CMOS-Compatible Ferroelectric Synapse Technology for Analog Neural Networks

Yanjie Shao, Elham R. Borujeny, Taekyong Kim,

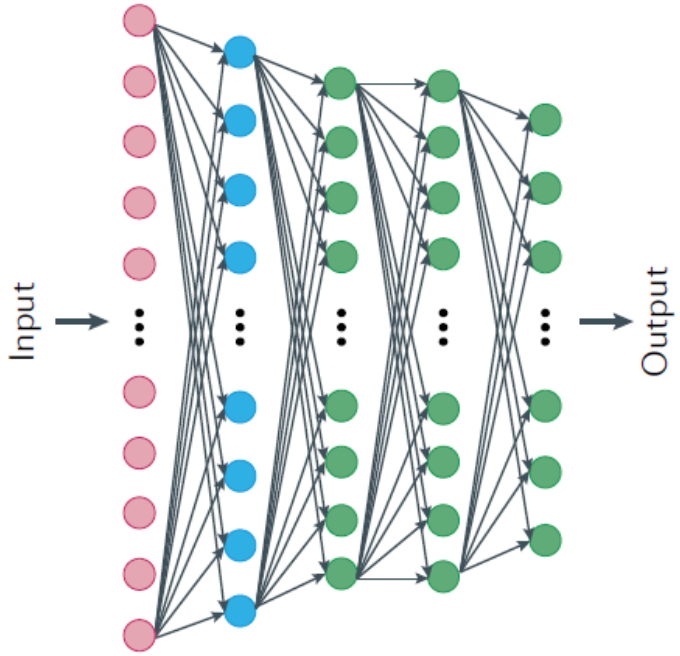
Dimitri A. Antoniadis, and Jesús A. del Alamo

MIT EECS & MTL

Sponsorship: Intel through AI HW program; SRC GRC program

MIT AI Hardware Program annual symposium
MIT, April 6, 2023

Ever-increasing computing needs of deep AI models

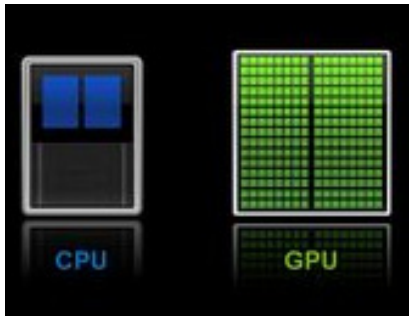


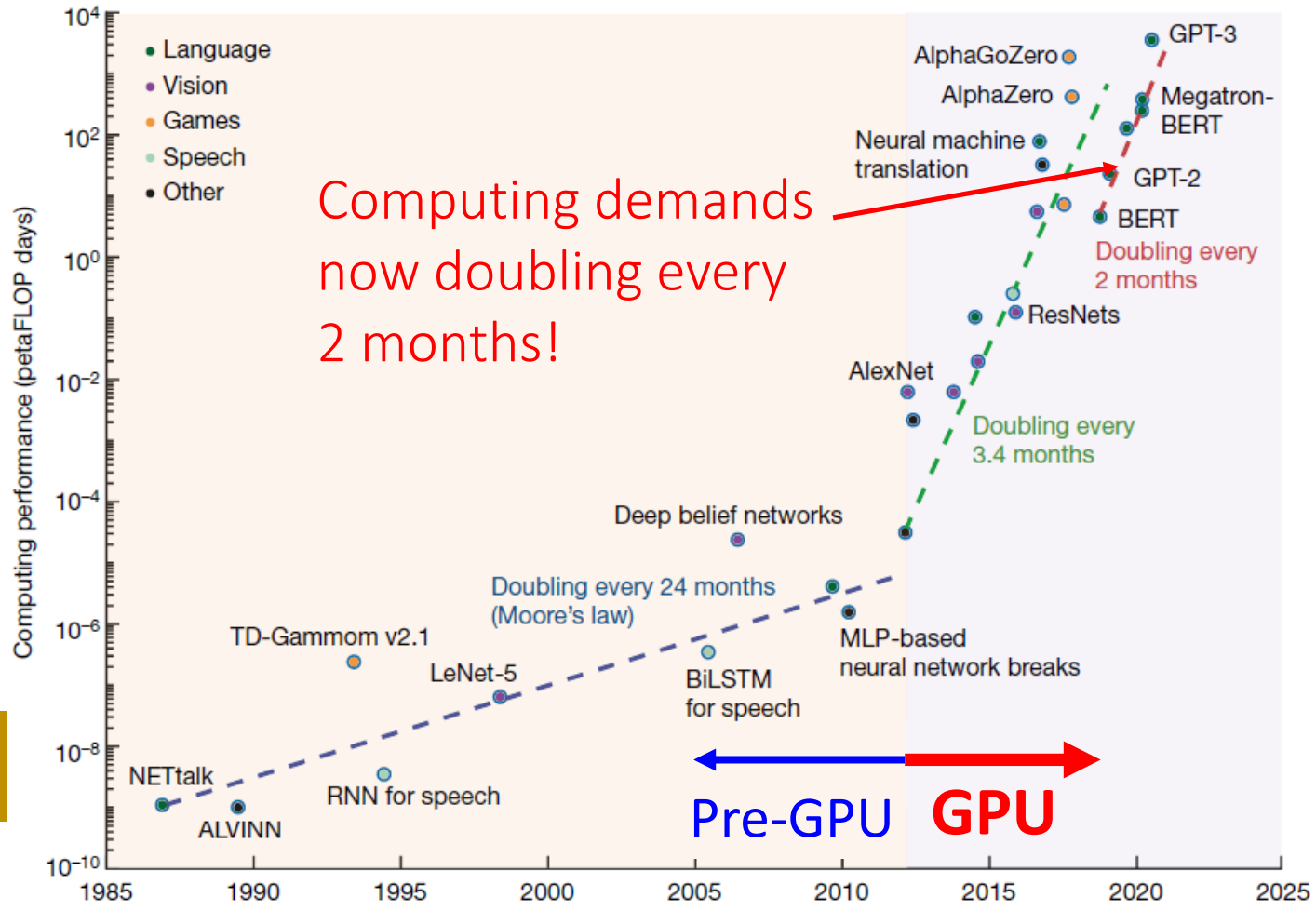
D. Marković *et al.*, *Nature Reviews Physics* 2, 499–510 (2020)

