Ferroelectric Hafnia Superlattices for Bio-Inspired Computing.

Laura Bégon-Lours*[1,2], Mattia Halter [1,2], Stefan Slesazeck[3], Donato Francesco Falcone[2], Bert Jan Offrein[2]

1. Integrated Systems Laboratory, ETH Zurich – D-ITET, Zurich, Switzerland
2. IBM Research Zurich, Rüschlikon, Switzerland.
3. Namlab Gmbh, Dresden, Gemany

*LBEGON@ethz.ch

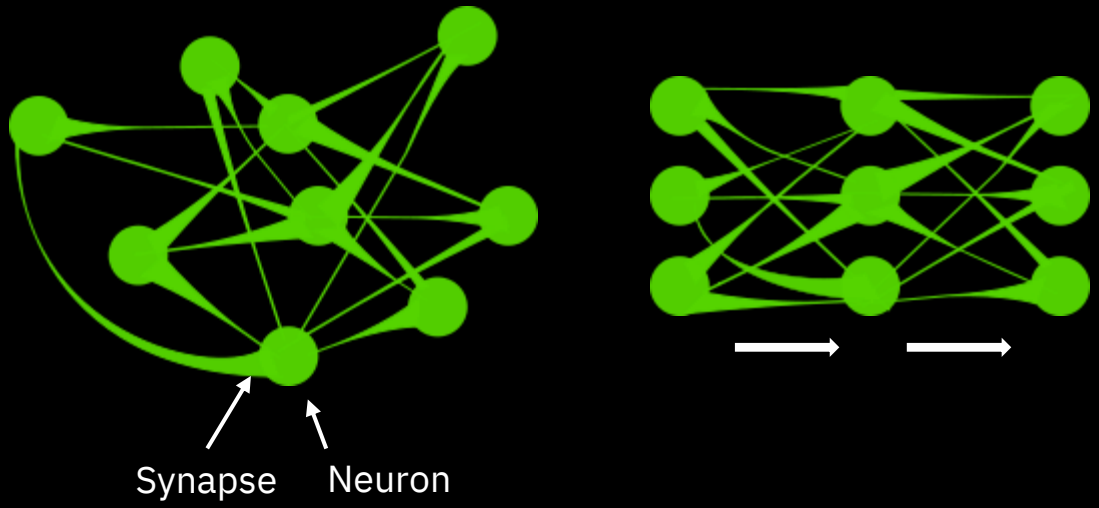
Acknowledgements:

- This work is funded by H2020 FREEMIND (840903), BeFerroSynaptic (871737), UNICO (ANR-19-CHR3-0006), ALMOND (SNF 198612)
- The authors acknowledge the *Binnig and Rohrer Nanotechnology Center* (*BRNC*)

1. Motivation:
    a. Analog resistive switching in $HfZrO_4$ devices for Artificial Neural Networks hardware
    b. HfO2 / ZrO2 superlattices for lower crystallization T.

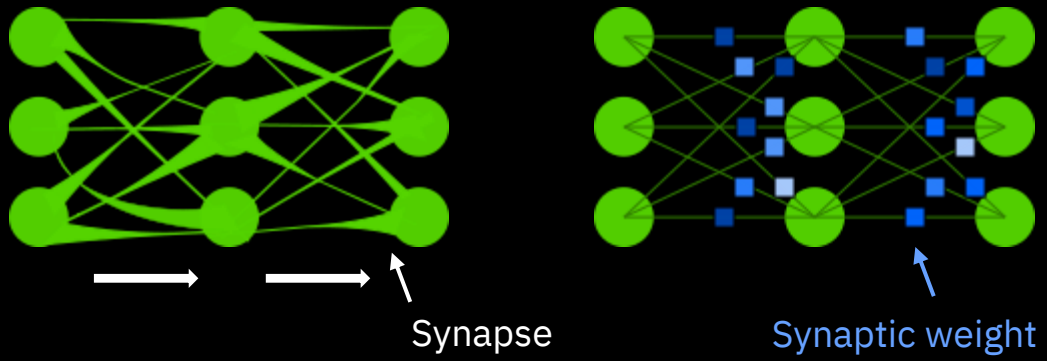
2. Results:
    a. 5x(0.5 nm HfO_2 / 0.5 nm ZrO_2) / WO_x // Si
    b. Co-integration of FTJ to CMOS in the BEOL

Bio Inspired Computing: Mimicking the brain, “Deep-Learning” algorithms are structured into layers of interconnected neurons.

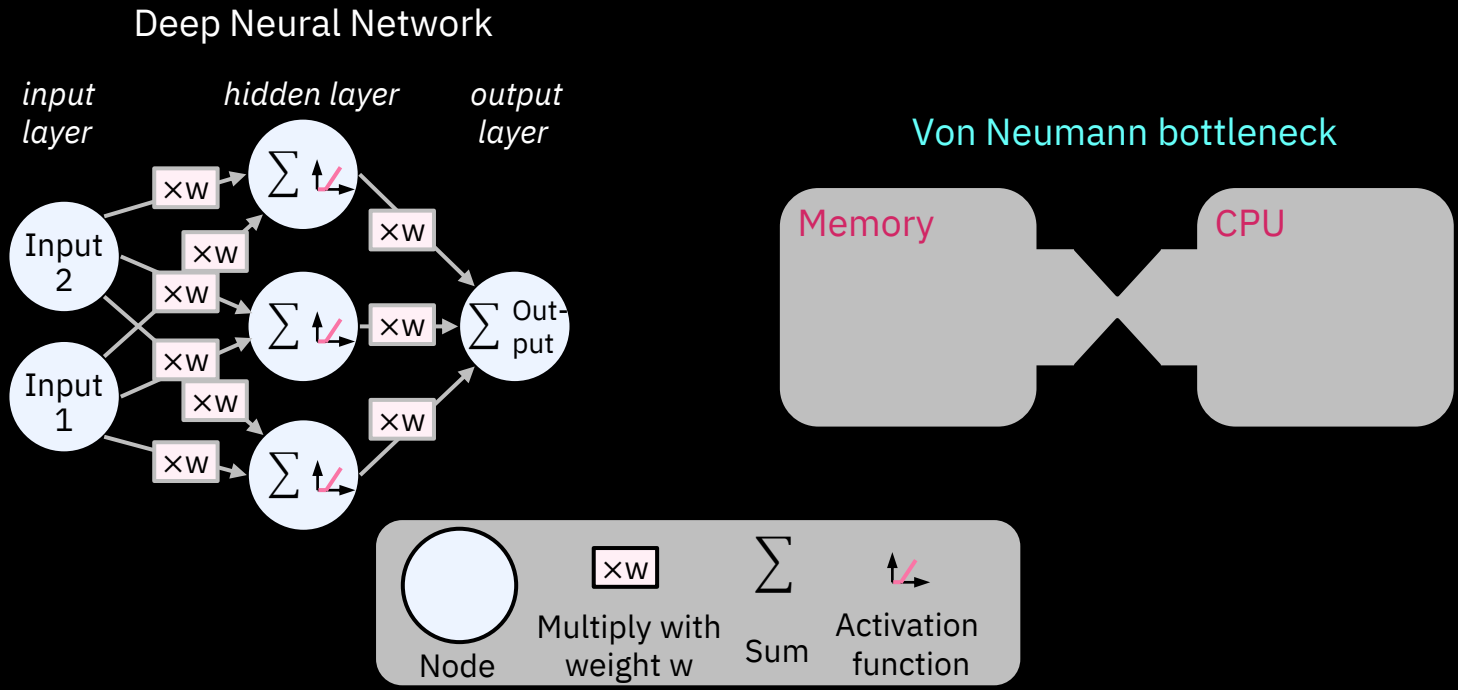


At each layer, a matrix vector multiplication is performed.
“Learning” is achieved by adjusting the matrix elements,
by analogy: the synaptic weights

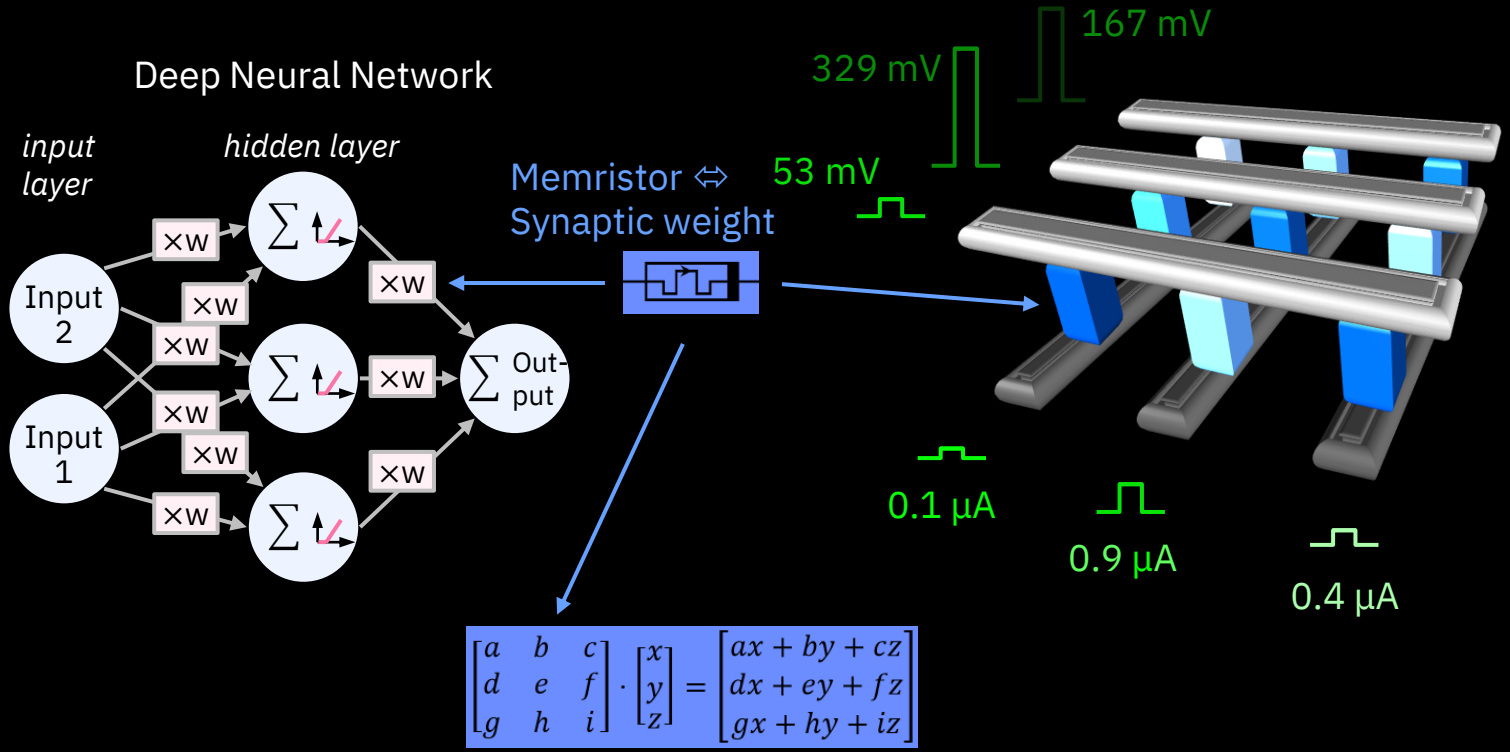
$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{bmatrix}$$



