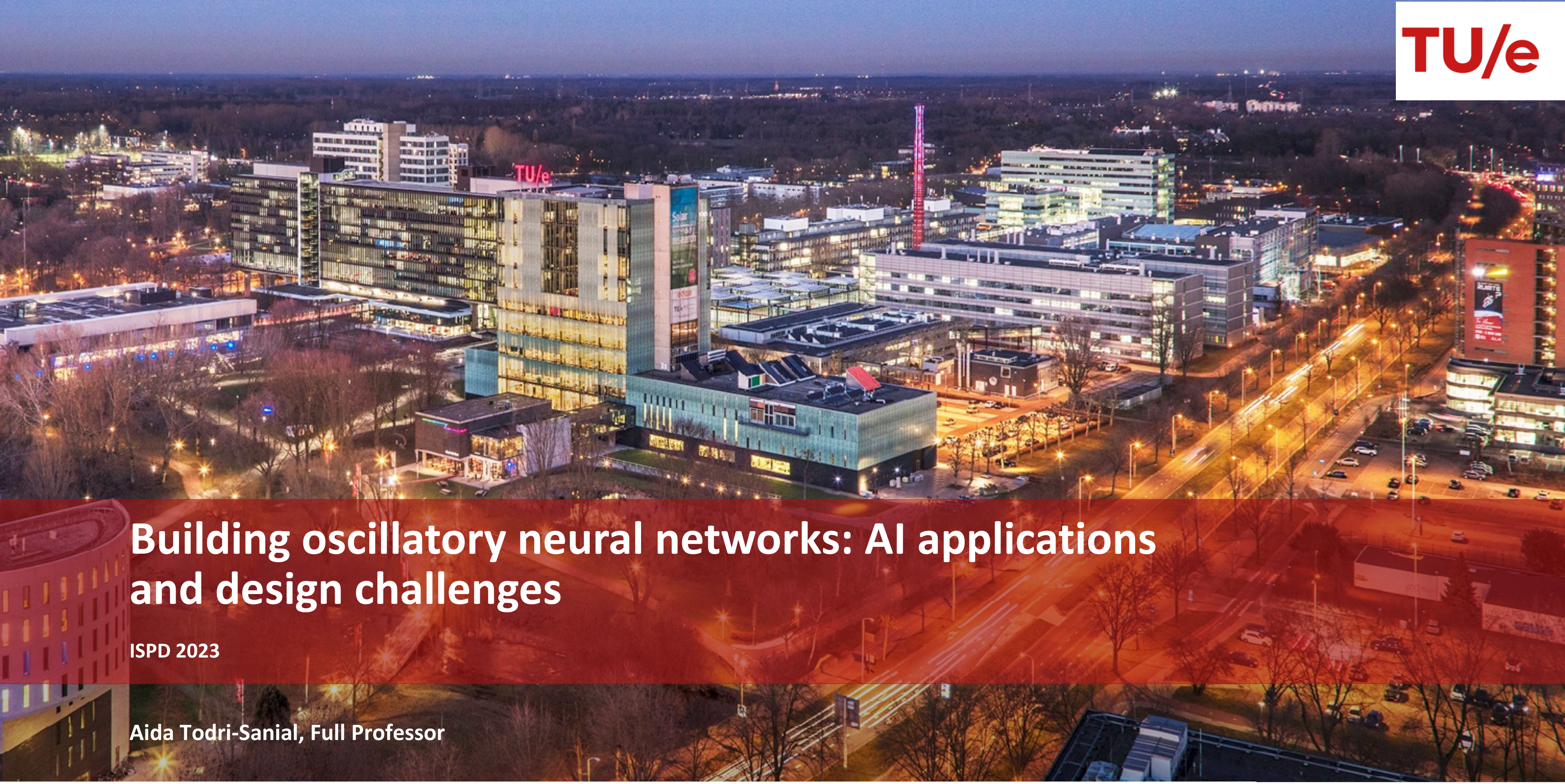


Building oscillatory neural networks: AI applications and design challenges

ISPD 2023

Aida Todri-Sanial, Full Professor

Electrical Engineering Department

EINDHOVEN UNIVERSITY OF TECHNOLOGY

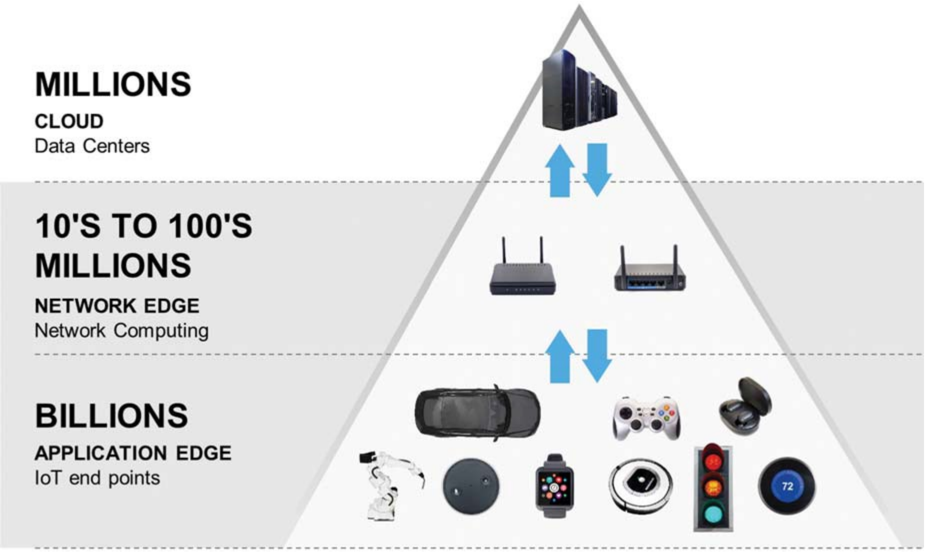
Outline – Take away messages

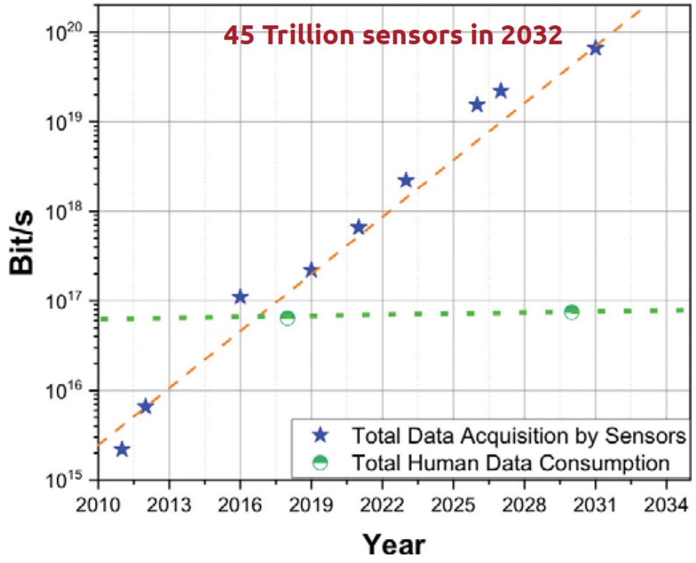
We need a novel computing paradigm that is energy efficient to enable edge AI computing for online learning and inference.

In this talk, I will present an alternative computing paradigm based on oscillatory neural networks.

How to compute with oscillators? – I'll cover some important aspects from materials, devices, architecture to applications.