Larger model

Larger dataset

Performance ↑
but… compute ↑↑





A. Mehonic and A. J. Kenyon, *Nature* 604, 255–260 (2022)

GPU is great, but…

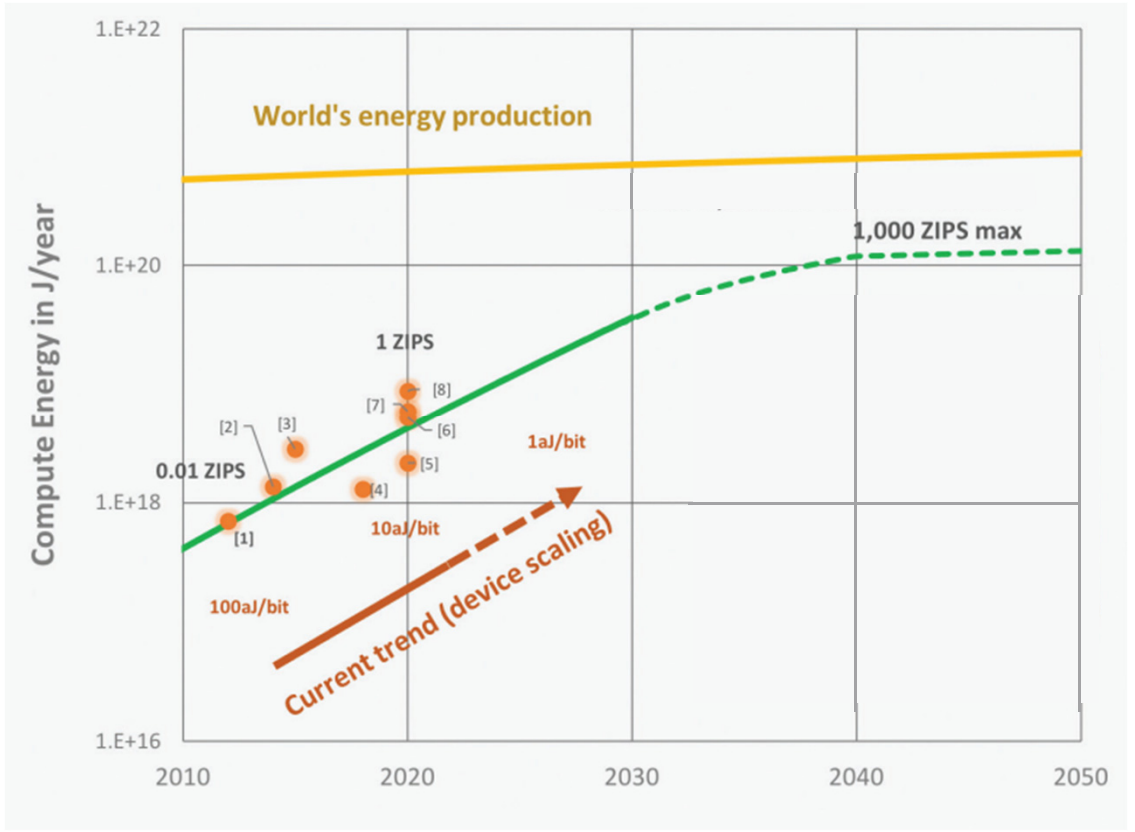
Need of energy-efficient AI hardware

WIRED

WILL KNIGHT BUSINESS JAN 21, 2020 7:00 AM

AI Can Do Great Things—if It Doesn't Burn the Planet

The computing power required for AI landmarks, such as recognizing images and defeating humans at Go, increased 300,000-fold from 2012 to 2018.



The Decadal Plan, SRC 2021

- Energy demands for compute now approaching total global energy production
- Specialized energy-efficient AI hardware in need

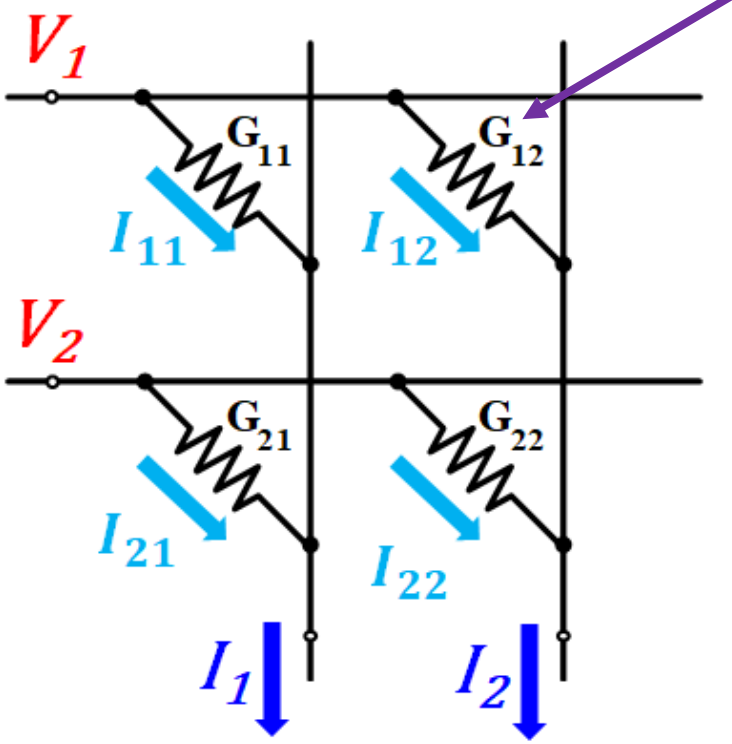
Analog hardware for deep learning

Multiply-Accumulate operation at heart of Matrix Multiplication

$$I_j = \sum_{i=1}^{N} G_{ij} V_i$$

Kirchoff's Law

Ohm's Law



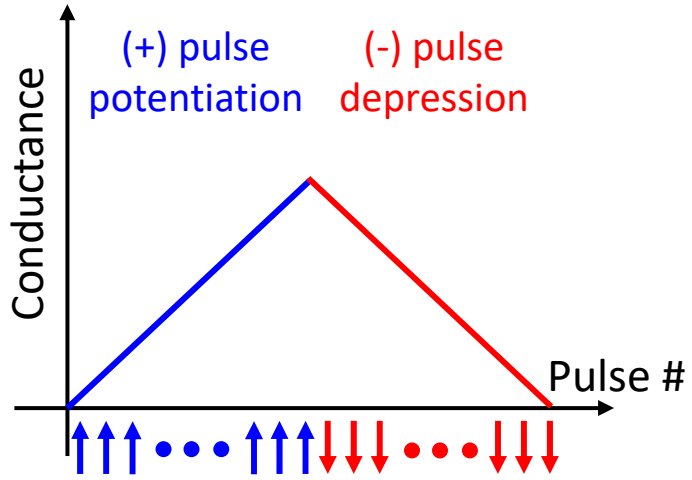
Programmable non-volatile resistor ("synapse")