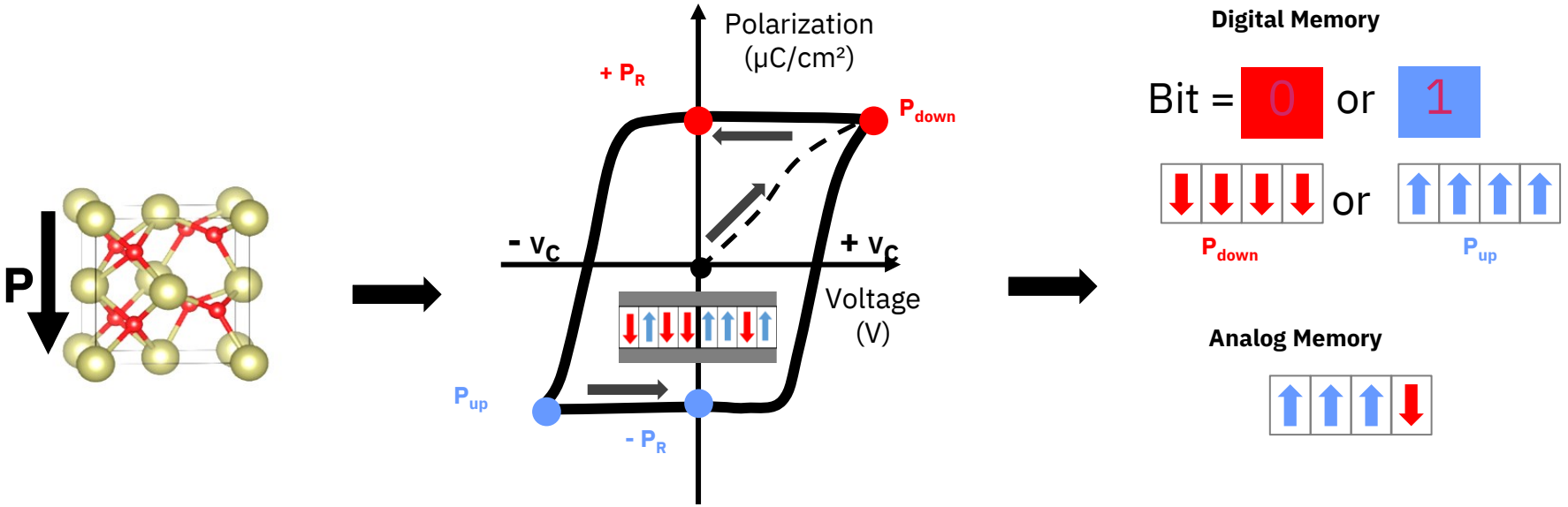
Artificial Neural Network: Solving the Von Neumann bottleneck with In-Memory Computing



Analog multiplication in an Artificial Neural Network: Parallel voltage drop through a cross-bar arrays of Memristors



Ferroelectricity



- **Ferroelectricity** → asymmetric orthorhombic cell→ dipoles (=Polarization)
- **Hysteresis** → Polarization switchable by an electric field
- **Coercive field, Ec** $\stackrel{\text{def}}{=}$ field to switch from one polarization state to the other
- **Remnant Polarization, Pr** $\stackrel{\text{def}}{=}$ polarization still available at null electric field

“Passive” crossbars arrays

Halter et al., Crossbar operation of BiFeO3/Ce–CaMnO3 ferroelectric tunnel junctions: From materials to integration, JMR 2023

Passive crossbar operation enabled by the sharp coercive field in $BiFeO_3$ ferroelectric.

Problem: difficult to industrialize (Pulsed Laser Deposition of single crystal thin films)

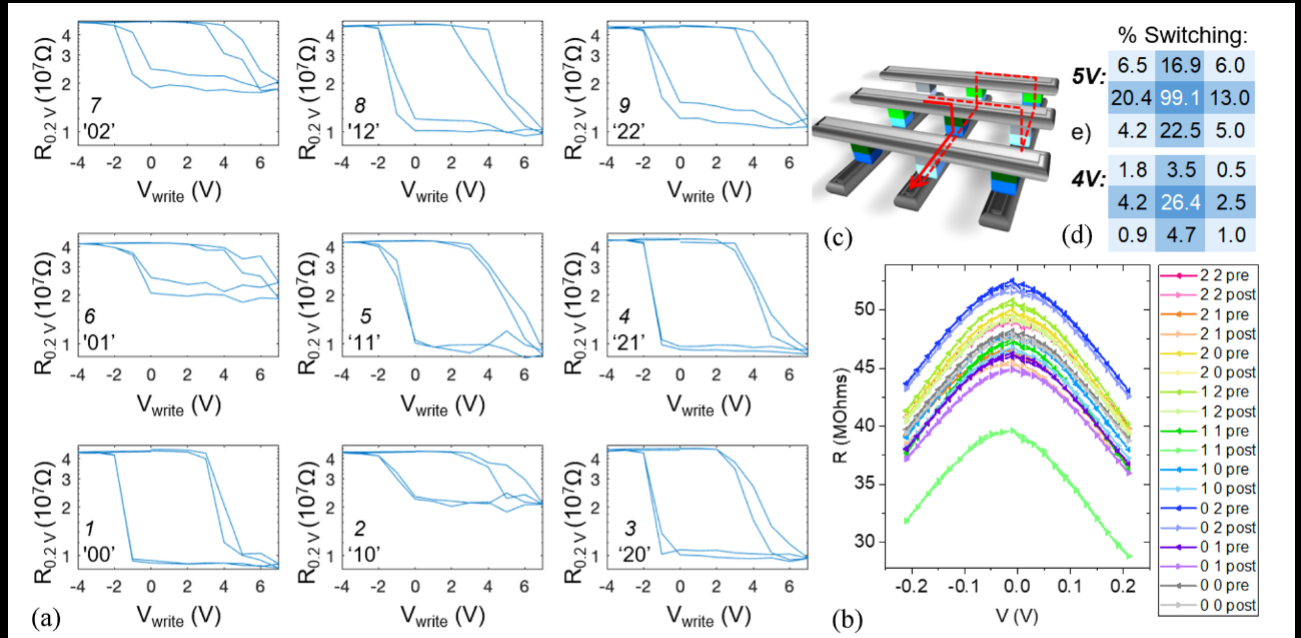
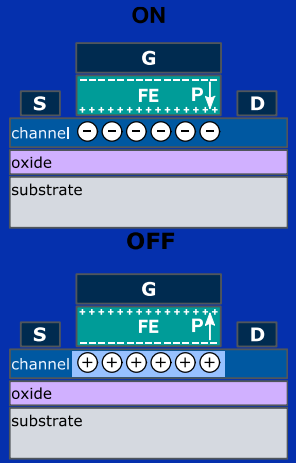


Figure 6: (a) Electro-resistance loops for nine FTJs in a 3×3 crossbar, numbered, ‘i, j’ with i, j in [0; 1; 2]. Consecutively: ‘00’, ‘10’, ‘20’, ‘21’, ‘11’, ‘01’, ‘02’, ‘12’, ‘22’. (b) R-V read sweeps before (-pre) and after (-post) applying 4 V to the element [1 1]: for this moderate programming voltage, the effect of sneak paths (c) is weak: as reported in the table in (d), neighboring elements have switched by $dR_{i,j}$ = 2–5% of the dynamic range measured in (a), and corner elements by 0–2%. (e) In contrast, for a programming voltage of 5 V to the central element [11], neighboring elements have switched by 13–22% of the dynamic range, and corner elements by 4–6%.

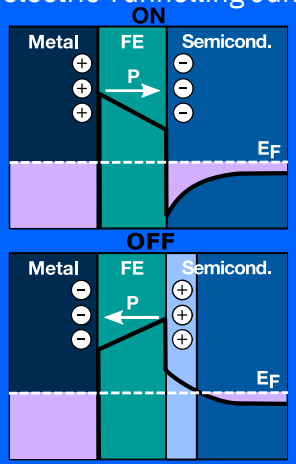
Invited Feature Paper
DOI:10.1557/s43578-023-01158-8

CMOS compatible memristive devices at ZRL: IBM-ETH Exploratory clean room fabrication capability ~ 950 m², class 100-10000