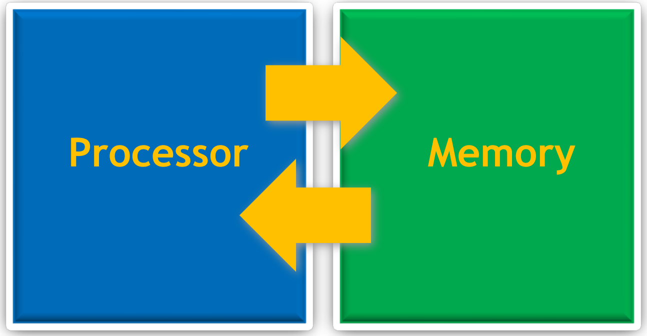
Trends – growing number of edge Iot

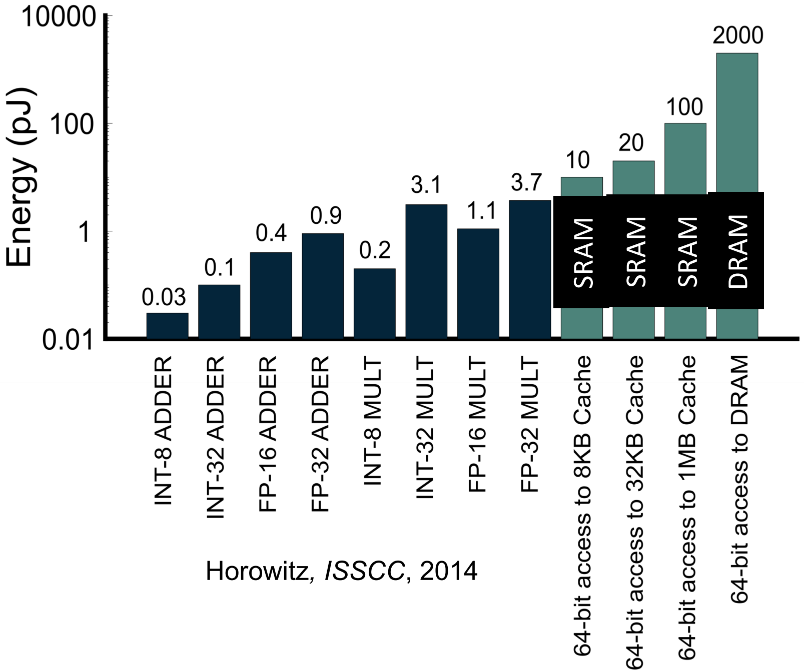




The cost of “moving data”

Conventional von Neumann computing architecture

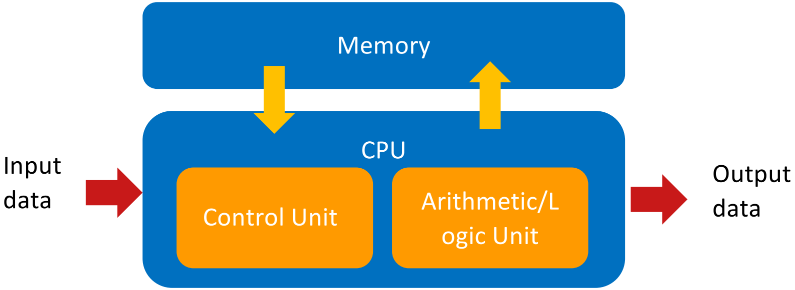


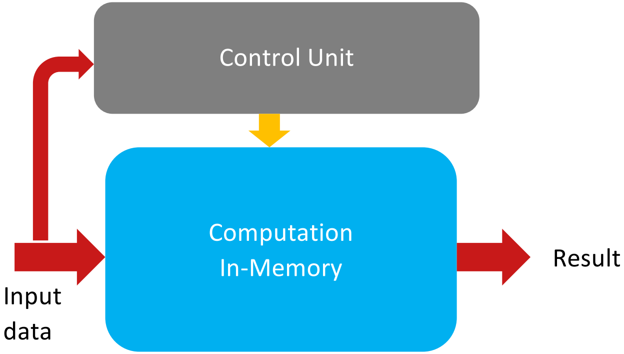


Horowitz, *ISSCC*, 2014

Beyond von Neumann architecture

~ 20W

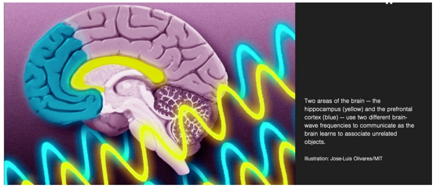




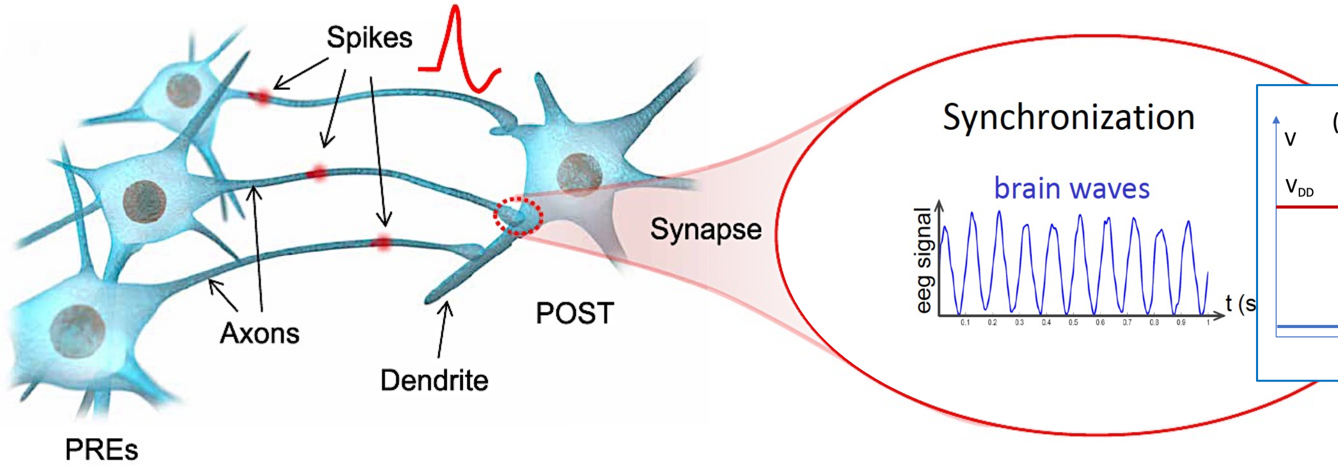
What are Neurons?

- Biological neural networks have rhythmic behavior, they spike and fire:
- ubiquitous in neocortex, hippocampus and olfactory bulb

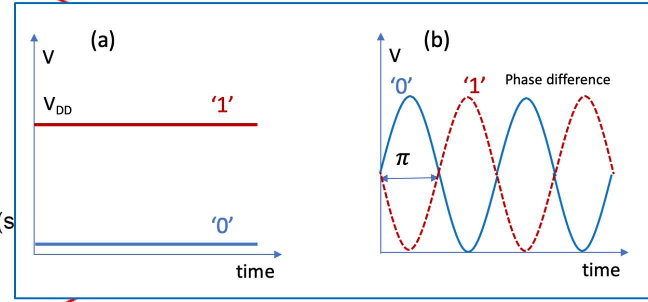
neuroscience



How brain waves guide memory formation
Neurons hum at different frequencies to tell the brain which memories it should store.