Analog vs. Digital NNs:

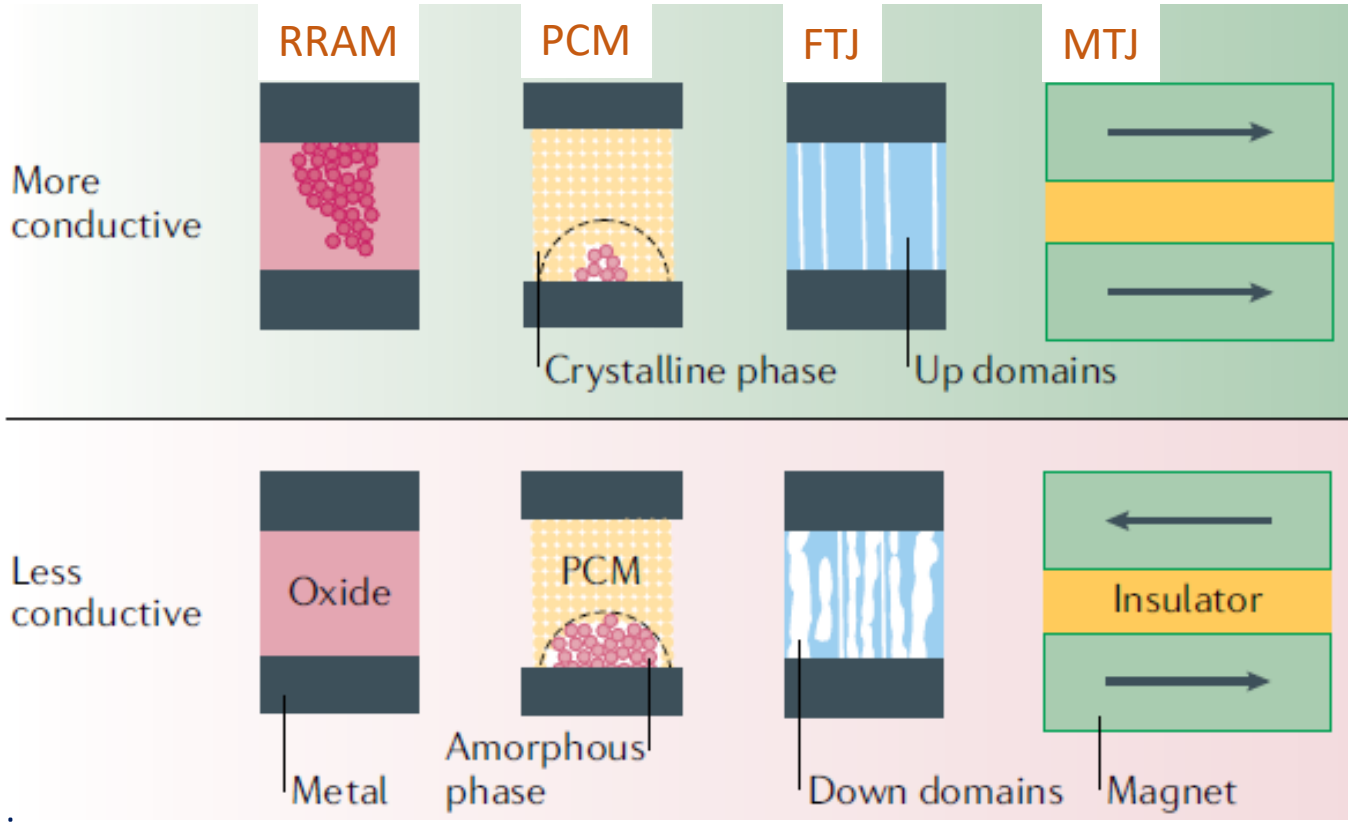
- Footprint ↓
- Latency ↓
- Energy efficiency ↑↑

Desired device characteristics

Ideal characteristics:



- G~0.1 μS
- High speed (ns)
- High efficiency (sub-pJ)
- Large dynamic range (~10x)
- Fine-tunable (~1000 levels)
- Low variability
- Symmetric modulation
- High endurance
- Good retention
- CMOS compatible



D. Marković *et al.*, *Nature Reviews Physics* 2, 499–510 (2020)

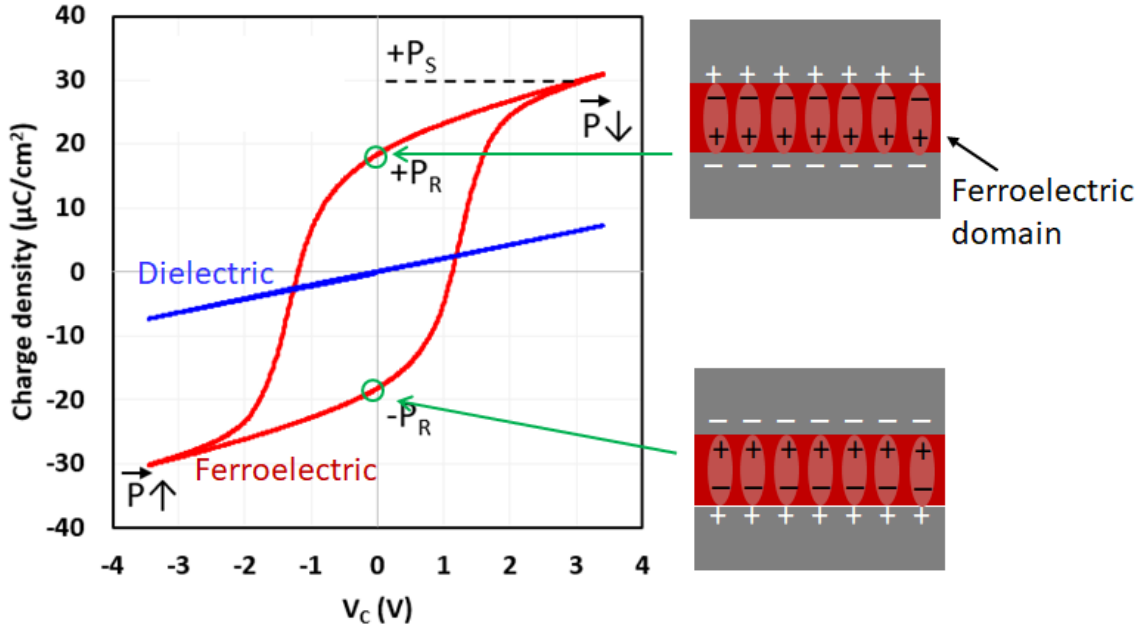
Device with desired characteristics does not exist today!

