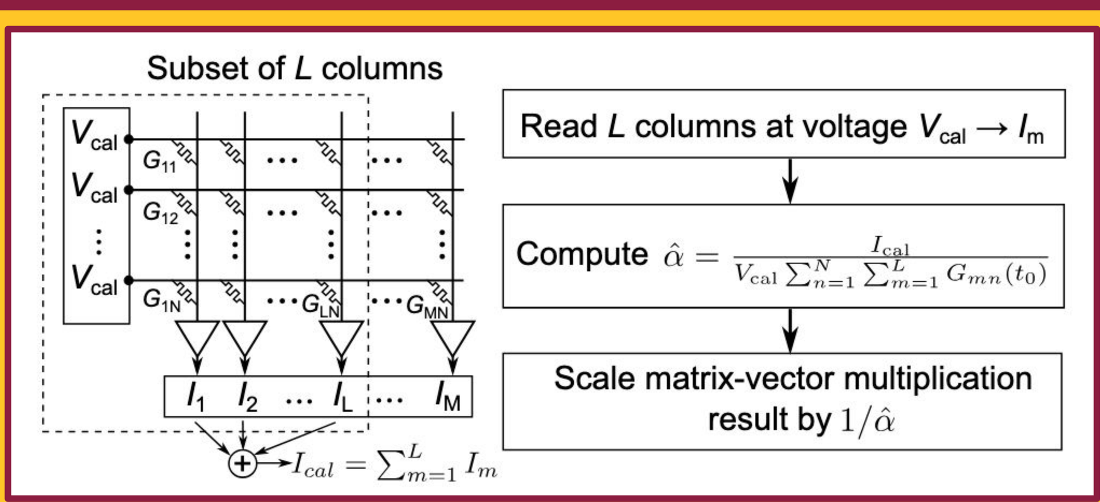


Figure 3: Alpha Correction Factor Computation [4]

- Multiply the output of the MVM by 1/α
- Denominator: Multiply a test V_cal matrix by the sum of the conductances
- Numerator: Sum the currents from multiplying V_cal by the irradiated SONOS chip

Results for Method 1 - Simulation

- Increased TID induces dot product errors
- Applying alpha greatly reduces the effect of drift

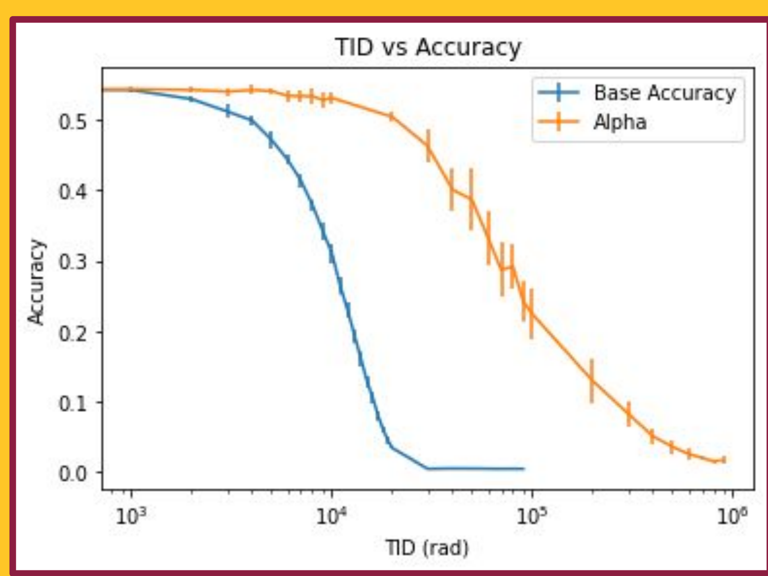
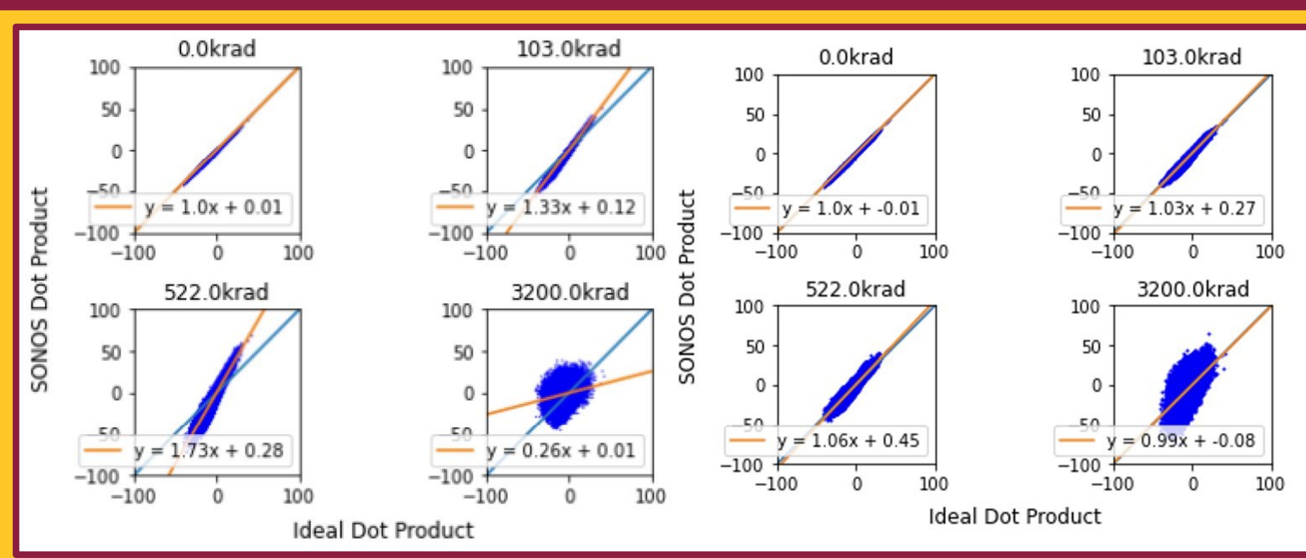