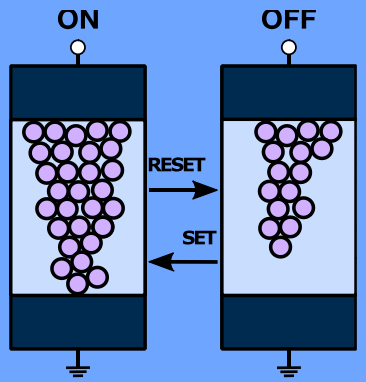
Ferroelectric Field Effect Transistor



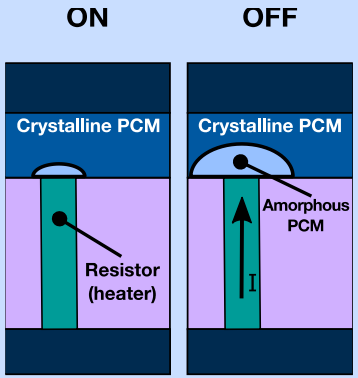
Ferroelectric Tunnelling Junction



Filamentary Resistive RAM



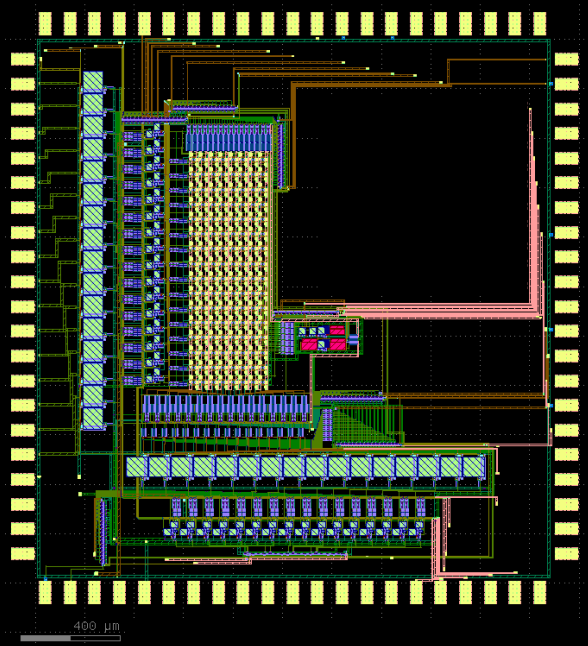
Phase Change Memory

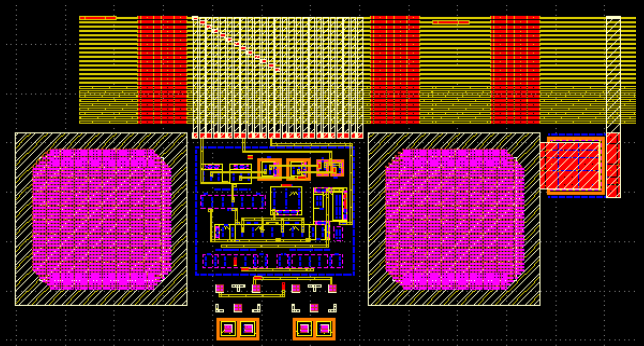


Neuromorphic chips:

UNICO: ANR-19-CHR3-0006

Voltage-Dependent Synaptic Plasticity (VDSP): Unsupervised probabilistic Hebbian plasticity rule based on neurons membrane potential
Nikhil Garg (CNRS, University of Sherbrooke)
DOI: 10.3389/fnins.2022.983950

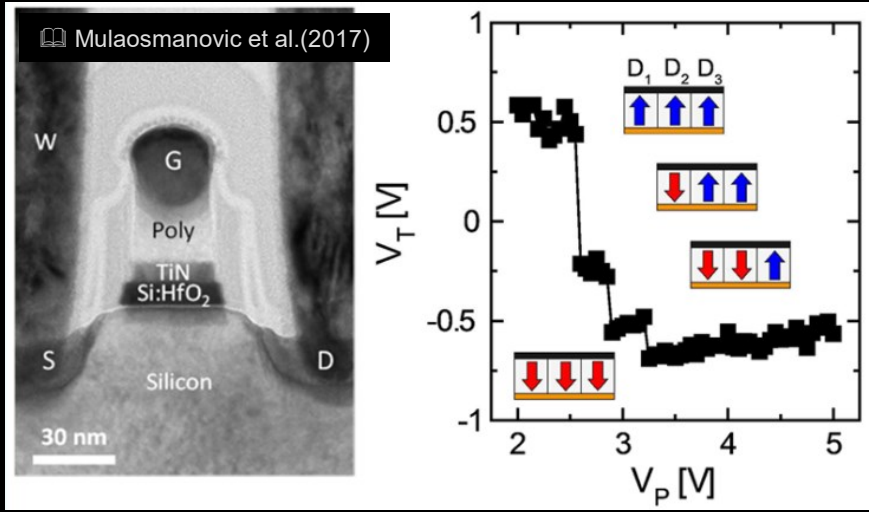




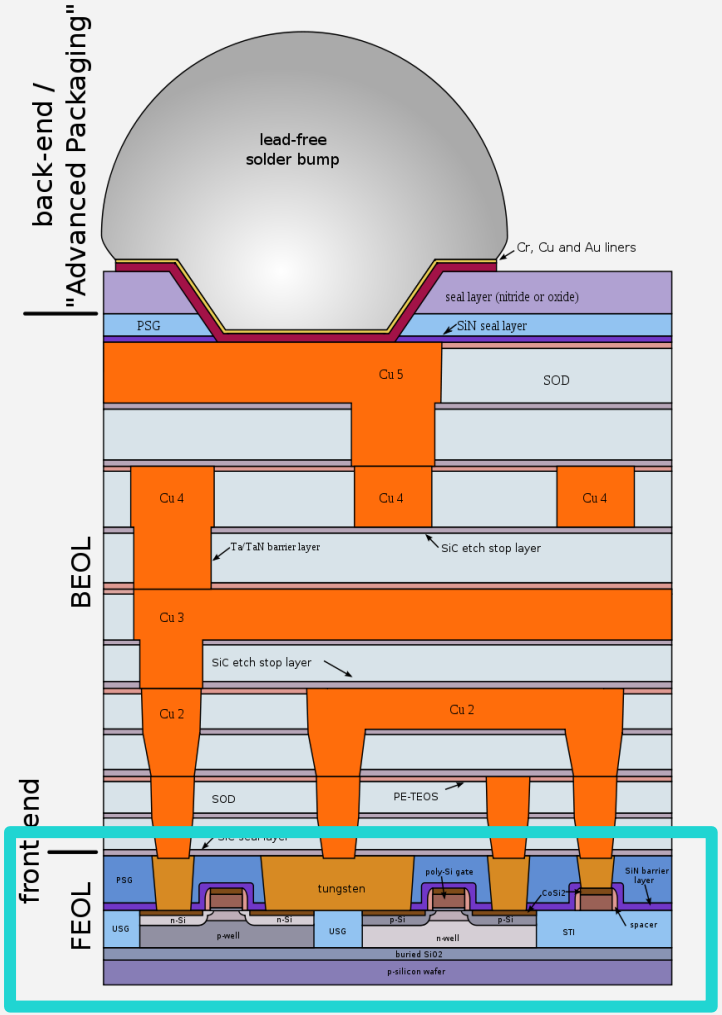
BeFerroSynaptic: H2020-ICT-2019-2

BEOL technology platform based on ferroelectric synaptic devices for advanced neuromorphic processors
Institute of Neuroinformatics (UZH), NaMLab, University of Groningen
https://beferrosynaptic.eu/publications

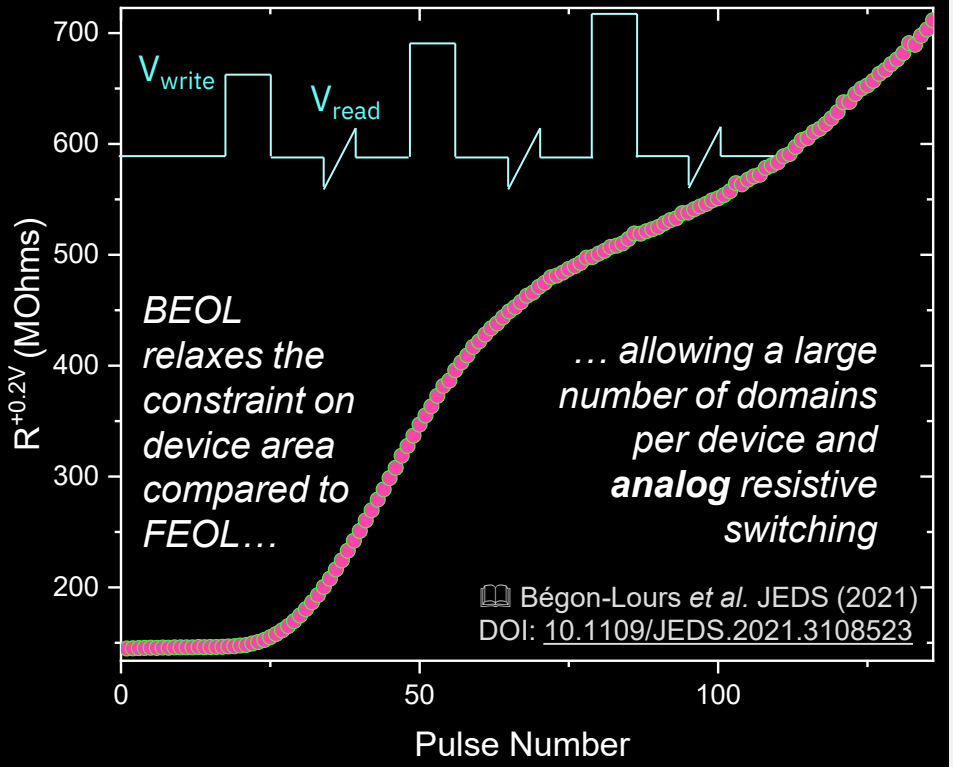
Integration of HfO_2 in the Front-End-Of-Line

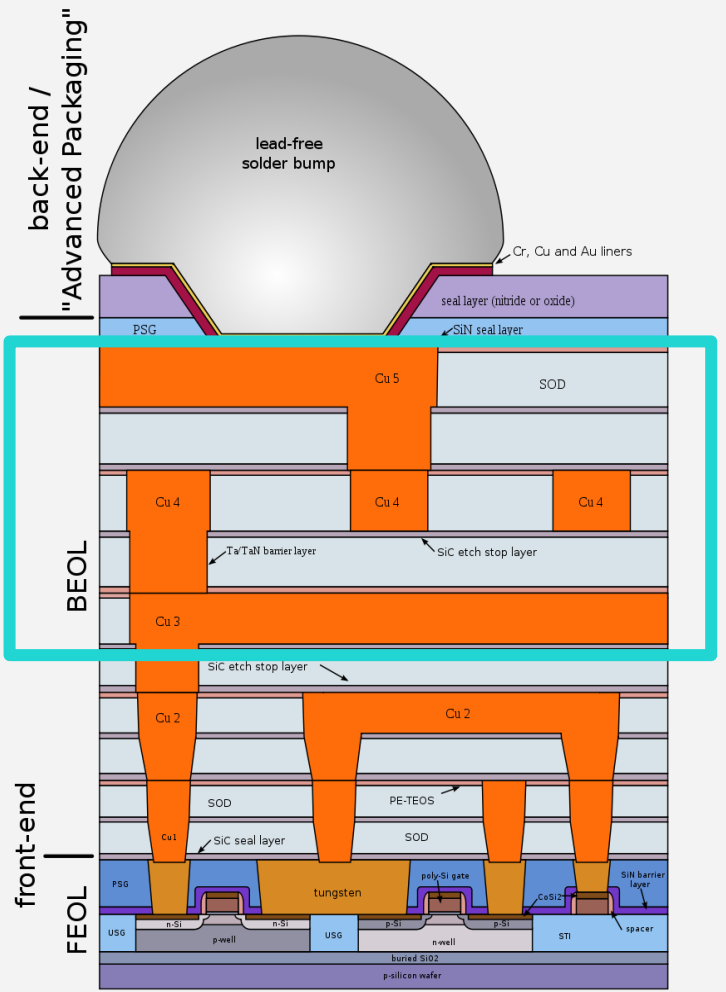


- Crystalline grains switch without sub-domains.
- In scaled devices, the number of grains is discrete