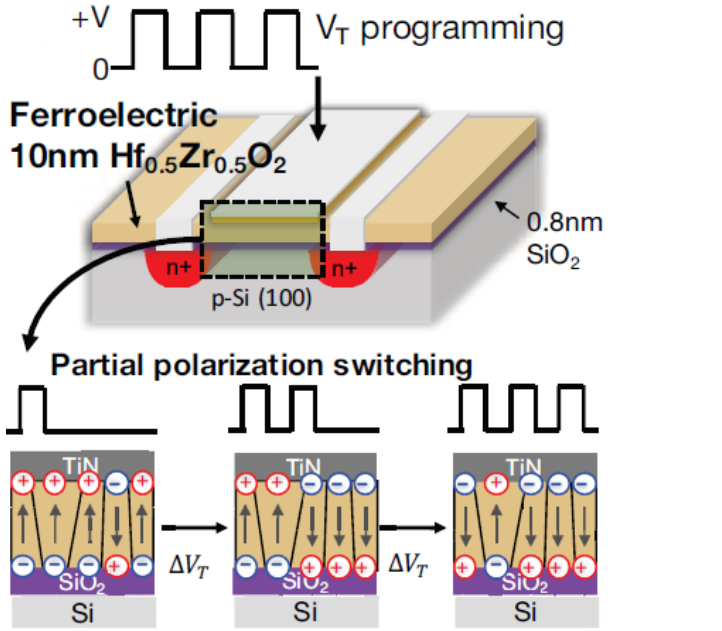
T. Gokmen and Y. Vlasov, *Frontiers in Neuroscience* 10: 333, 2016

Programmable resistor based on ferroelectric effect

Ferroelectric $Hf_xZr_{1-x}O$ (HZO):

- stores polarization charge due to unique crystal structure
- CMOS compatibility





M. Jerry *et al.*, *IEDM* 6.2.1-6.2.4 (2017)

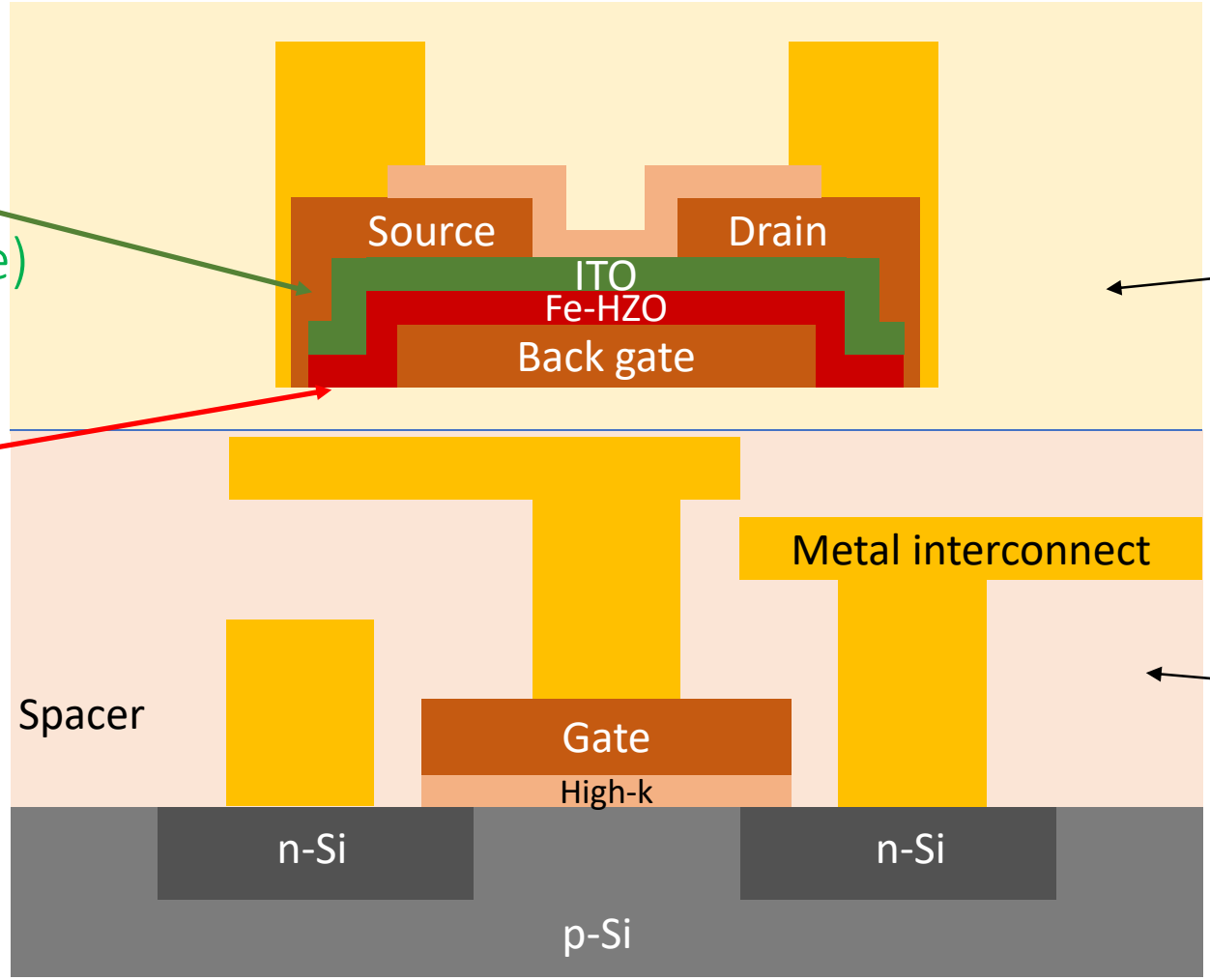
Multi-domain structure: partial polarization switching → analog behavior