Figure 5: ResNet-32 Image Accuracy



No Alpha Alpha

Figure 4: Dot Product Errors

- ResNet-32 model image recognition accuracy
- Applying alpha to the model results in a marked increase in accuracy at higher TID levels

Method 2 - Training with Noise

- Injects noise into neural network during training
- Optimizes the model at a specific TID
- Tradeoff: Less accuracy at other TIDs

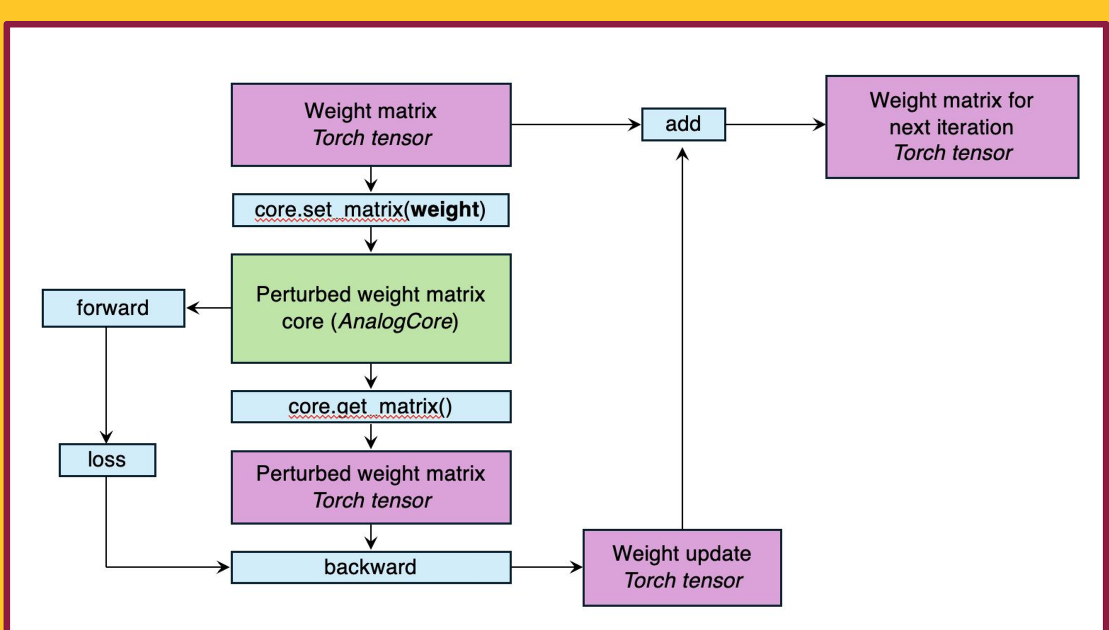


Figure 6: Model Training Loop

Results for Method 2 - Simulation

Three models shown:

- Base (0 TID, no radiation)
- 4k TID
- 10k TID

The simulation shows that the 10k model demonstrates clear improvements when used in a chip with higher radiation exposure, but poorer performance in a brand-new chip when compared to the base model.
The effect is less noticeable for smaller doses.