Integration of $HfZrO_4$ in the Back-End-Of-Line





Laura Bégon-Lours / 29.10.2024 / ETHZ & IBM Research / LBEGON@ethz.ch

Oxide Interfaces

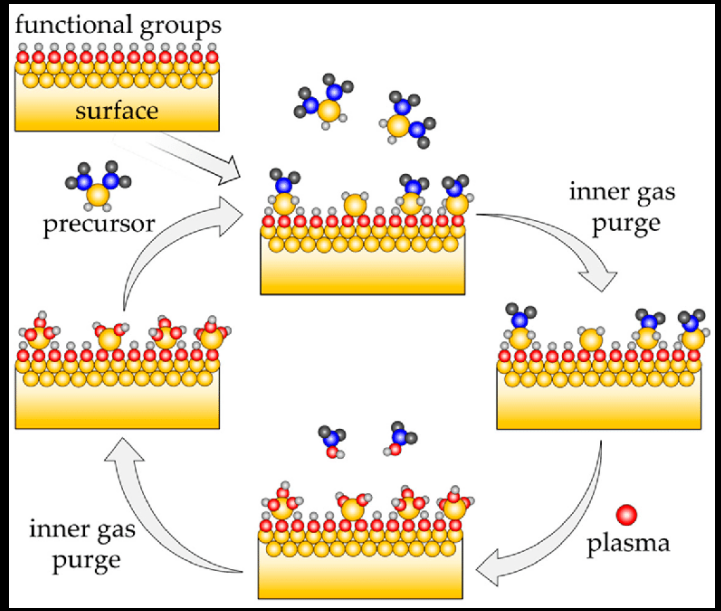
Example of the Ferroelectric Tunnel Junction

HZO deposition: **O_2 plasma** at 300C. $TiNO_x$ is expected: M/FE/DE/M stack.

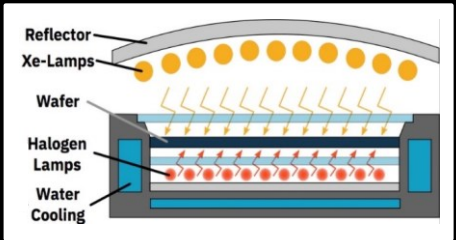
Charge-trapping in this layer might be detrimental to the **endurance**

$TiNO_x$ will also form, but **not** enhanced by O_2 plasma.

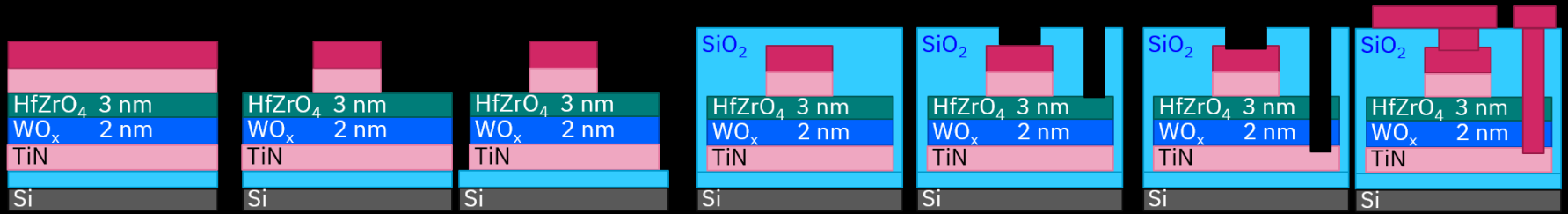
$M/FE/MO_x/M$ stack, no charge-trapping.



Ferroelectric tunnel junctions at IBM: $HfZrO_4/WO_x$ bilayers



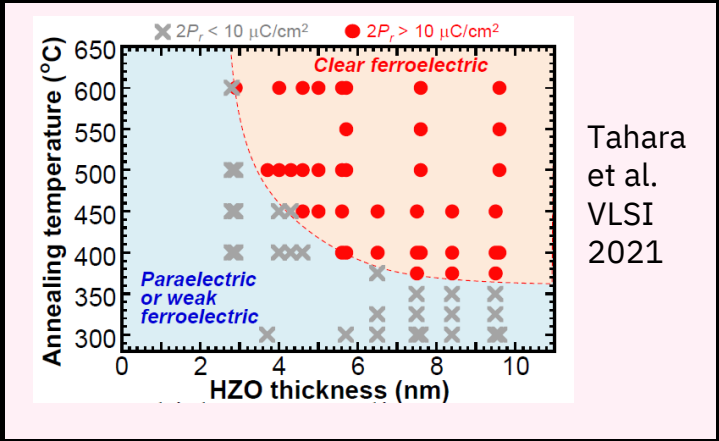
- Atomic Layer Deposition of TiN / $HfZrO_4$ / WO_x / TiN
- Crystallisation by millisecond Flash Lamp Annealing
  📖 É. O'Connor *et al.* (2018)
- Device definition by Reactive Ion Etching & Ion Beam Etching



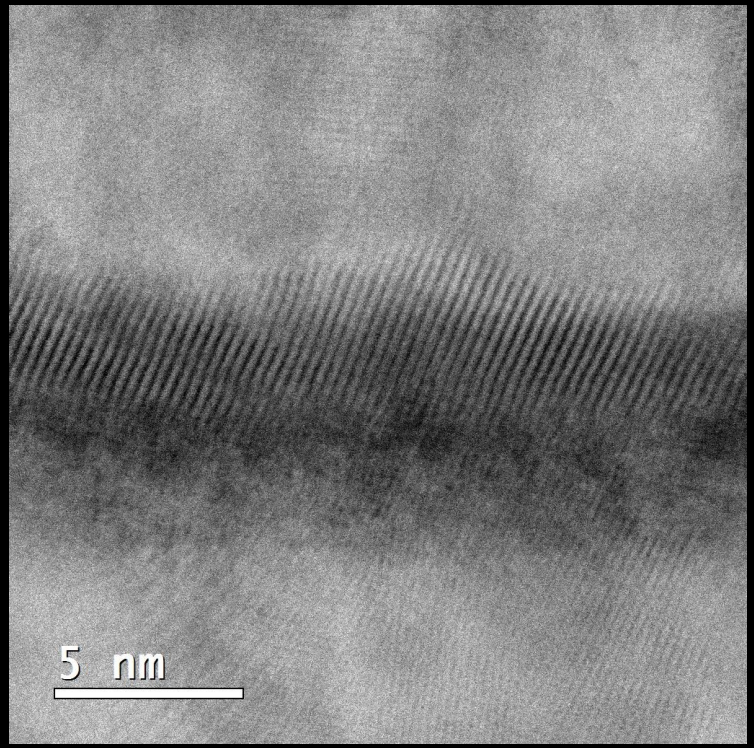
Back-End-Of-Line integration: constraint on the thermal budget (<400C)

Neuromorphic chips design: constraint on the device footprint (< $(10\mu m)^2$) + constraint on I_{OFF} & I_{ON}

Ultra-thin HZO (<3 nm): meet footprint + I_{OFF} requirement, but crystallises above 400°C.



Tahara et al. VLSI 2021



Bégon-Lours *et al.* Neuromorphic Computing and Engineering (2022) https://doi.org/10.1088/2634-4386/ac5b2d

TiN / HZO 4nm / WO_x / TiN / SiO_2 // Si junctions: Meet all the requirements for synaptic weights in the BEOL

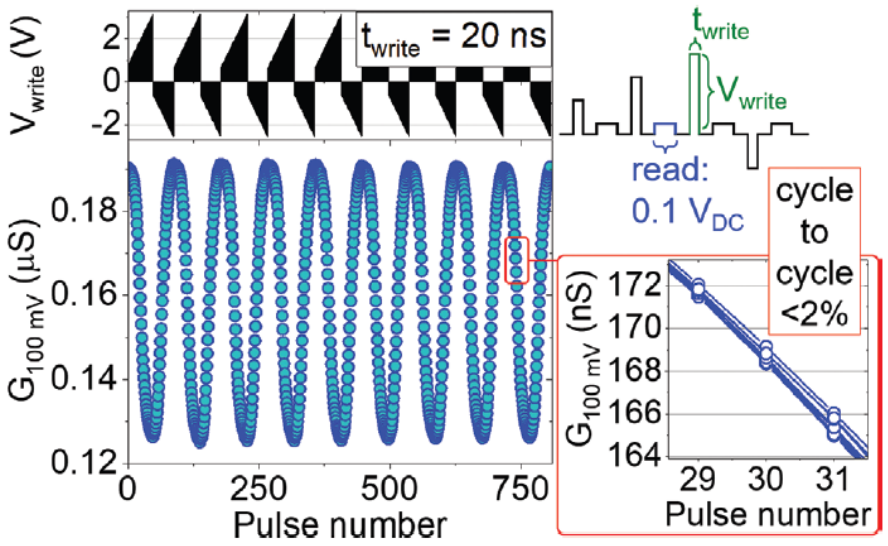


Figure 4. Synaptic potentiation and depression with pulses of increasing amplitude. The pulse amplitude V_{write} and the pulse duration t_{write} are defined as in the measurement scheme (top right). The inset shows the cycle-to-cycle variation.