Our approach: non-volatile thin-film ferroelectric MOSFET

Critical requirement: Back-End CMOS compatibility

InSnO channel
(room temperature)

Ferroelectric
HfZrO gate stack
(T ~ 400 °C)



FE Synapse
fabricated in
CMOS Back-End

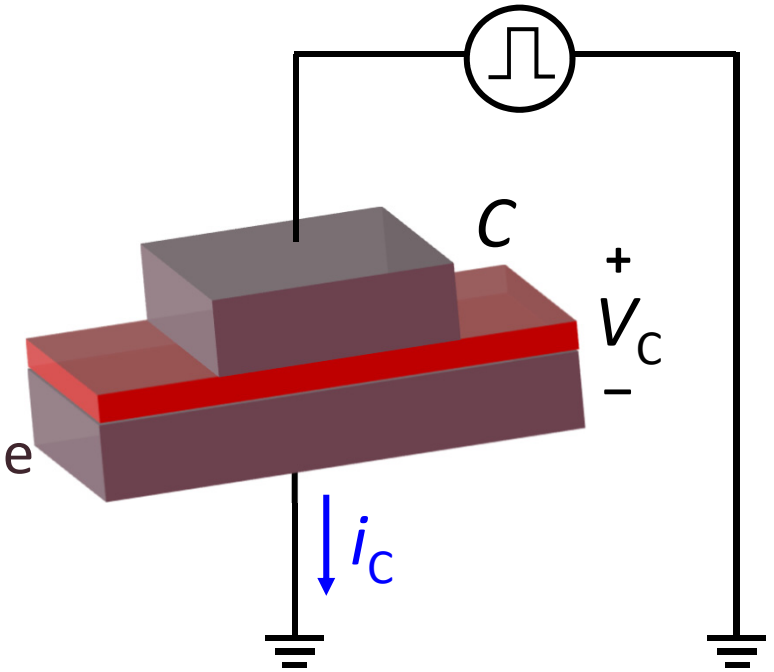
CMOS
Front-End

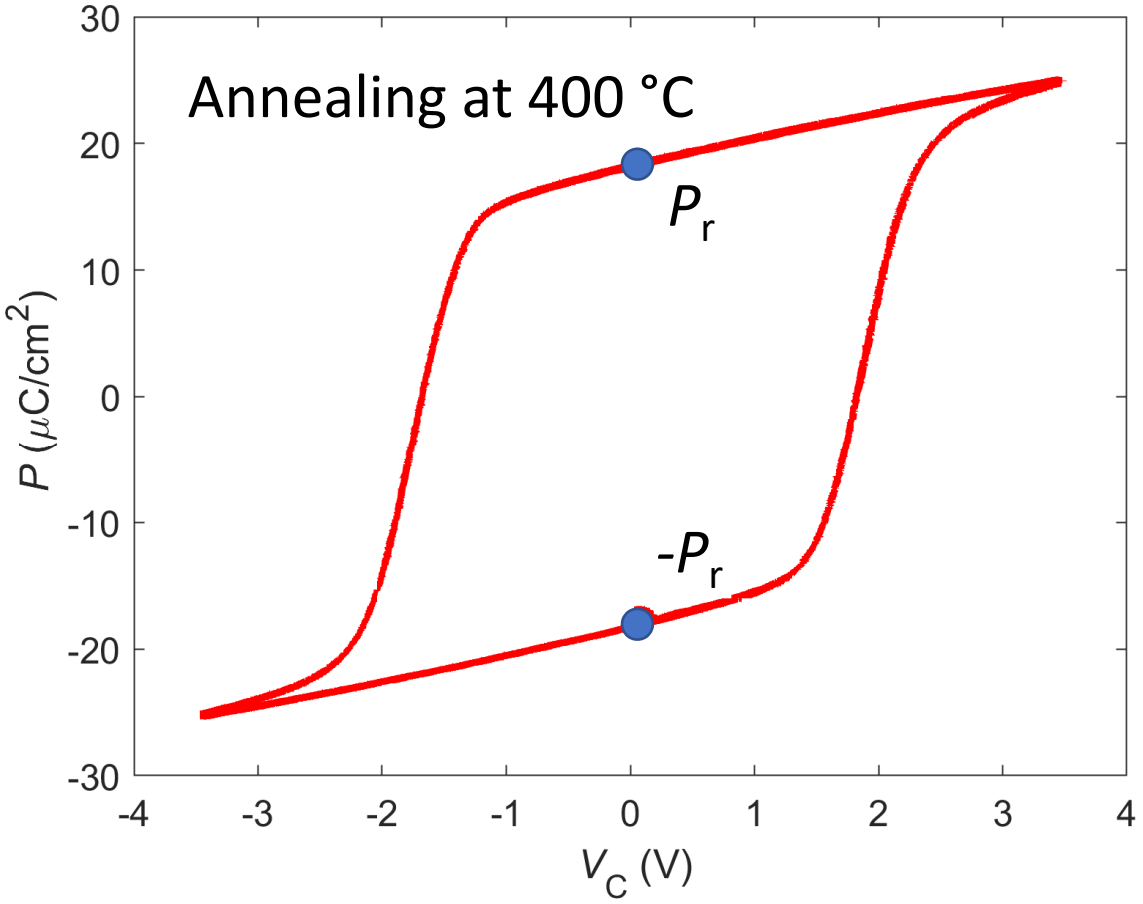
Enabling component #1: high-quality FE HZO with T ≤ 400 °C