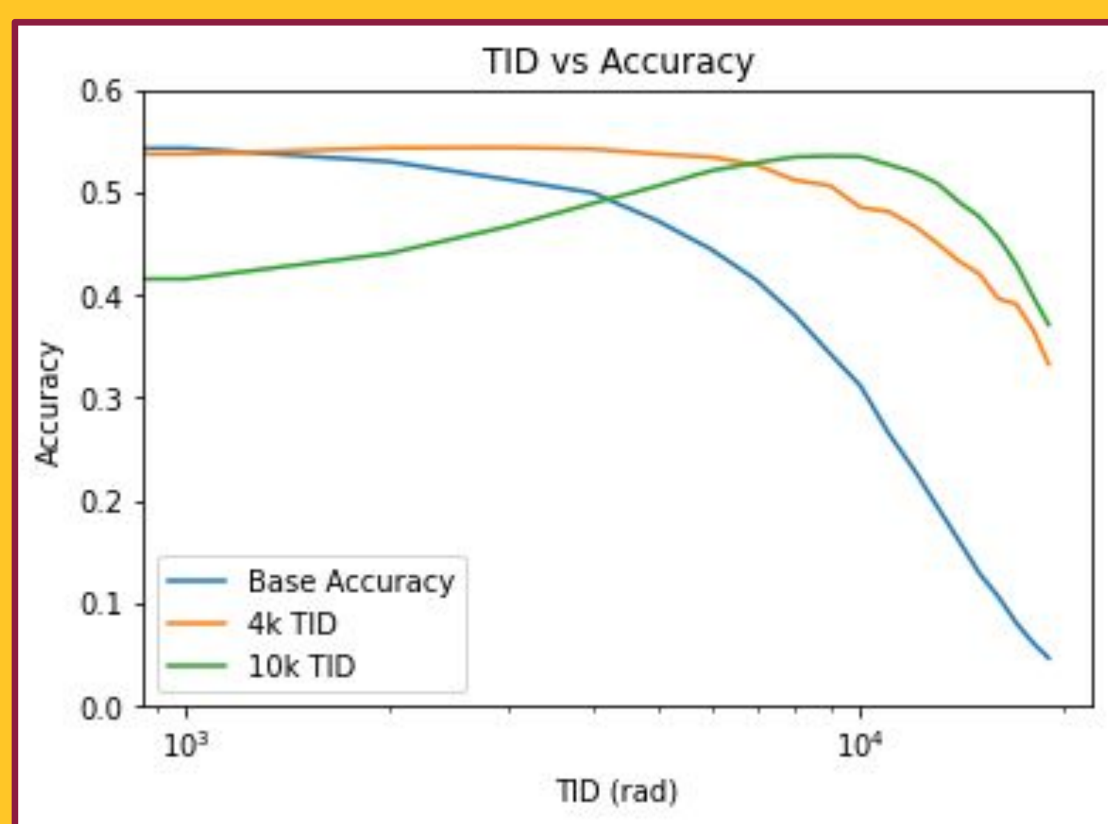


Figure 7: Different models accuracy at varying TID

Hardware

- Setup
  - The chip is placed in one of the sockets
  - The controller board connects PC and the chip/board
- Capabilities
  - Program the chip and read back the data
  - Calculate MVM
  - Read column currents

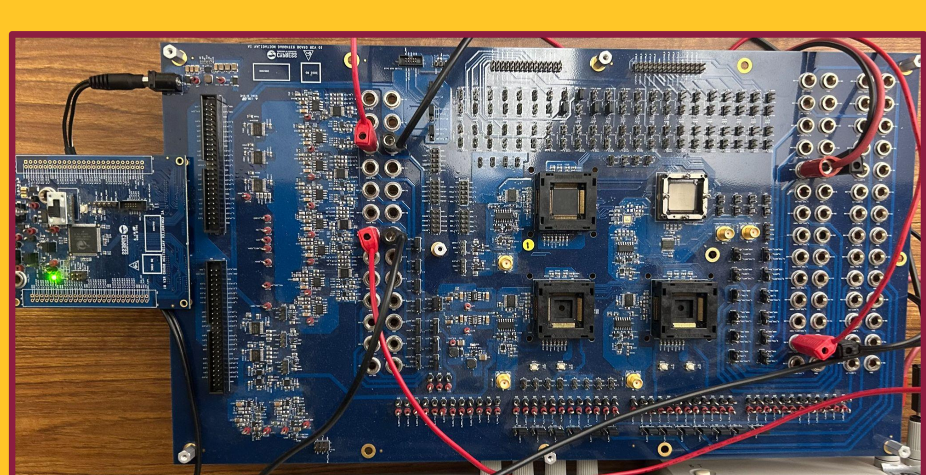


Figure 8: Memristive chip in PCB

Conclusion

- SONOS chips are highly efficient at performing machine learning, but radiation can cause conductance drift and reduce accuracy
- Methods such as Alpha Correction Factor and Training with Noise can help correct for that and greatly improve accuracy

References:

[1] Sophia Balasubramanian - ASU News: "ASU's Sol ranks among top-performing supercomputers globally"
[2] T. Patrick Xiao *et al.* "In-situ analog in-memory computing under ionizing radiation exposure,"
[3] Richard H. Maurer *et al.* "Harsh Environments: Space Radiation Environment, Effects, and Mitigation"
[4] Joshi Vinay *et al.* "Accurate deep neural network inference using computational phase-change memory."