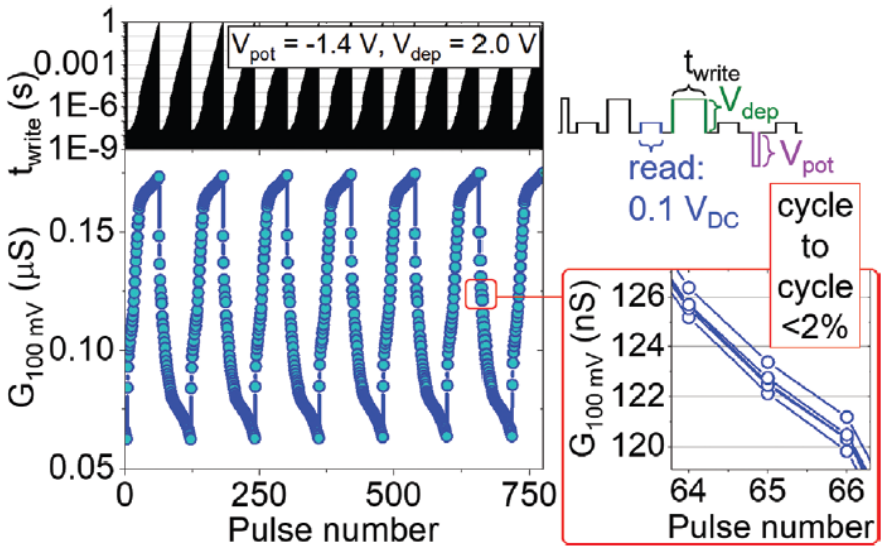
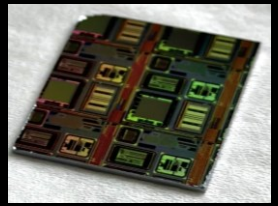


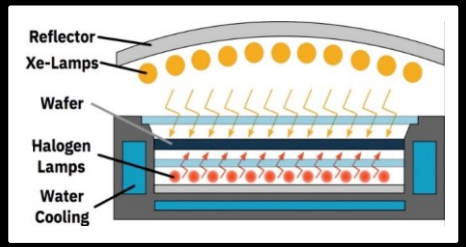
Figure 5. Long-term potentiation and depression for pulses with constant amplitudes V_{pot} and V_{dep} (as defined in the measurement scheme) and increasing duration t_{width}. The inset shows the cycle-to-cycle variation.

Laura Bégon-Lours / 29.10.2024 / ETHZ & IBM Research / LBEGON@ethz.ch

📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2022) https://doi.org/10.1002/aelm.202101395.

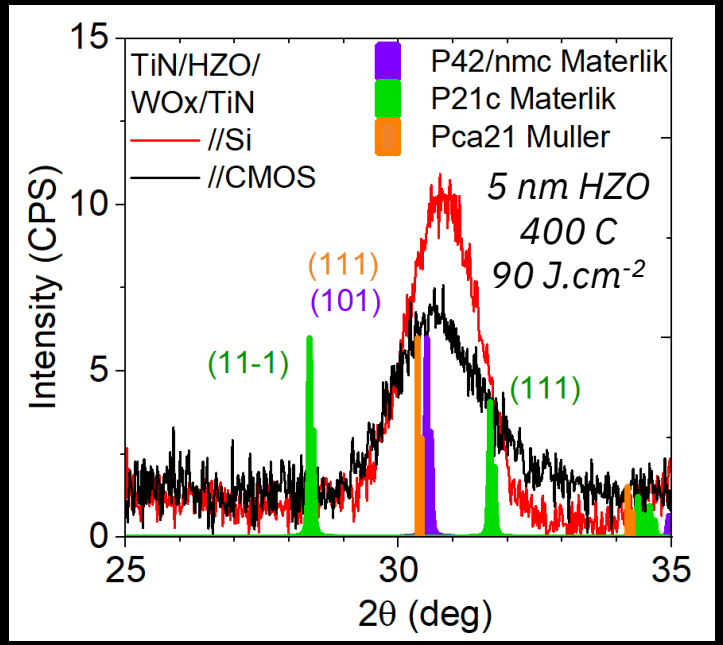
Back-End-Of-Line integration of TiN / HZO 4 nm / WO_x / TiN ferroelectric synaptic weights:

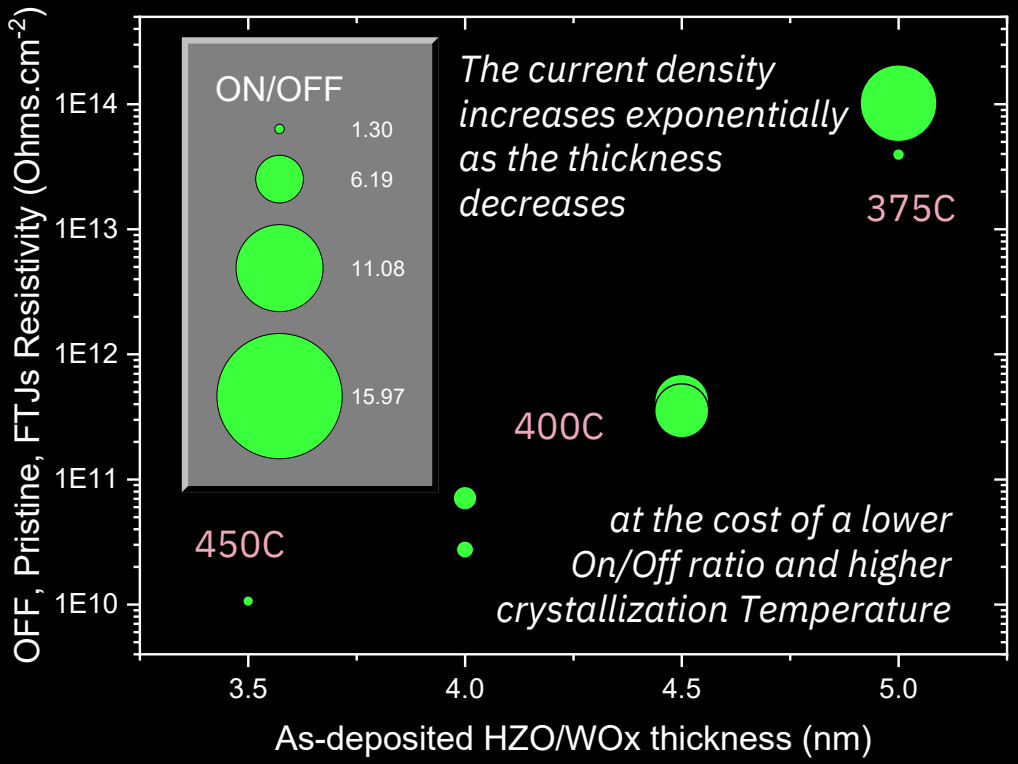




At 400C, crystallization of TiN / HZO 4 nm / WO_x / TiN / SiO_2 // Si using a millisecond flash lamp annealing (90 J/cm2), but not on the CMOS substrate.

TiN / HZO 5 nm / WO_x / TiN partially crystallises on the CMOS, but doesn't meet the footprint / I_{OFF} requirements.





Laura Bégon-Lours / 29.10.2024 / ETHZ & IBM Research / LBEGON@ethz.ch

Could ZrO_2 inclusion favor a lower crystallization temperature ?

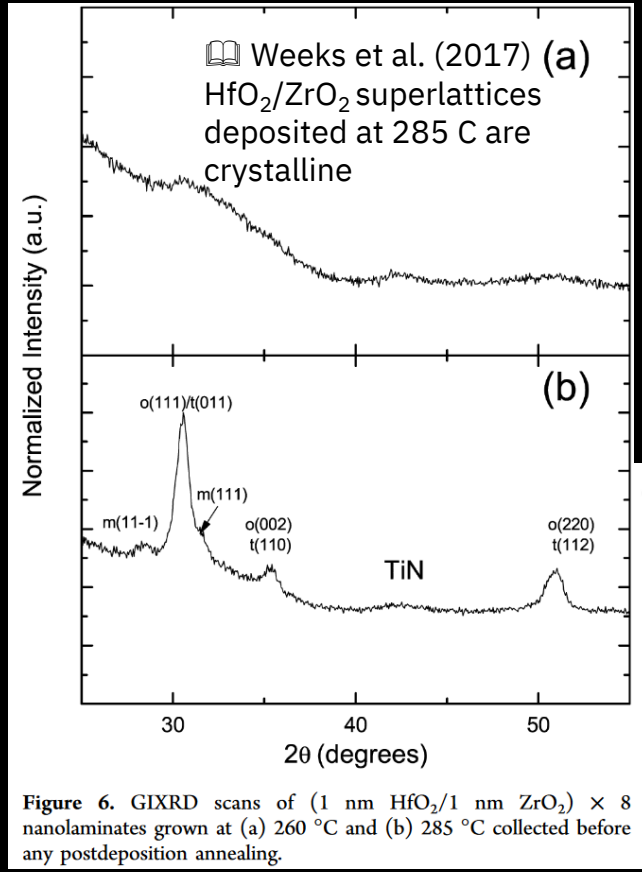
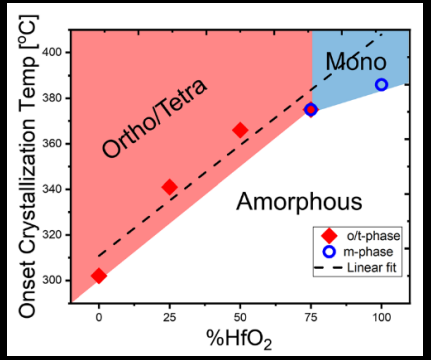


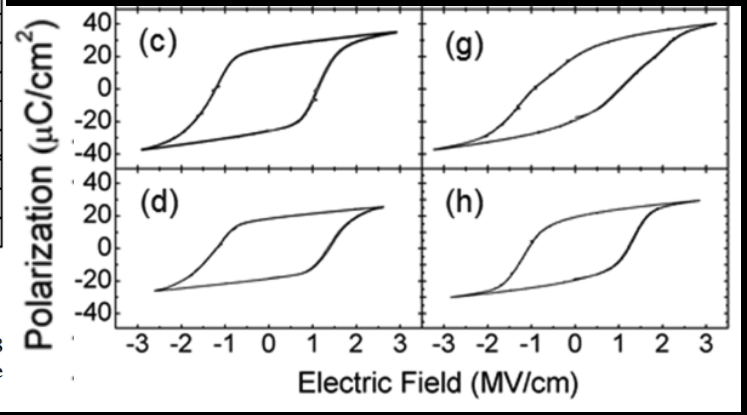
Figure 6. GIXRD scans of (1 nm HfO_2/1 nm ZrO_2) × 8 nanolaminates grown at (a) 260 °C and (b) 285 °C collected before any postdeposition annealing.

📖 Hsain et al. (2020): ZrO_2 has a lower crystallization temperature than HfO_2.