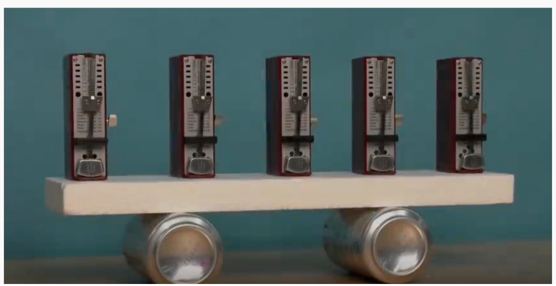
Computing in Phase



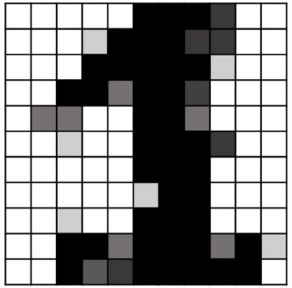
Opportunity: rather than potential voltage over time, we leverage **neurons biological behavior as the frequency of oscillations**

Computing with Oscillators

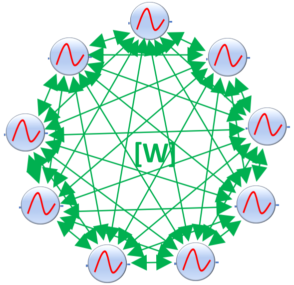
www.neuronn.eu



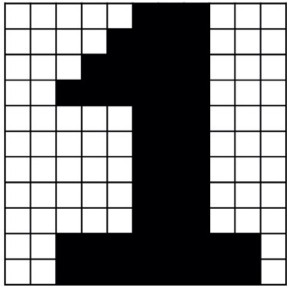
Each pixel is an oscillator



Coupling weights between oscillators is the memory



Output encoded on phase difference between oscillators



Associative memory

Neuromorphic Computing

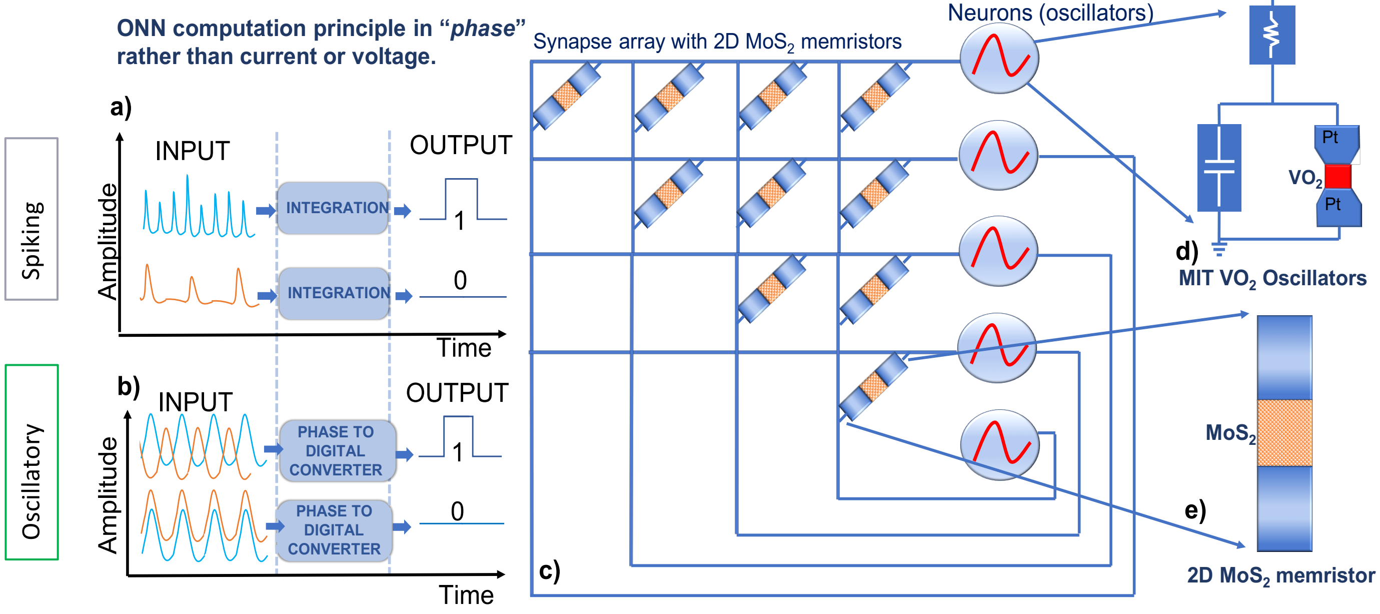
EU H2020 NEURONN
(Grant No 871501)

www.neuronn.eu

Computing in Phase – Oscillatory Neural Networks

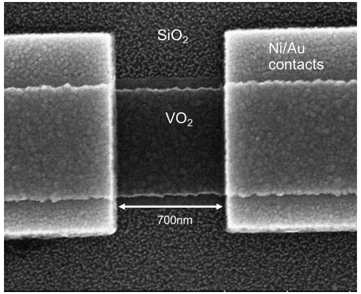
ONN Implementation for 5 neurons

ONN computation principle in “*phase*” rather than current or voltage.



Neurons -> VO_2 devices

Synapse -> MoS_2 devices

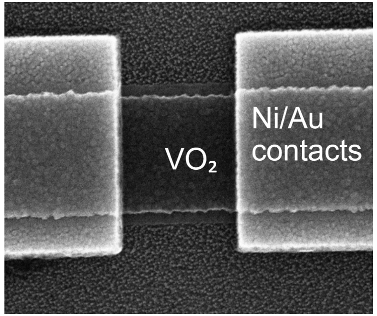


E. Corti, et al. IEEE EDL, 2019.

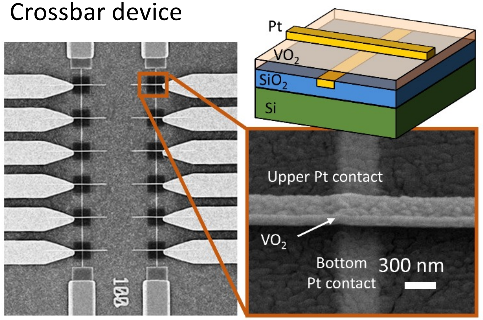
A look into materials & Devices

VO_2 Insulator-Metal Transition

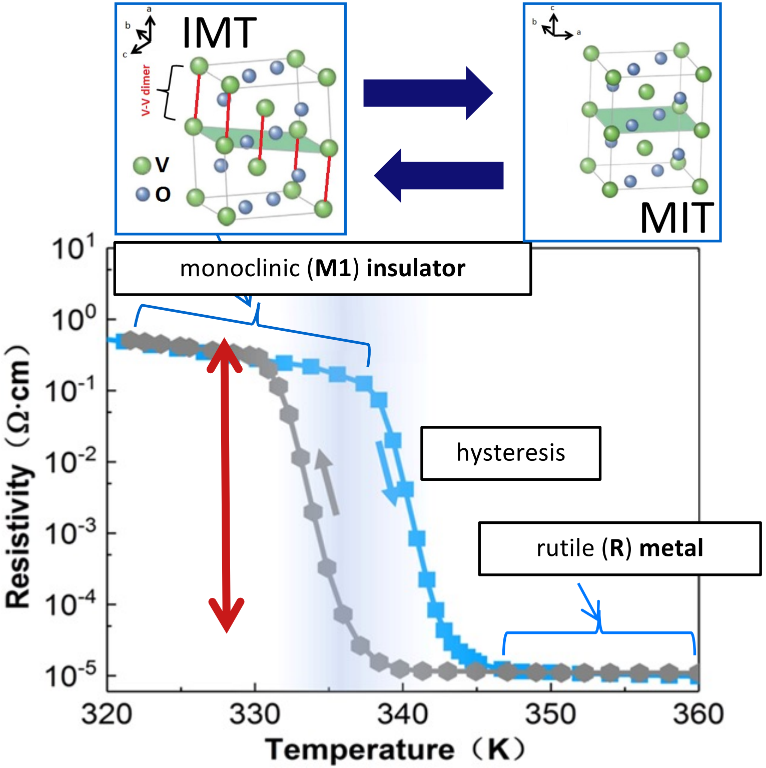
Planar device



Crossbar device



E. Corti, et al, IEEE EDL 2020
E. Corti et al, SSE 2020



Phase-change triggered by Joule Heating

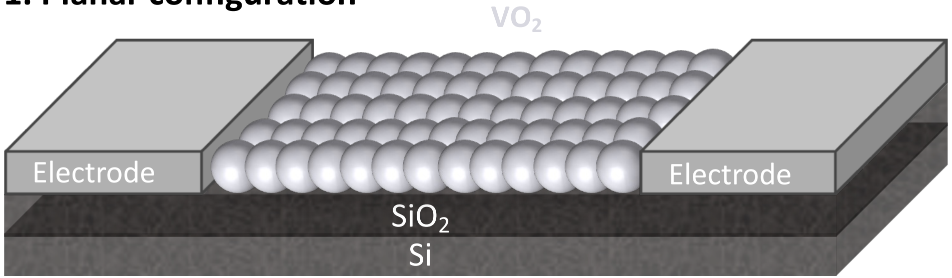
Insulator to metal transition at 68°C (340K)