W top electrode

10 nm FE HZO

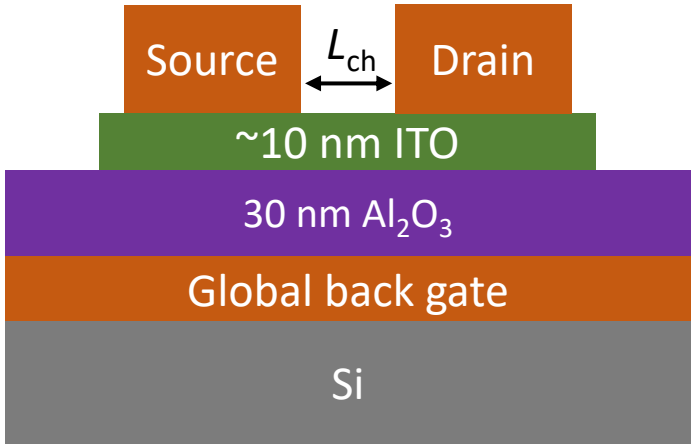
W bottom electrode





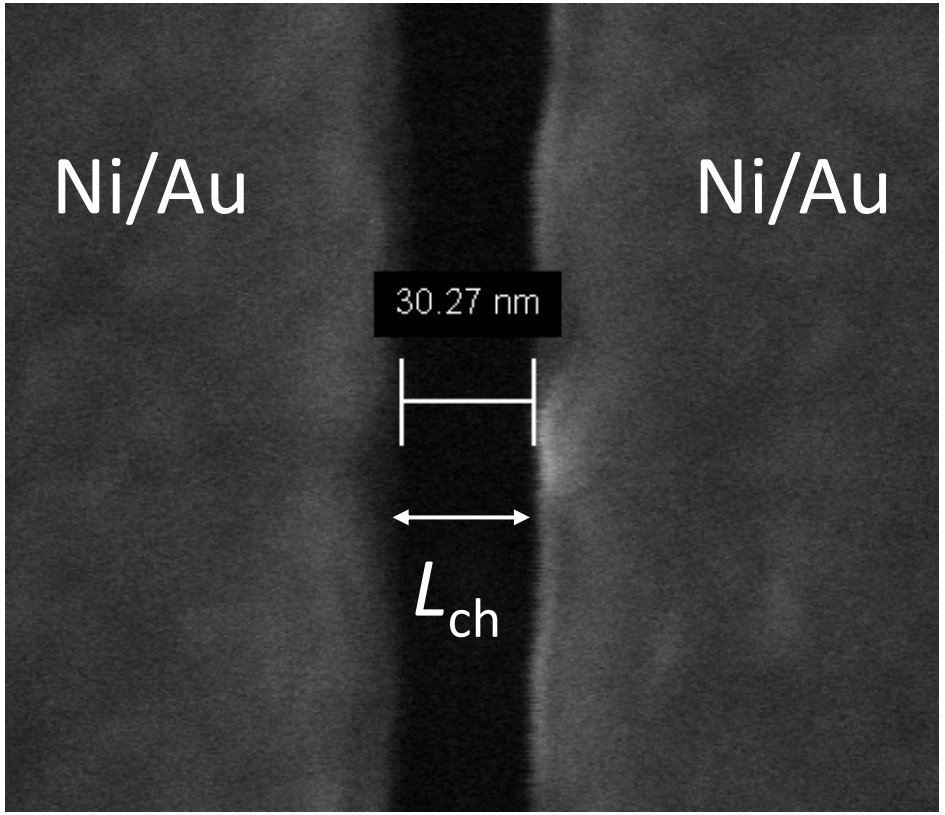
- Remnant polarization (P_r) ~ 18 µC/cm^2 with ±3.5 V sweep range

Enabling component #2: ITO thin-film transistor



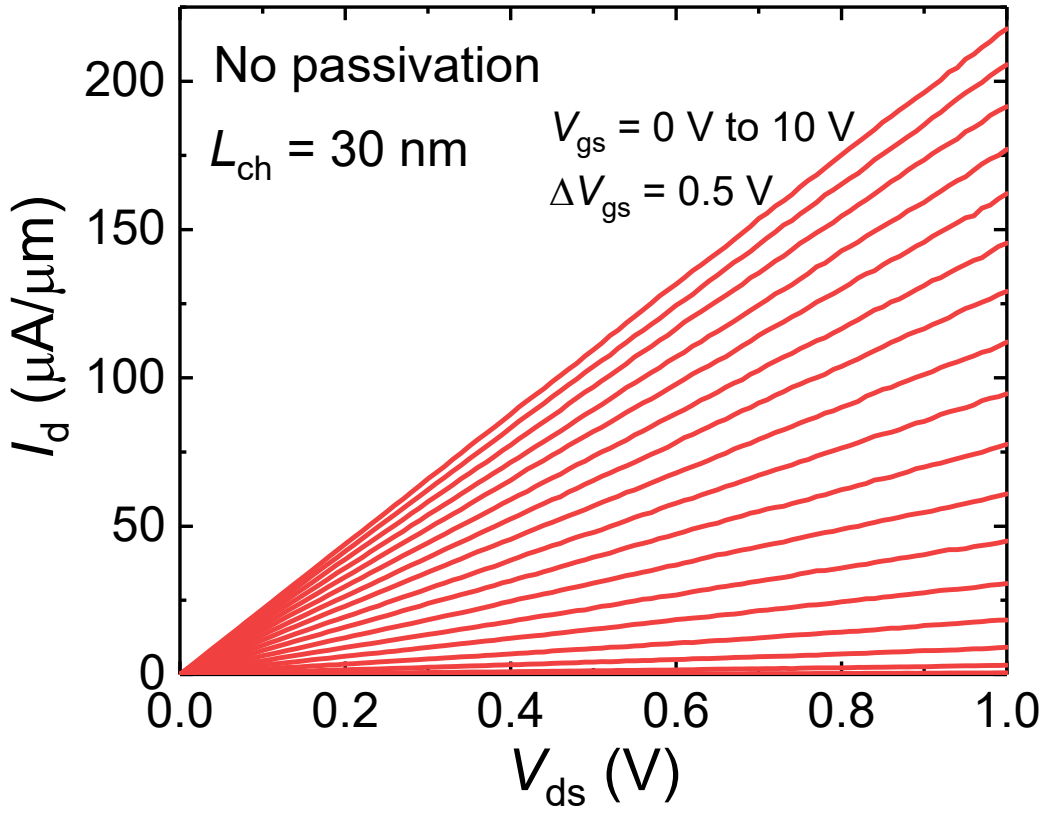
ITO thin film:

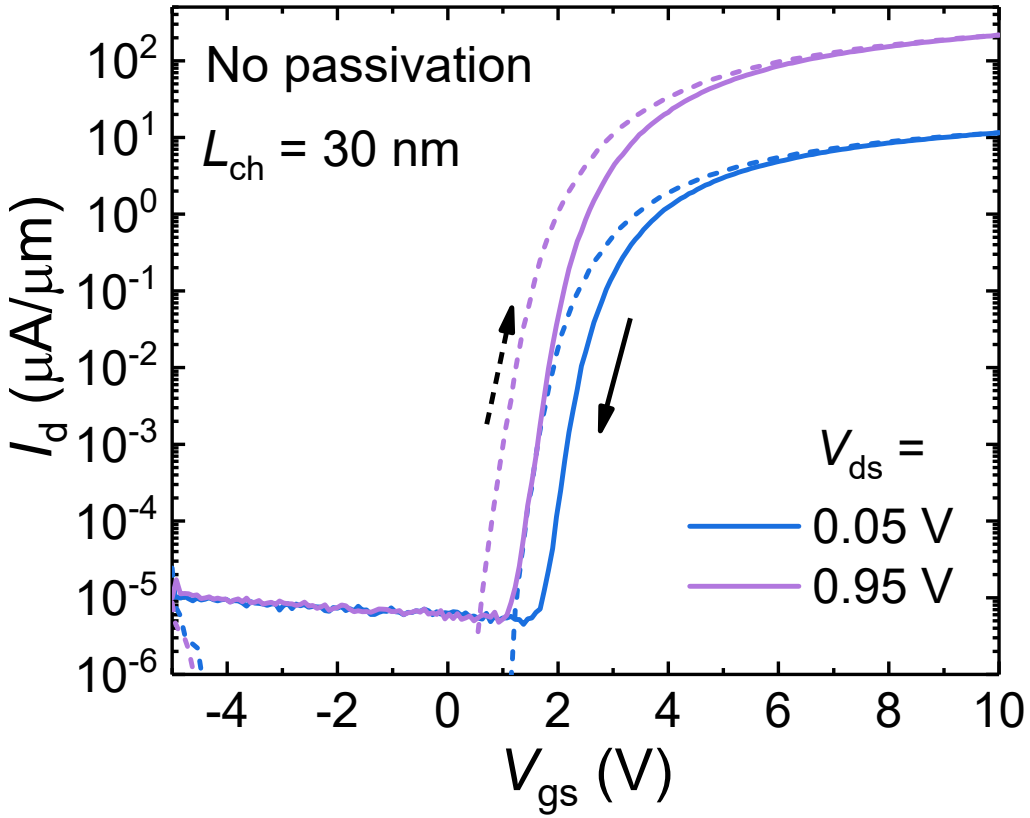
- Sputtering at room temperature
- No thermal annealing



- Global back gate transistors with scaled channel length

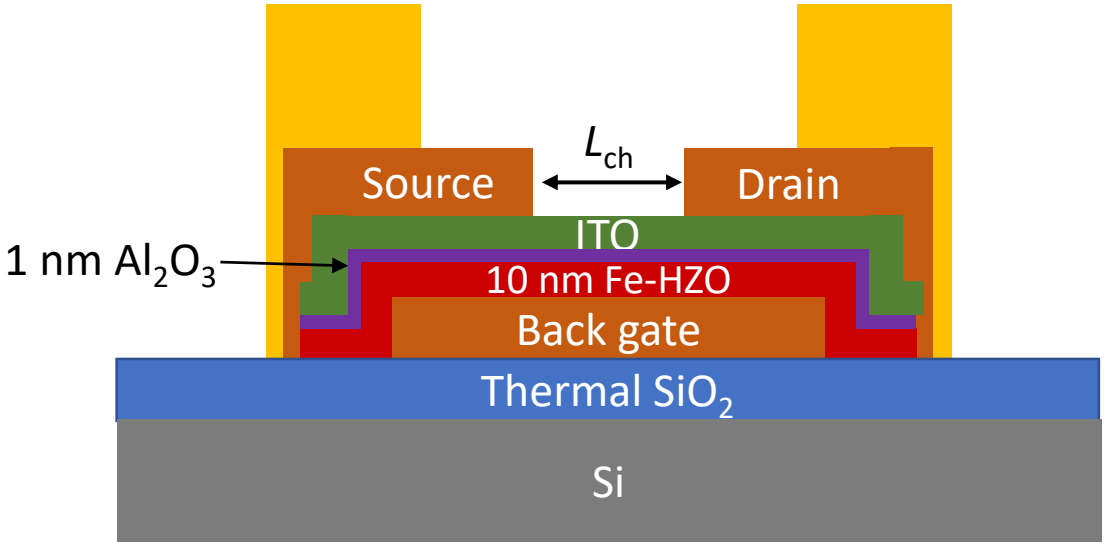
Enabling component #2: ITO thin-film transistor