ALD: 285C + annealing

ALD: 260C + annealing



(1 nm HfO_2 / 1nm ZrO_2)x4

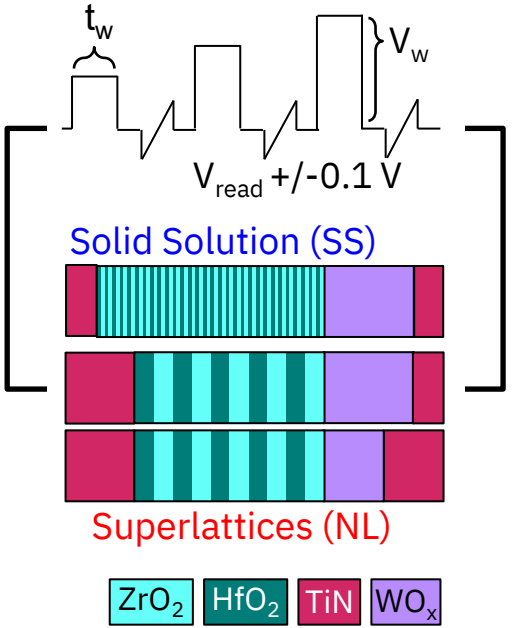
$HfZrO_4$

1. Motivation:
    a. Analog resistive switching in $HfZrO_4$ devices for Artificial Neural Networks hardware
    b. HfO2 / ZrO2 superlattices for lower crystallization T.

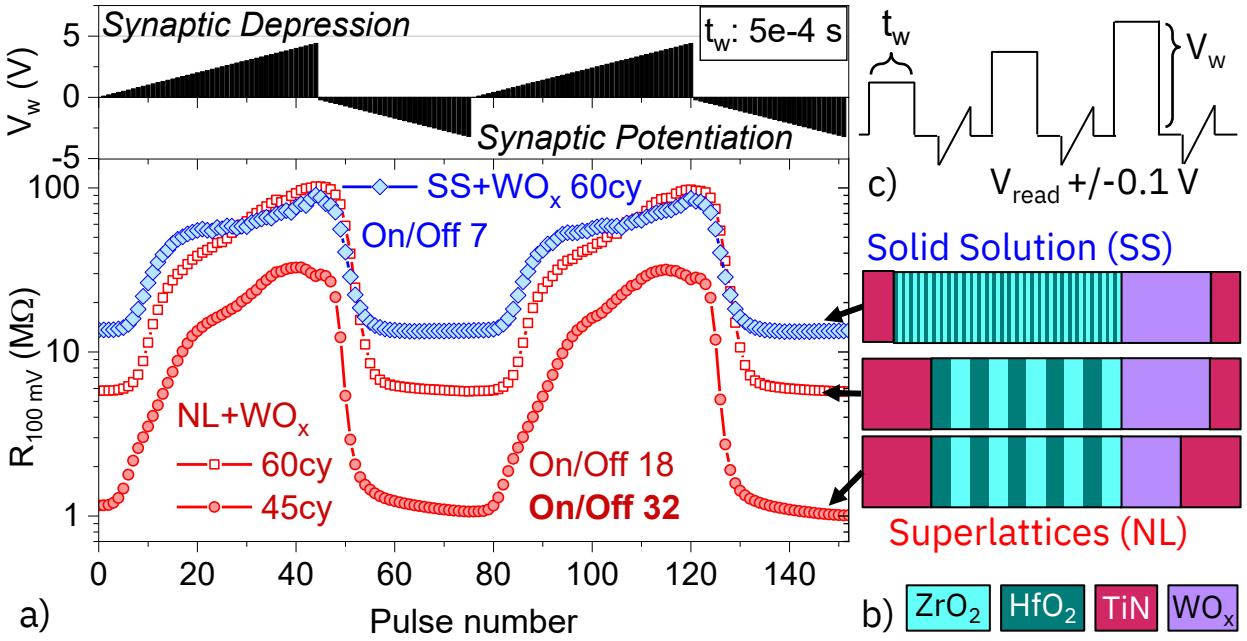
2. Results:
    a. 5x(0.5 nm HfO_2 / 0.5 nm ZrO_2) / WO_x // Si
    b. Co-integration of FTJ to CMOS in the BEOL

HfO_2 / ZrO_2 superlattices by Atomic Layer Deposition

- 5 repetitions instead of 1 repetition in each supercycle.
- Millisecond flash lamp annealing at 400C for superlattices (4 nm) & solid solution (5 nm)
- Functionality test:
  - Writing: pulses of increasing amplitude and constant duration,
  - Reading: IV Sweep



HfO_2 / ZrO_2 superlattices by Atomic Layer Deposition



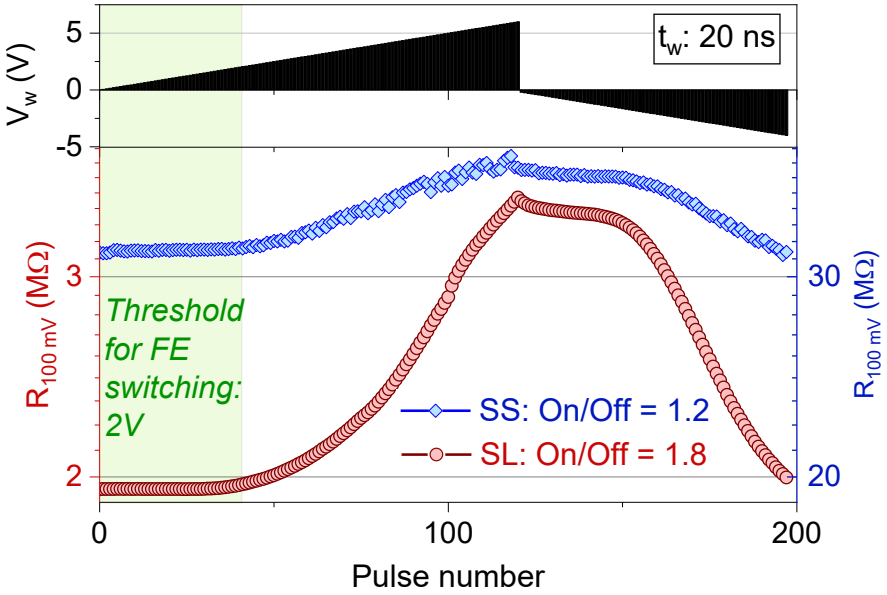
- HZO-SS ->HZO-SL. For equal I_{Off}, I_{on} increases + increase of the On/Off ratio from 7 to 18.
- Optimization of the WOx interlayer thickness: a 50% thinner WOx layer further decreased the RON, the ROFF and **increased the On/Off ratio to 32**

📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

Ultra-fast switching (20 ns programming pulses)

HZO-SS and HZO-SL synaptic weights on Si are operated under the increasing amplitude scheme (+6V/-4V, step size: 50 mV) with a **pulse width of 20 ns.**

Also in this ultra-fast regime, it is one order of magnitude smaller for the SL compared to the SS, while exhibiting a larger On/Off.

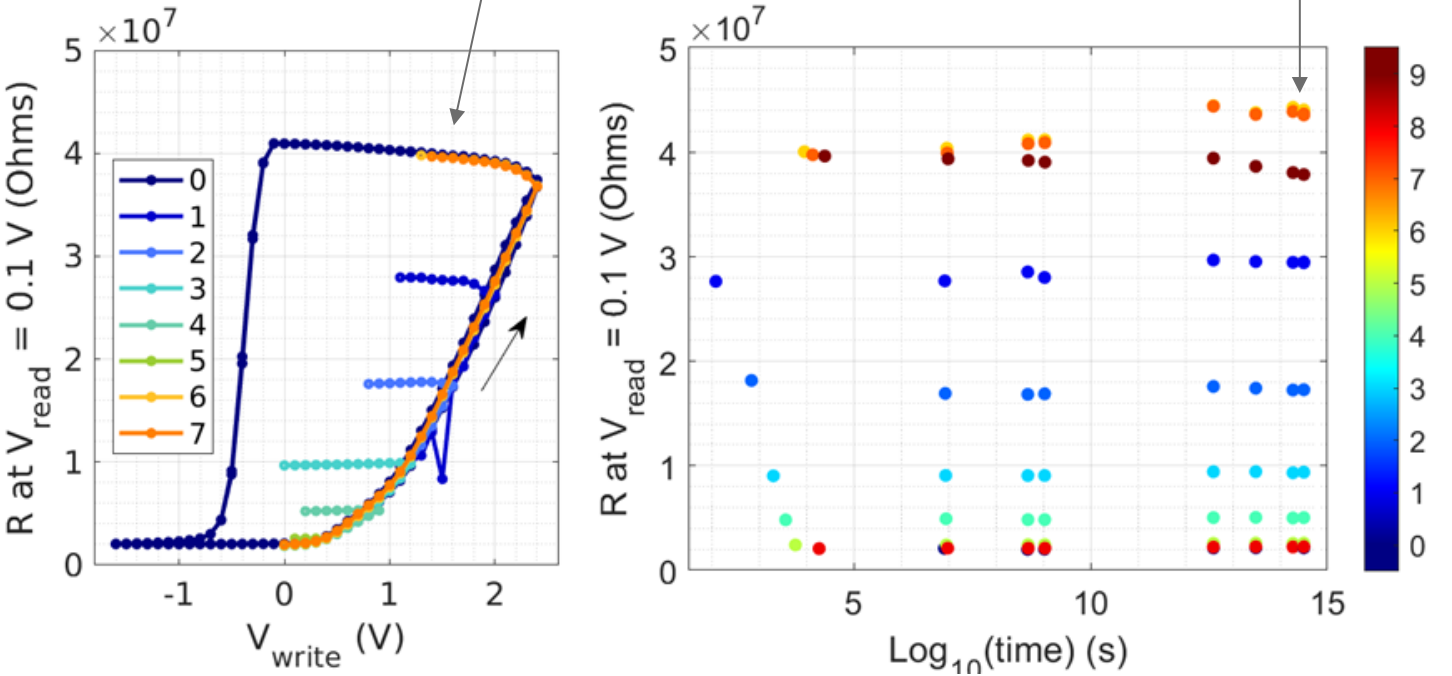


📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

Retention of the programmed states

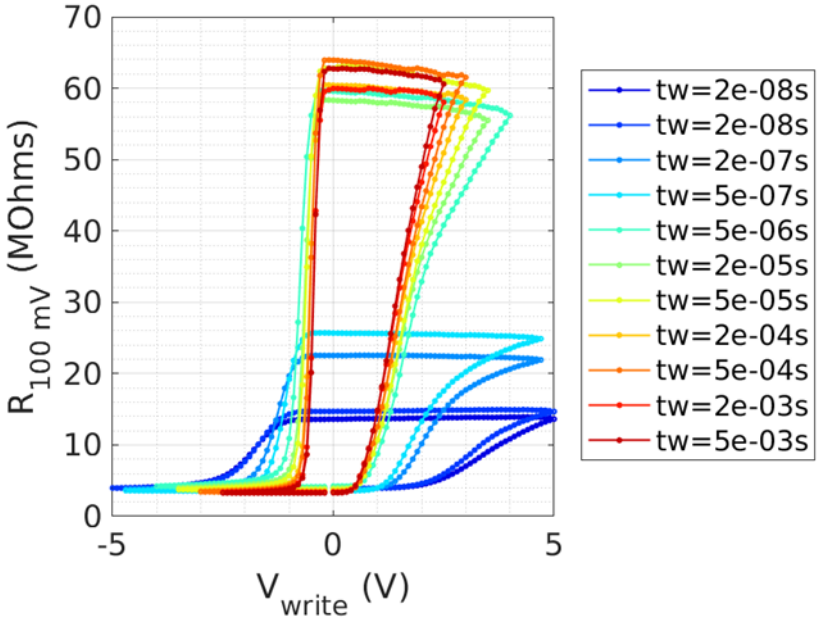
Resistance increases upon pulses of decreasing amplitude: else than ferroelectricity ?