Fabrication of VO_2-based oscillators

Two types of device configurations studied:

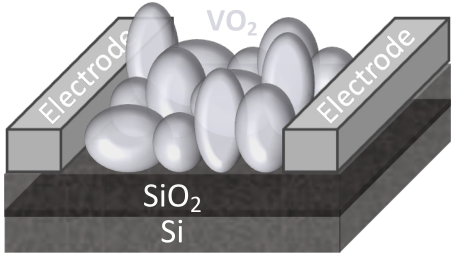
- Integration on Si to allow for CMOS processes compatibility
- Nanoscale devices for low-power and high frequency operation

1. Planar configuration

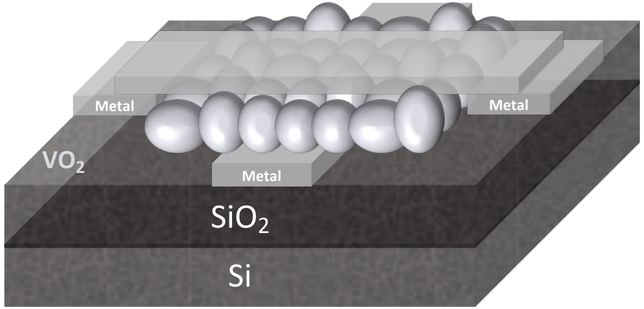


From µ-scale to nanoscale

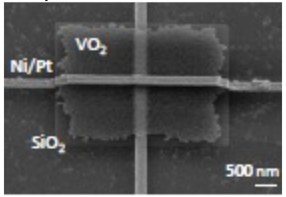
Growth on SiO_2 introduces granular films



2. Crossbar configuration

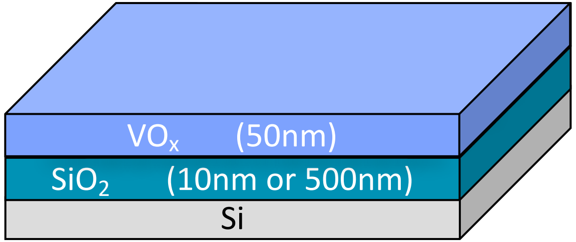


- VO_2 sandwiched between top and bottom electrodes
- Goal:
  - Smooth film
  - Thin monolayer of grains
  - Grains densely positioned

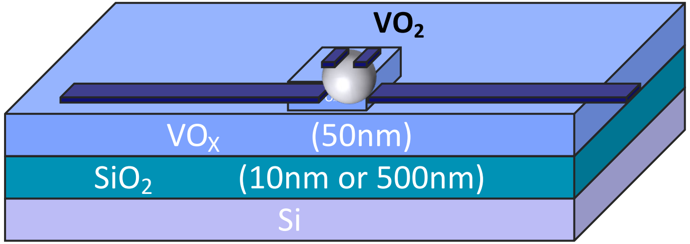


E. Corti, S. Karg, IBM

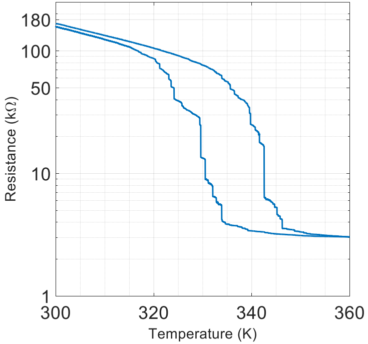
Annealing results

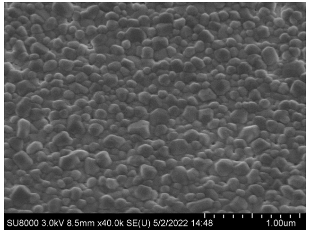


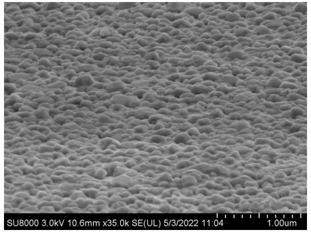
Granular film



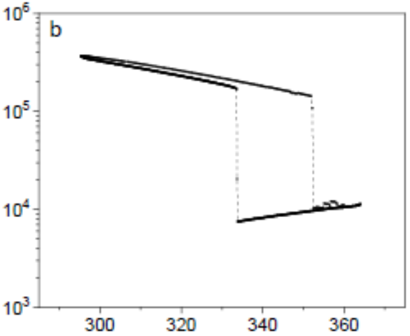
Large single grains
(isolated nanoparticles)



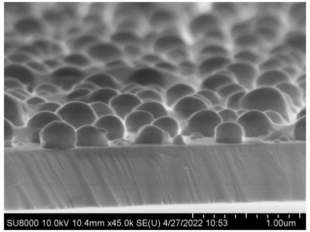


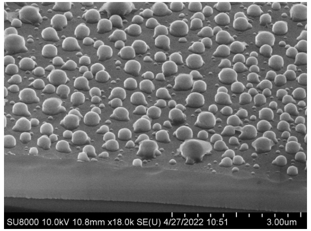


Granular film response

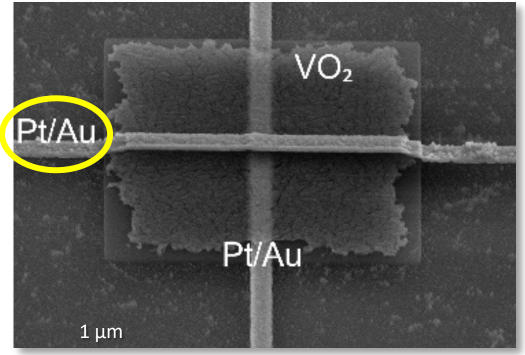


Nanoparticle response

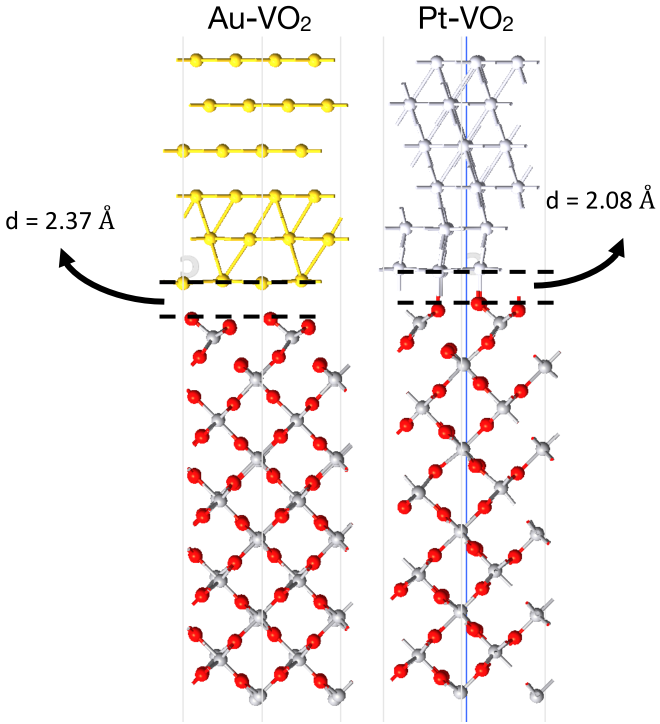




QM Simulations on Vo_2-Metal Contacts



Experimental cross-bar device developed in NeurONN.



