ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

XXXV Cycle

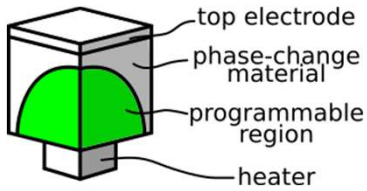
Non-conventional energy-efficient data processing architectures

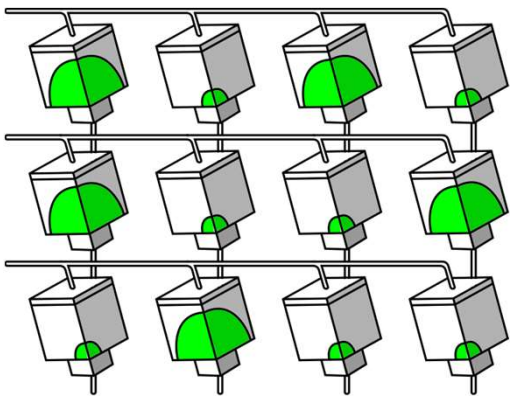
Carmine Paolino

Supervisor: Prof. Gianluca Setti

Research context and motivation

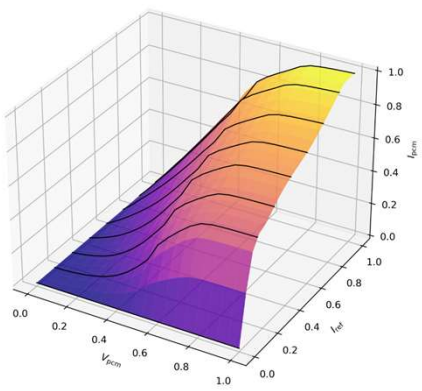
- **Phase Change Memories** (PCM) have found place as robust Non-Volatile Memories.Their successful integration in CMOS fabrication processes, their high throughput and read/write endurance make them suitable candidates for **Analog In-Memory Computing** applications.
- Neural networks are nowadays considered as the tool of choice for problems where large amounts of data is available and more traditional processing techniques have reached a performance plateau.
- Active research is carried out into the robustness of neural networks towards imperfect implementation, with a notable example being the quantization of weights and signals to realize compact models suitable also for hardware with limited resources.

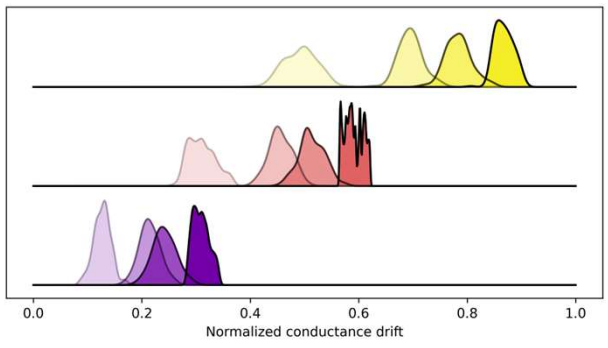




Addressed research questions/problems

- The **nonidealities** of PCM devices, whilst manageable in digital applications, heavily constrain the performance of analog ones, with limits imposed by the voltage-dependent conductance, programming error and conductance drift.
- In-Memory Computing applications must become **robust** towards these effects in order to make the technology usable. This, in turn, would specify an application-level requirement that technologists/designers could pursue in their implementations.





List of attended classes

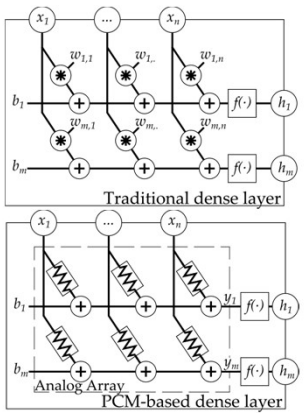
- 01QORRV Writing Scientific Papers in English (20/02/2020, 3 CFU)
- 01QWFBG Signal processing: methods and algorithms (24/04/2020, 6 CFU)
- 01UNRRV Entrepreneurship and start-up creation (03/07/2020, 8 CFU)
- 01QSXRU The measurement of electrical impedance (10/03/2021, 2 CFU)
- 01DOBRV Mathematical-physical theory of electromagnetism (06/06/2022, 3 CFU)
- 01UMNRV Advanced deep Learning (didattica di eccellenza) (15/06/2021, 6 CFU)
- 01UZHPQ Aspetti psicologici ed educativi dello sviluppo e dell'apprendimento (23/06/2022, 6 CFU)
- 02SIOPQ Didattica, tecnologie e ricerca educativa (18/07/2022, 6 CFU)