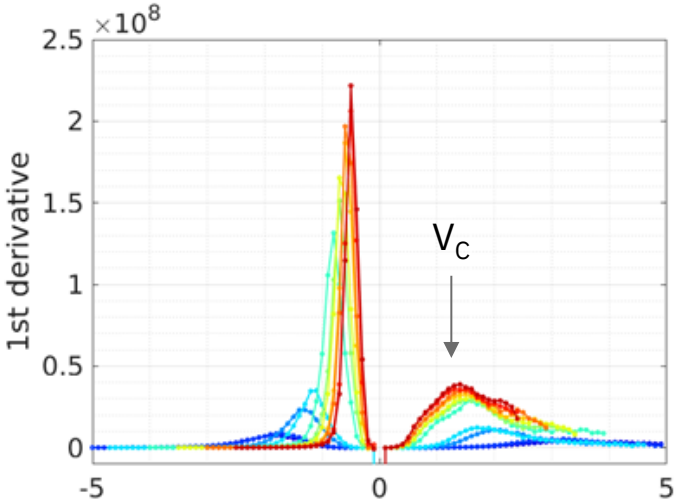
3 weeks: excellent retention of the programmed states

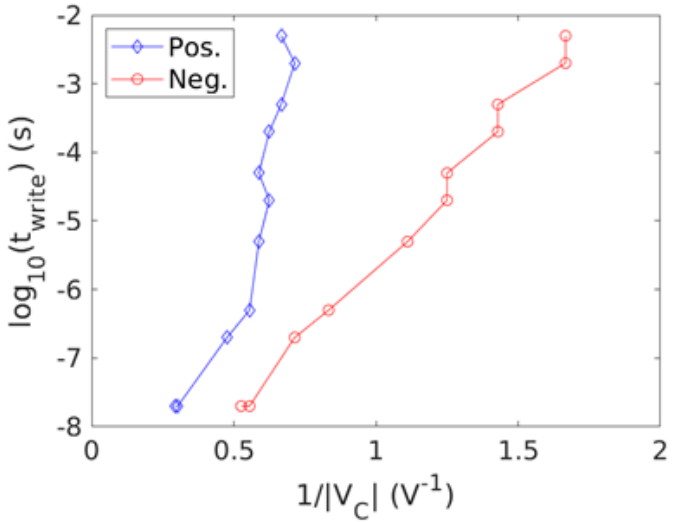


📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

Merz law in ferroelectrics:



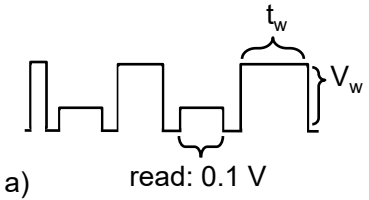


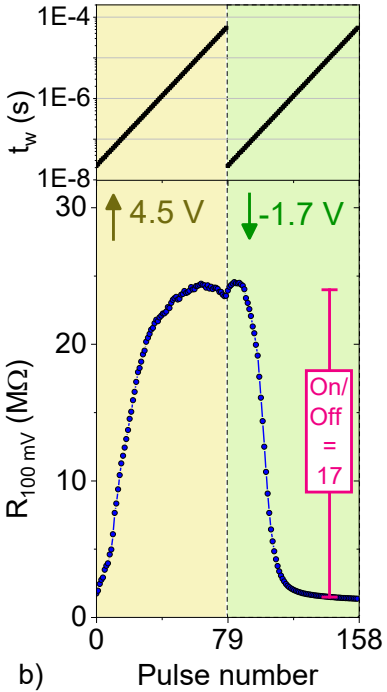


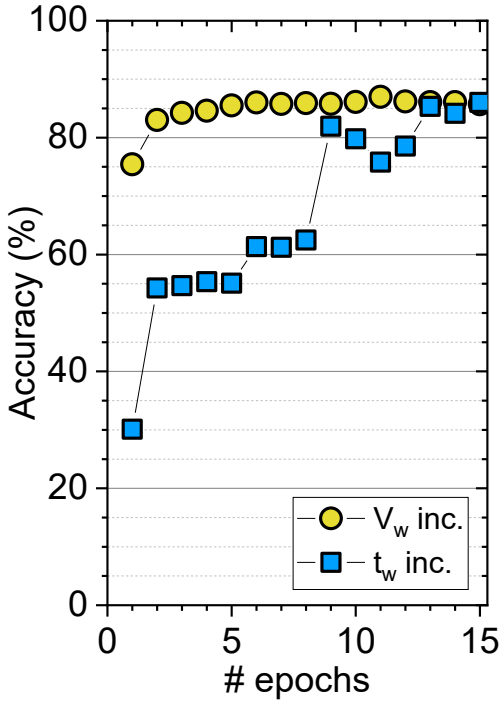
The coercive field depends on the programming pulse duration, but asymmetrically.

Increasing pulse width scheme: slower convergence than increasing amplitude scheme, same accuracy.

NeurosimMLP+V3.0 simulation framework (Luo 2019)
Estimated accuracy of a ANN with 400, 100 and 10 neurons in the input, hidden and output layers (MNIST dataset).

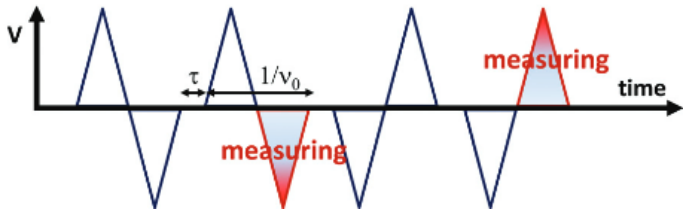




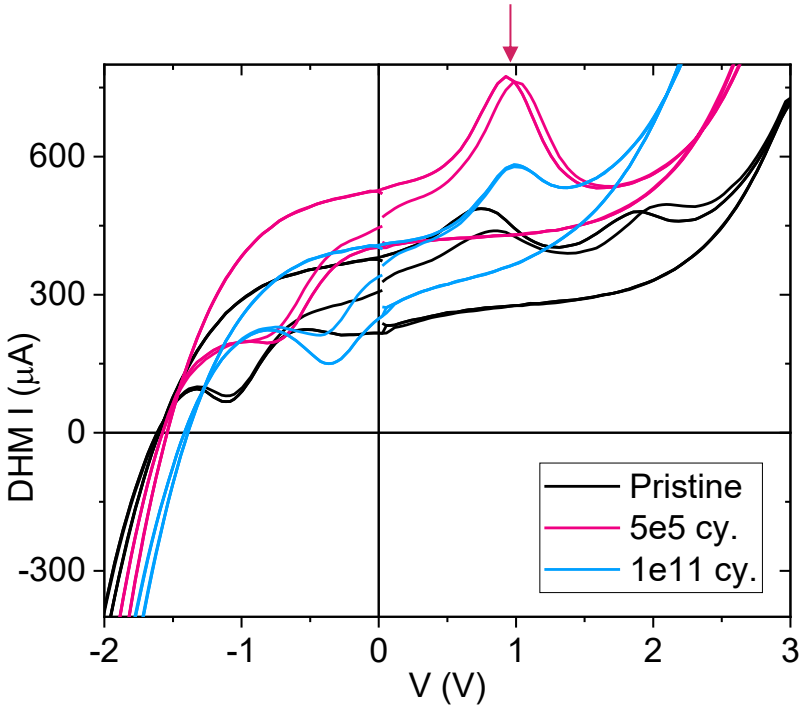


📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

Ferroelectric characterization:
Dynamic Hysteresis Mode & 3V endurance up to 10^{11} cycles.



Switching peak observed at 1V, but no saturation in the electroresistance loop: indicates non-ferroelectric contributions.



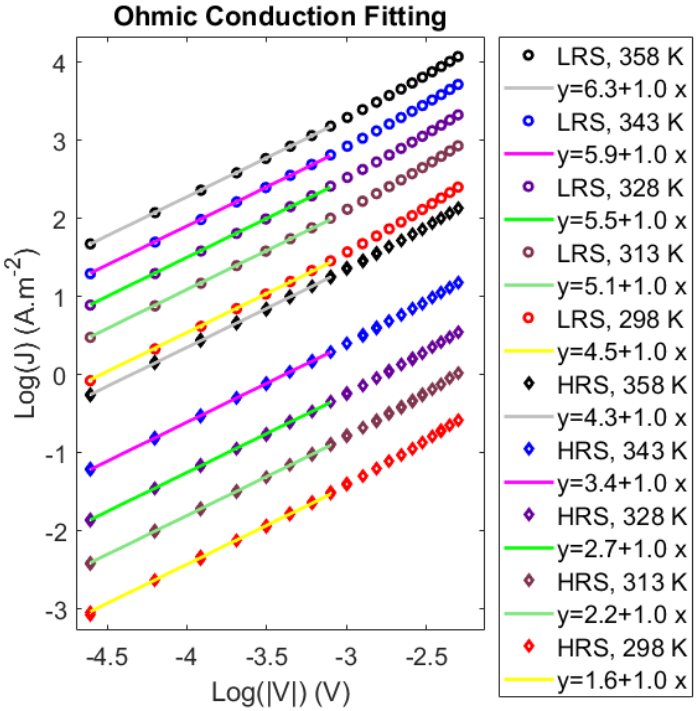
📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

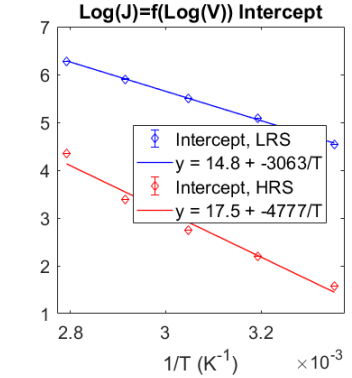
Linear weight read for analog computing

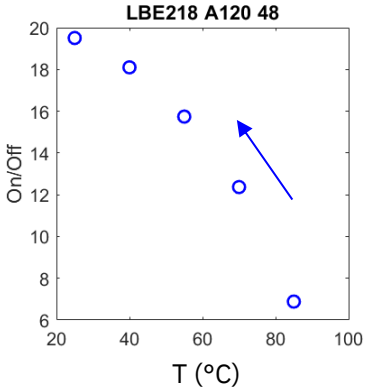
Linear (Ohmic) conduction, ideal for implementing vector-matrix multiplications.

$$J = q\mu N_C \exp(-(E_C - E_F)/kT)E$$

Temperature dependent measurement in the On and Off states:
Cumulated contributions of a change in the energy barrier (0.26 / 0.41 eV) and a change in the carrier concentration * mobility product (8.7e14 / 1.2e16 $(cm.V.s)^{-1}$ at RT).