DFT-optimised structures of VO_2 interface models with Au and Pt metal contacts.

VO_2 forms closer contact with Pt compared with Au

Evidence of chemical reactivity Pt-VO_2

Thermal resistance Pt-VO_2 expected to be lower

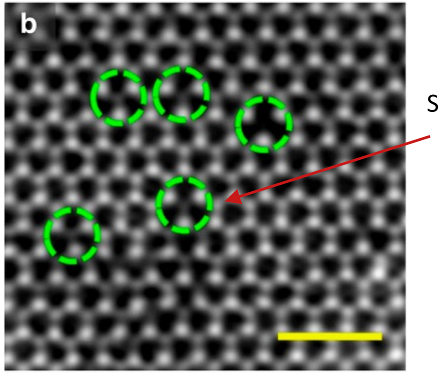
*S. Carapezzi, et al., IEEE JETCAS 2021

Investigation of MoS_2 memristive mechanism

Several mechanisms have been proposed in the literature depending on the device architecture.

Surface vacancies and impurities in MoS_2 seem to play the central role in the HRS-LRS switch.

Typical concentrations are high (~10^{12}-10^{13} cm^{-2}).



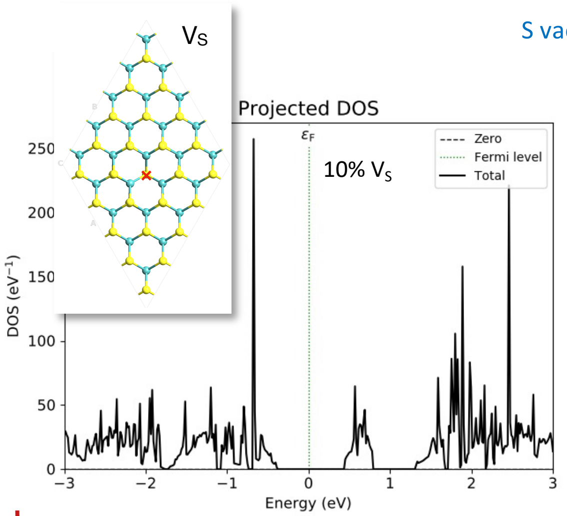
ADF-STEM characterisation of defects

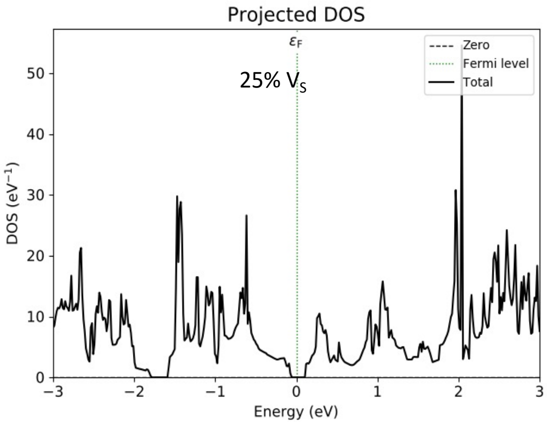
Effect of defect concentration

Defect concentration from **low** coverage (2%) up to **high** coverage (10%) and **very high** coverage (25%).

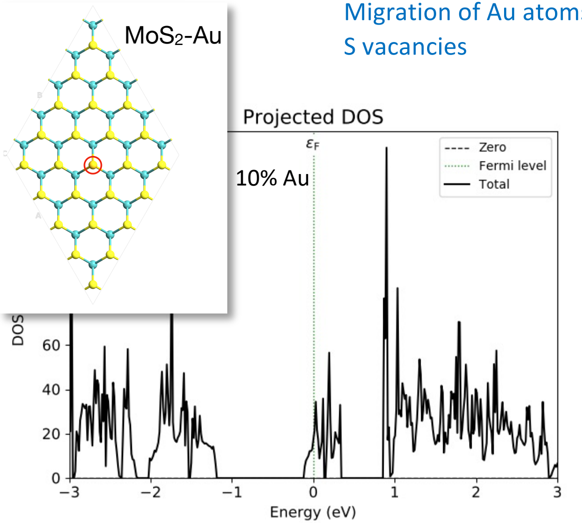
At high coverage S vacancies not enough to close the material's band gap, whereas with **Au substitution the material turns metallic!**

S vacancies





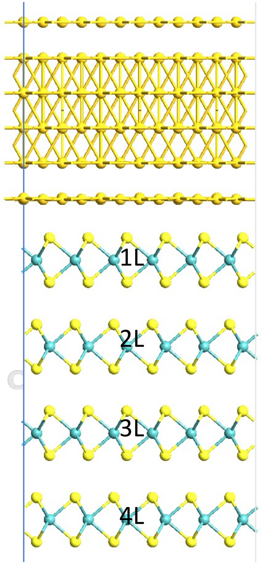
Migration of Au atoms to fill S vacancies

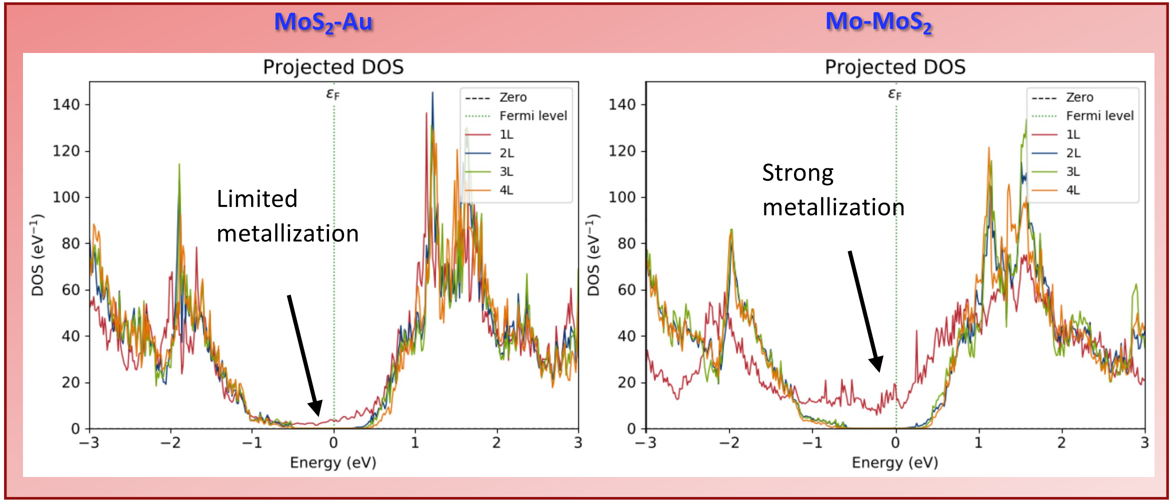


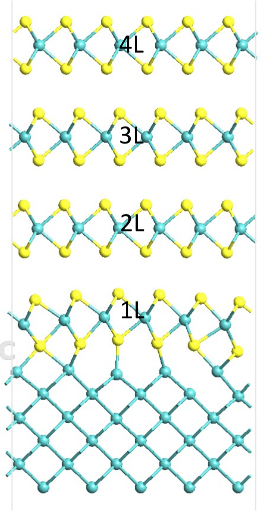
Semiconductor

Metallic

Effect of MoS_2 layer thickness







Strong vs. limited **metallization of MoS_2**.
Effect of metal electrode limited to the topmost MoS_2 layer.
Monolayer MoS_2 not sufficient to achieve memristive behavior!

G. Boschetto, S. Carapezzi, A. Todri-Sanial, International Conference on Materials Chemistry, July 2021

A look into Oscillator devices

Simulation models & tools

Atomistic, TCAD to compact modeling and simulation
Develop predictive models to explore manufacturing and design space.