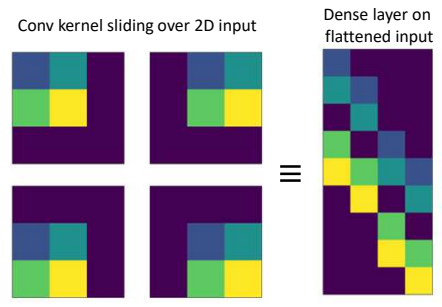
External:

- Summer School on Signal Processing (S3P-2019) @ IEEE SPS / EURASIP / ISIF – University of Genova (04/11/2019, 21 hours)
- Functional Analysis @ Scuola Matematica Universitaria 2020 - University of Perugia (07/08/2020, 18 hours)
- Mathematical Statistics @ Scuola Matematica Universitaria 2020 - University of Perugia (07/08/2020, 18 hours)

Novel contributions

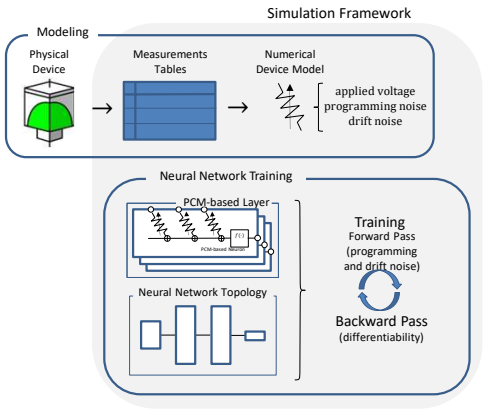
- (Previously) Proposed an iterative Compressed-Sensing decoding procedure to cope with the voltage dependency of PCM conductances
- Proposed a **general procedure to train neural networks on arbitrary synapse models**, but tested specifically with PCM models.
- Analyzed how quantization techniques can be adapted to the generation of robust networks against device-dependent programming noise and weight drift.
- Observed the **benefits of convolutional layers unrolling** to introduce redundancy and tolerance towards device nonlinearity.

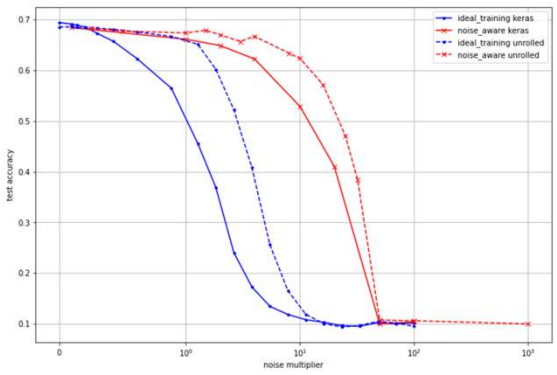




Adopted methodologies

- Iterative programming and characterization of a batch of PCM, with the subsequent characterization both in the I/V domain and in terms of conductance drift.
- Development of **numerical models** to describe the most critical device properties.
- Training of neural networks, employing these computational models as the core synaptic model, with additional techniques to make them robust against the observed nonidealities.





Future work

- Introduce energy consumption as an additional regularization function in the training process, leading to models whose energy footprint is lower.
- Validation of the proposed techniques on the actual devices, requiring updates on the test chip structure for meaningful observations.
- Application of the newly observed device properties on previously addressed problems, i.e. robust compressed sensing decoding.
- Verification of the proposed procedures on different NVM technologies to create a benchmark for the comparison of different devices

Submitted and published works

- Paolino C., et al, "*Practical Architecture for SAR-based ADCs with Embedded Compressed Sensing Capabilities*", PRIME 2019, Lausanne, Switzerland
- Paolino C., et al, "*A passive and low-complexity Compressed Sensing architecture based on a charge-redistribution SAR ADC*", Integration, vol. 75, November 2020, pp. 40-51
- Paolino C., et al, "*Asymptotic expressions of mismatch variance in interdigitated geometries*", ISCAS 2020, Seville, Spain
- Paolino C., et al, "*An architecture for ultra-low-voltage ultra-low-power compressed sensing-based acquisition systems*", NORCAS 2021, Oslo, Norway
- Paolino C., et al, "*Stability and mismatch robustness of a leakage current cancellation technique*", ISCAS 2021, Daegu, Korea
- Paolino C., et al, "*Compressed sensing by phase change memories: Coping with encoder non-linearities*", ISCAS 2021, Daegu, Korea
- Paolino C., et al, "*Phase-Change Memory in Neural Network Layers with Measurements-based Device Models*", ISCAS 2022, Austin, Texas

Politecnico di Torino 1859

PhD program in
Electrical, Electronics and Communications Engineering