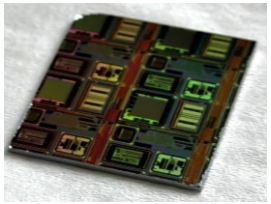




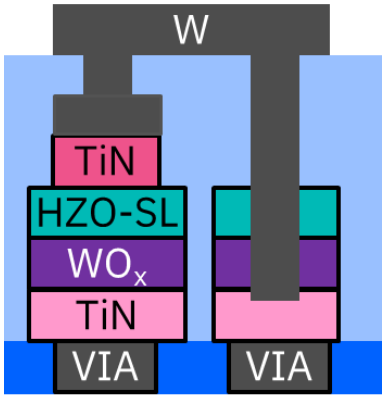


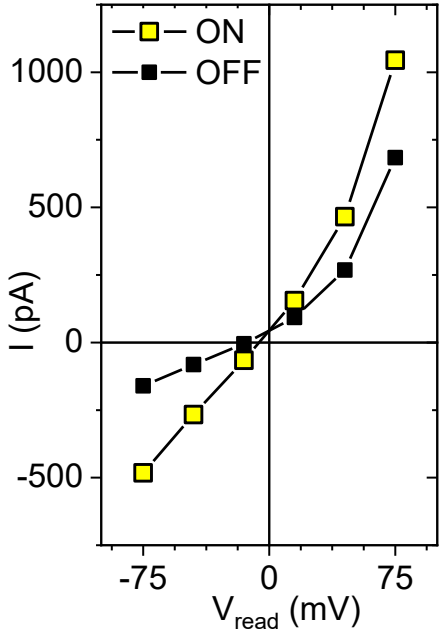
📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

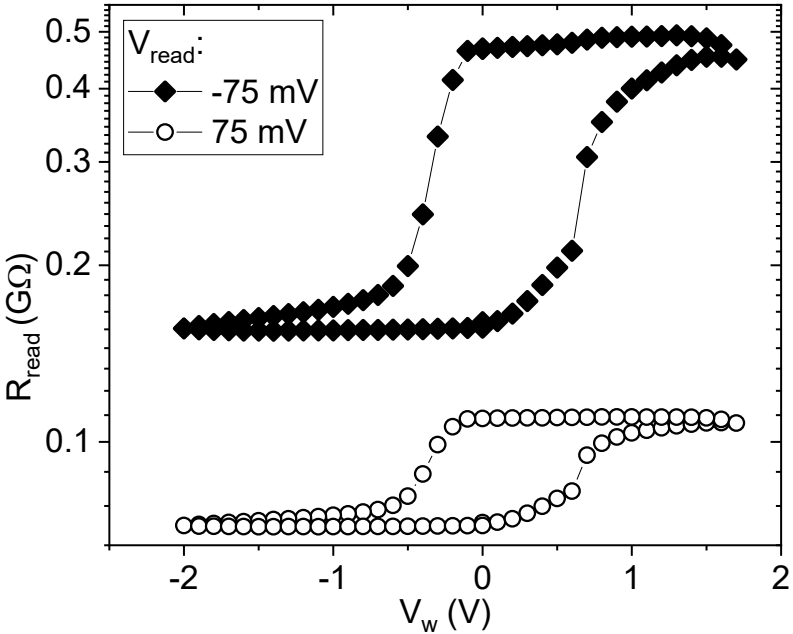
Successful co-integration of HZO-SL synaptic weights to XFAB 180 nm CMOS: analog resistive switching



BeferroSynpatic tape-out





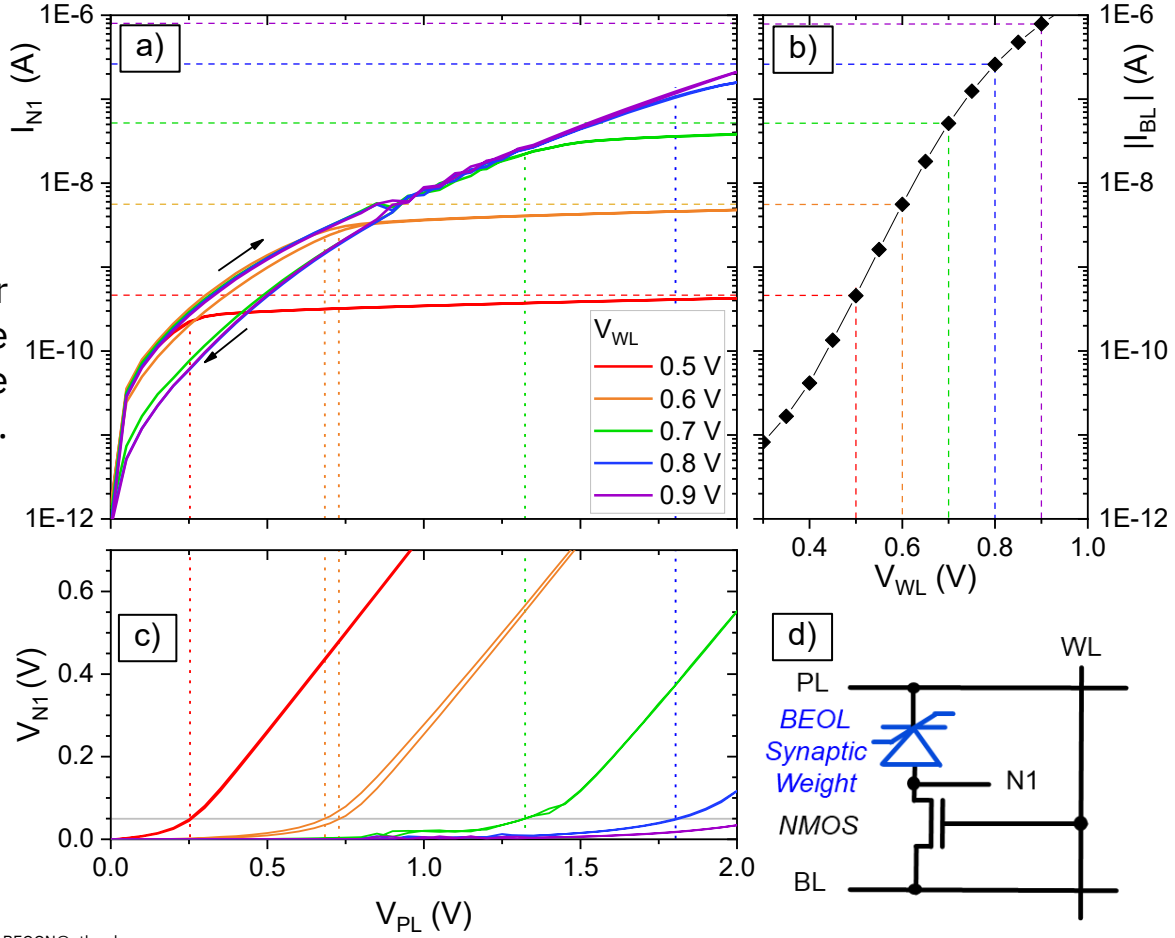


📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

Multi-level, 1T-1R operation.

$I_{N1} = f(V_{PL})$, for various gate voltage V_{WL} on the NMOS transistor.

The compliance gives rise to a potential V_{N1} at the source of the NMOS.



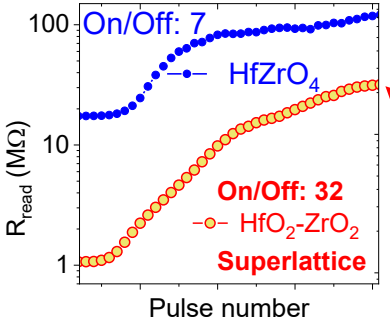
NMOS Id-Vg: the compliance observed in a) originates from the NMOS

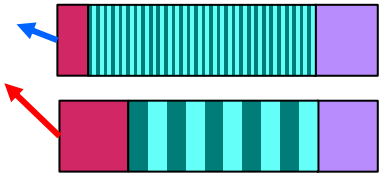
📖 Bégon-Lours *et al.*, under review in Adv. Elec Materials

📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

Summary: ultra-thin HfO_2-ZrO_2 nanolayers

- CMOS integration of HZO was compromised by the temperature / footprint / I_{OFF} constraints.
- The novel HfO_2/ZrO_2 superlattice synaptic weights exhibit:
  - Large On/Off up to 32
  - x8 improve in scalability,
  - Ultra-fast switching (20 ns),
  - Linear read out (in On and Off states),
  - High endurance (10^{11} full switching cycles).
- The cointegration of HZO-Superlattice to CMOS was successful:
  - Analog Resistive Switching
  - Multi-Level 1T-1R operation
- Outlook: hardware for artificial neural networks, supporting bio-inspired computing.





📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

📖 Bégon-Lours *et al.* Adv. Elec. Mat. (2024) DOI 10.1002/aelm.202300649.

📖 Hamming-Green *et al.* IEEE EDTM (2024) DOI 10.1109/EDTM58488.2024.10511719

Stay in touch !
Email: LBEGON@ethz.ch
Linkedin: Laura Bégon-Lours