Quantum mechanics

- DFT: Investigate contacting MoS_2 and VO_2 materials

Continuous models

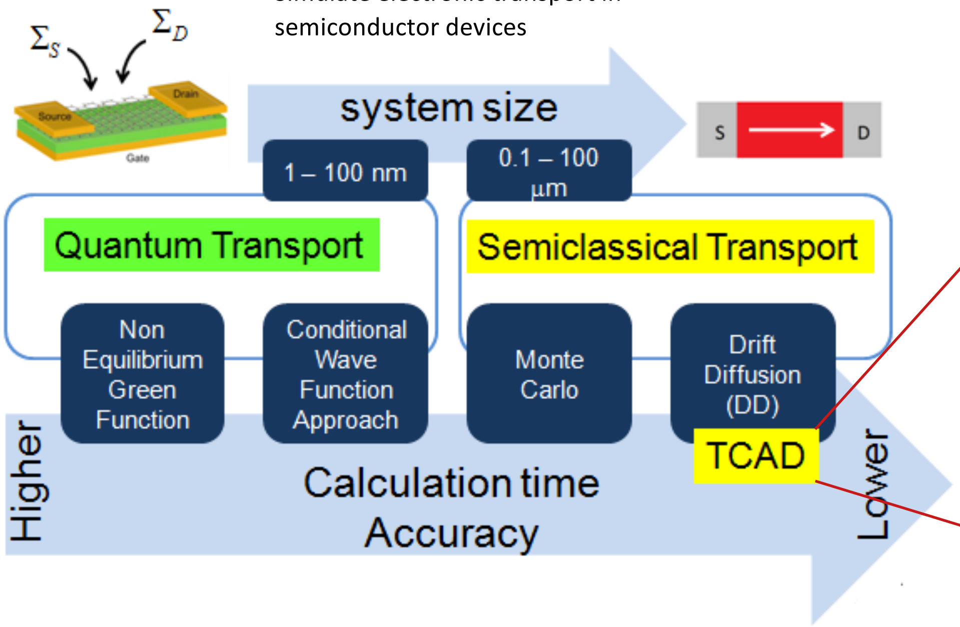
- TCAD: Device geometry impacts carrier/temp transport & resistivity

Circuits

- Compact models: Implement salient device physics for circuit and system design

Technology Computer-Aided Design (TCAD)

Simulate electronic transport in semiconductor devices



Self-Consistent Solution of Equations:

- Poisson's Equation: $-\varepsilon_s \nabla^2 \psi = \left(p - n + N_D^+ - N_A^-\right)$
- Continuity Equation:

$$\frac{\partial n}{\partial t} = \frac{1}{q}\vec{\nabla} \cdot \vec{J}_n + R_n$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\vec{\nabla} \cdot \vec{J}_p + R_p$$

R_n, R_p : the net generation - recombination rate of e, h

- Current Equation:

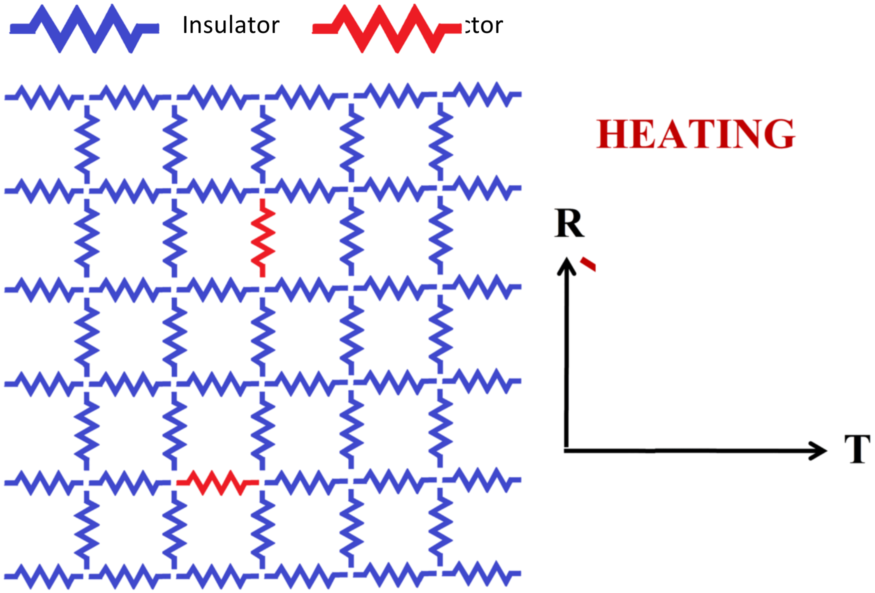
$$\vec{J}_n = qn\mu_n \vec{E} + qD_n \vec{\nabla} n$$

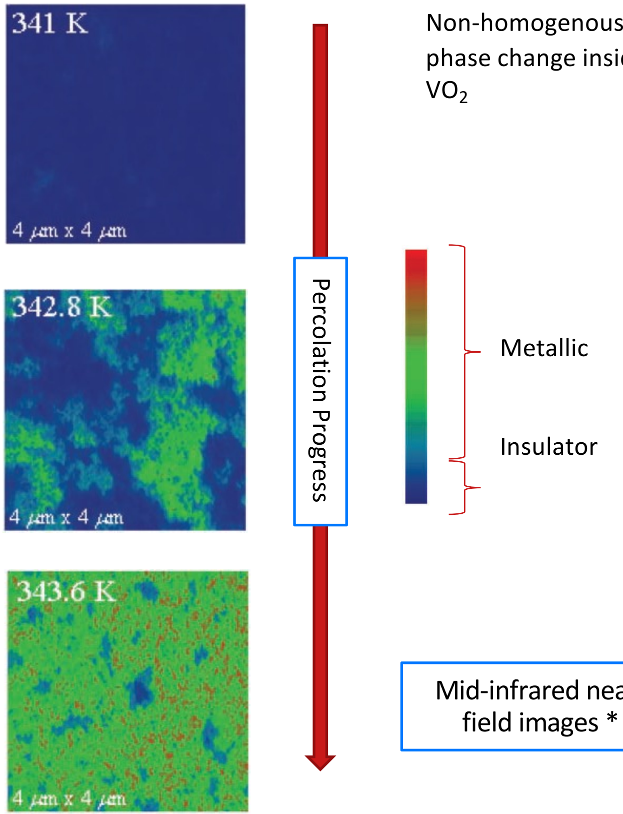
$$\vec{J}_p = qp\mu_p \vec{E} - qD_p \vec{\nabla} p$$

n, p : electron, hole density; μ : low - field mobility; D : diffusion coefficient

* After E. G. Marinn et al., IEEE Transaction on Electron Devices, Vol. 65, No. 10, 2018