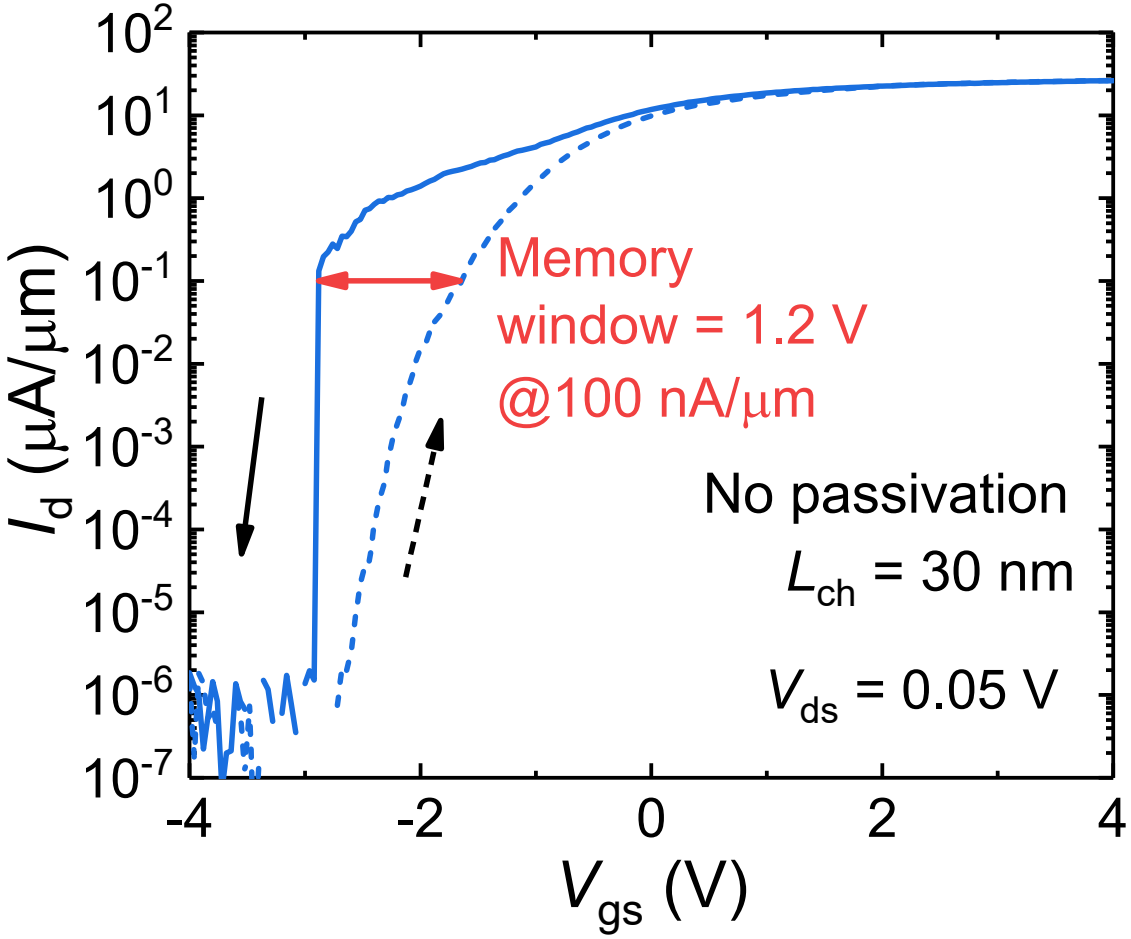
- Scaled L_{ch} enables I_{on} > 200 uA/um at V_{ds} = 1 V

Ferroelectric thin-film transistor: a first look



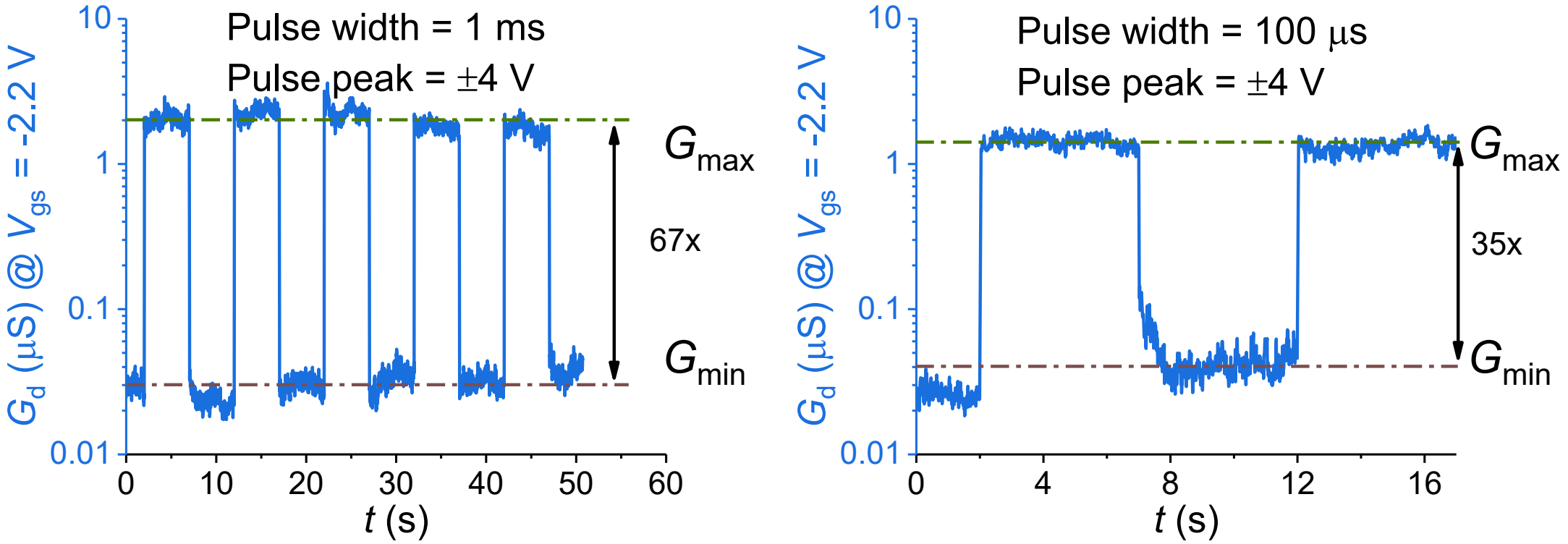
Process details:

- ITO thickness ~10 nm
- Highest process temperature = 400 °C



- FeTFT with a scaled L_{ch} and a large memory window

Ferroelectric thin-film transistor: two-level switching



- Large G_{max}/G_{min} ratio
- Relatively fast switching, considering the switching of all FE domains

Summary of progress and future work

- Progress:
  - ➢Well-behaved ferroelectric HZO and channel material ITO with CMOS compatibility
  - ➢Demonstration of high-performance ferroelectric thin-film transistor (FeTFT)
- Future work:
  - ➢Detailed study of synaptic behavior in FeTFT
  - ➢Detailed study of gradual domain switching in FeTFT
  - ➢Simulation and modelling of HZO multi-ferroelectric-domain dynamics