VO_2 as a resistor network



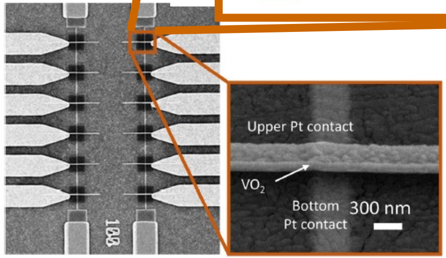


Non-homogenous phase change inside VO_2

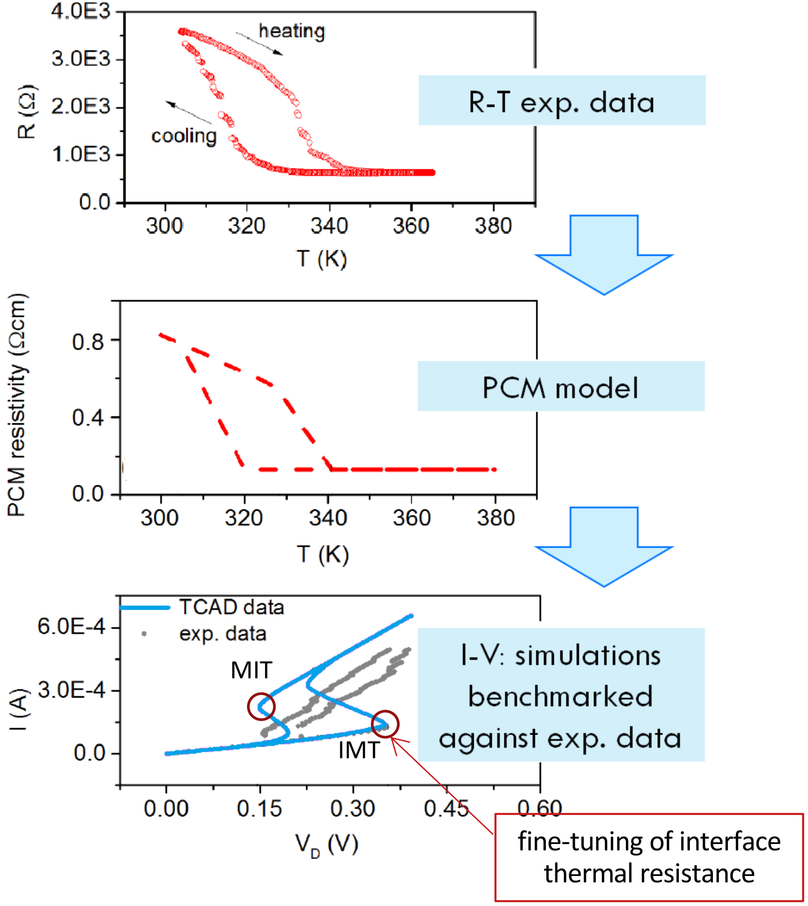
Mid-infrared near-field images *

* Qazilbash et al. Science 2007, Vol. 318, Issue 5857, pp. 1750-1753

CALIBRATION AND VALIDATION OF TCAD MODEL

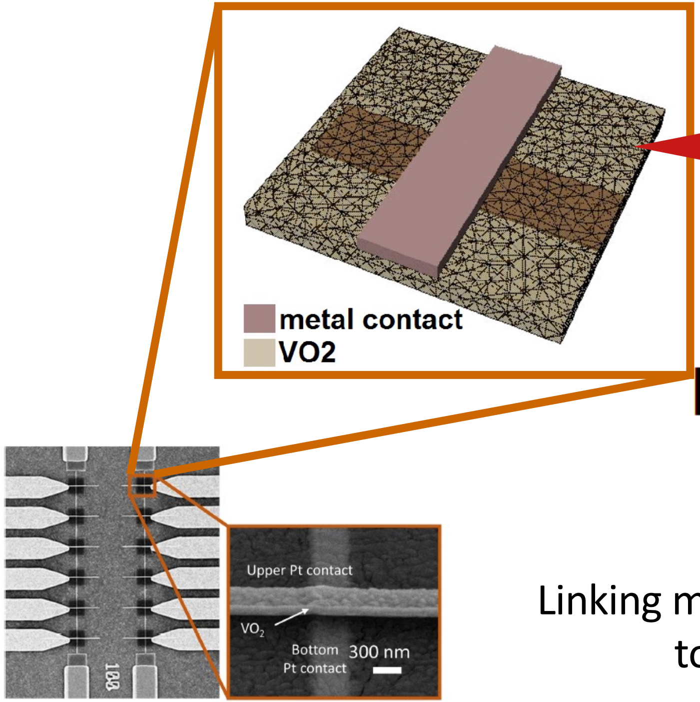


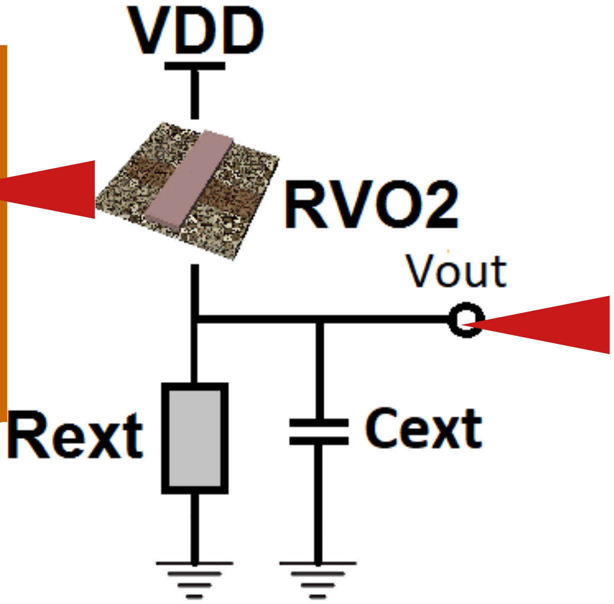
CB better control of switching region

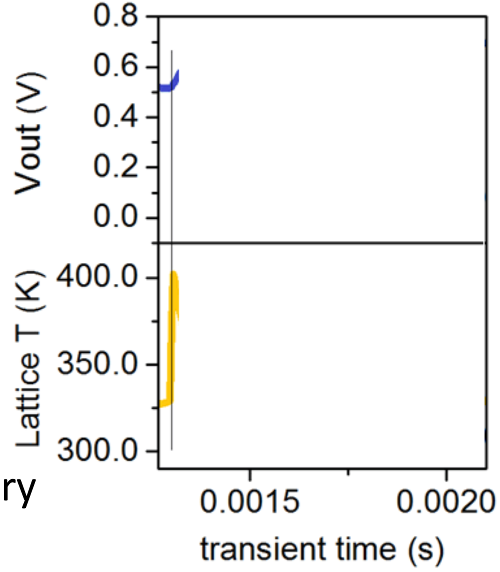


E. Corti et al., Frontiers in Neuroscience, vol. 15, p. 19, 2021
S. Carapezzi et al., *IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2021*
S. Carapezzi et al. 2021 19th IEEE International New Circuits and Systems Conference (NEWCAS)

TCAD SIMULATION OF VO_2 OSCILLATOR







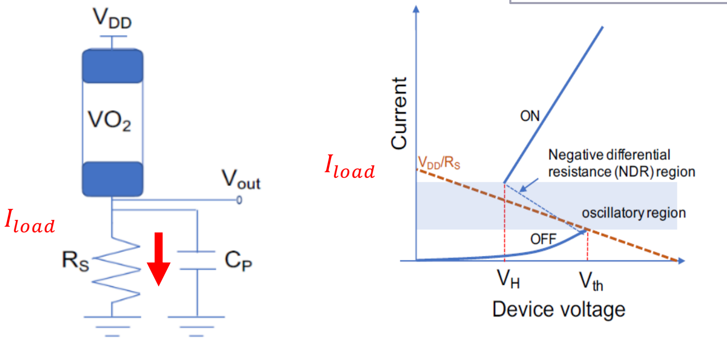
Linking material properties, device geometry to oscillator's figures of merit

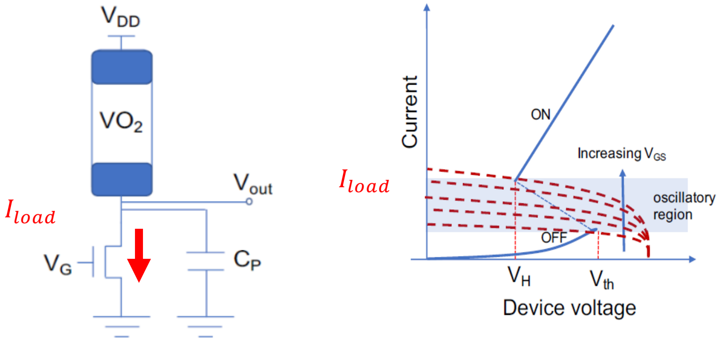
E. Corti et al., Frontiers in Neuroscience, vol. 15, p. 19, 2021;
Carapezzi et al., *IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2021;* Carapezzi et al. 2021 19th IEEE International New Circuits and Systems Conference (NEWCAS)

From Device to Circuit-level Simulation

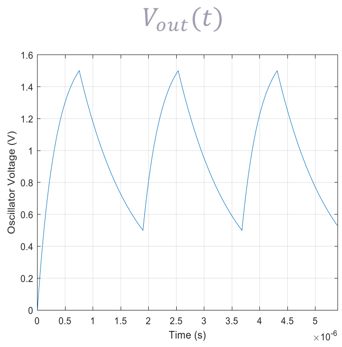
Oscillatory Circuit for ONN Implementation

Single Oscillator Case





2 different methods for oscillator load



Oscillator output voltage with distinctive:

- Period, T_{osc}
- Signal amplitude (V_L, V_H)

VO_2 oscillates in its NDR region

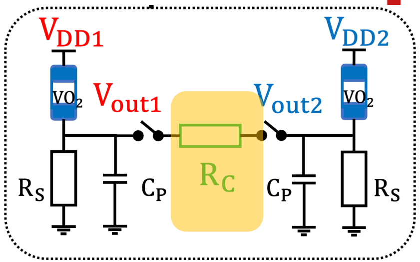
VO_2 oscillator dynamics:

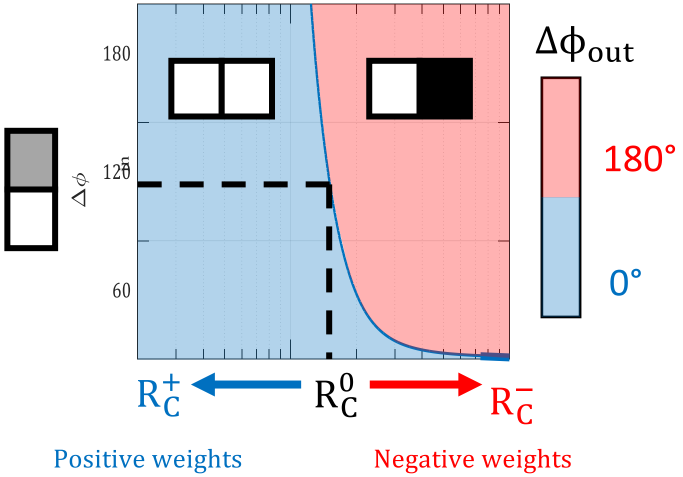
$$\begin{cases} \dfrac{dG^{VO_2}(t)}{dt} = f\left(G^{VO_2}(t), V_{out}\right) \\ C_P \dfrac{dV_{out}}{dt} = \left(V_{DD}(t) - V_{out}\right) G^{VO_2}(t) - I_{load} \end{cases}$$

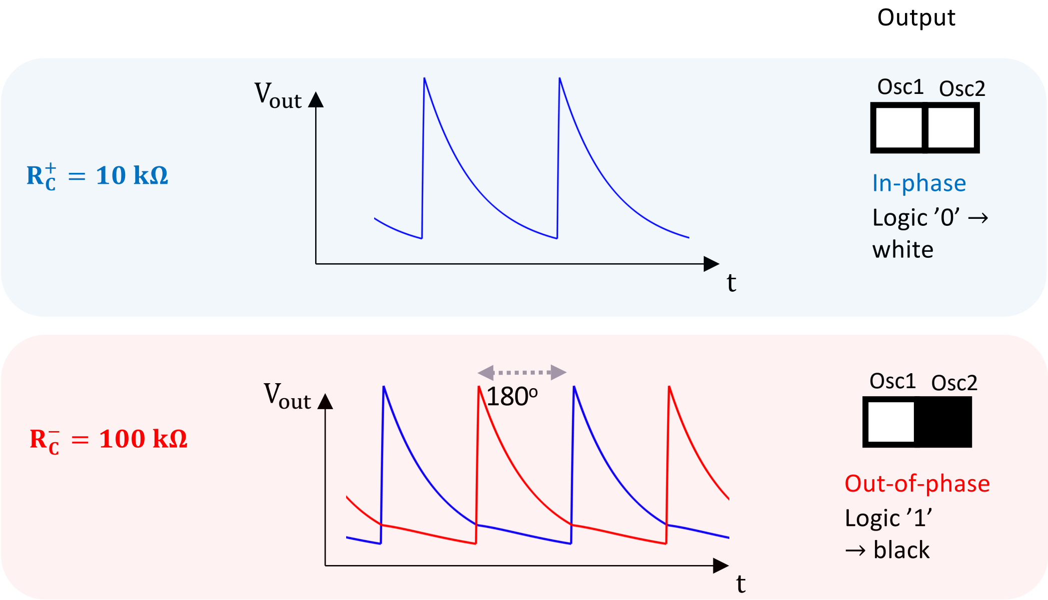
A. Todri-Sanial et al., TNNLS 2021

How to set Synaptic resistors ?

Coupling resistor R_C can implement synaptic weight



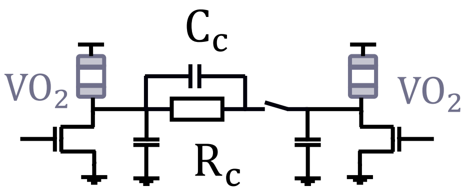


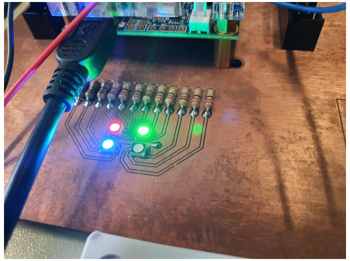


Mapping Hebbian Learning Rules to Coupling Resistances for Oscillatory Neural Networks, C. Delacour and A. Todri-Sanial, Frontiers of Neuroscience, 2021

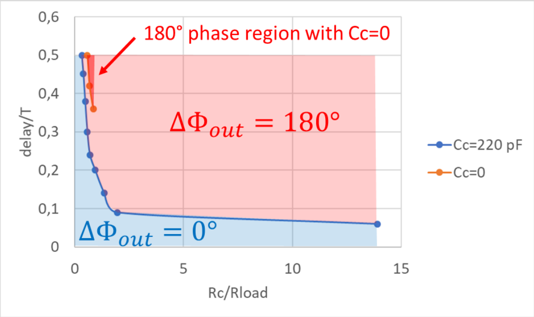
Role of synaptic capacitors

Adding coupling capacitors enhances ONN performances

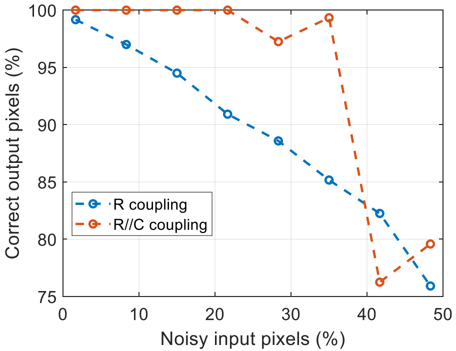




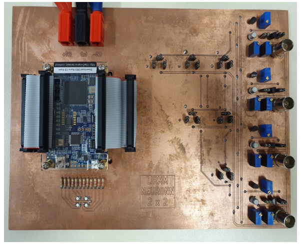
Robust phase locking



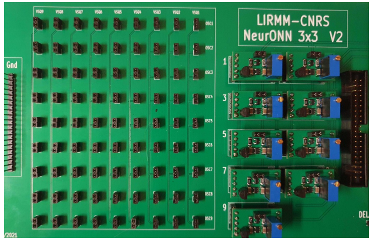
Higher accuracy



Demonstration on 2x2 ONN PCB

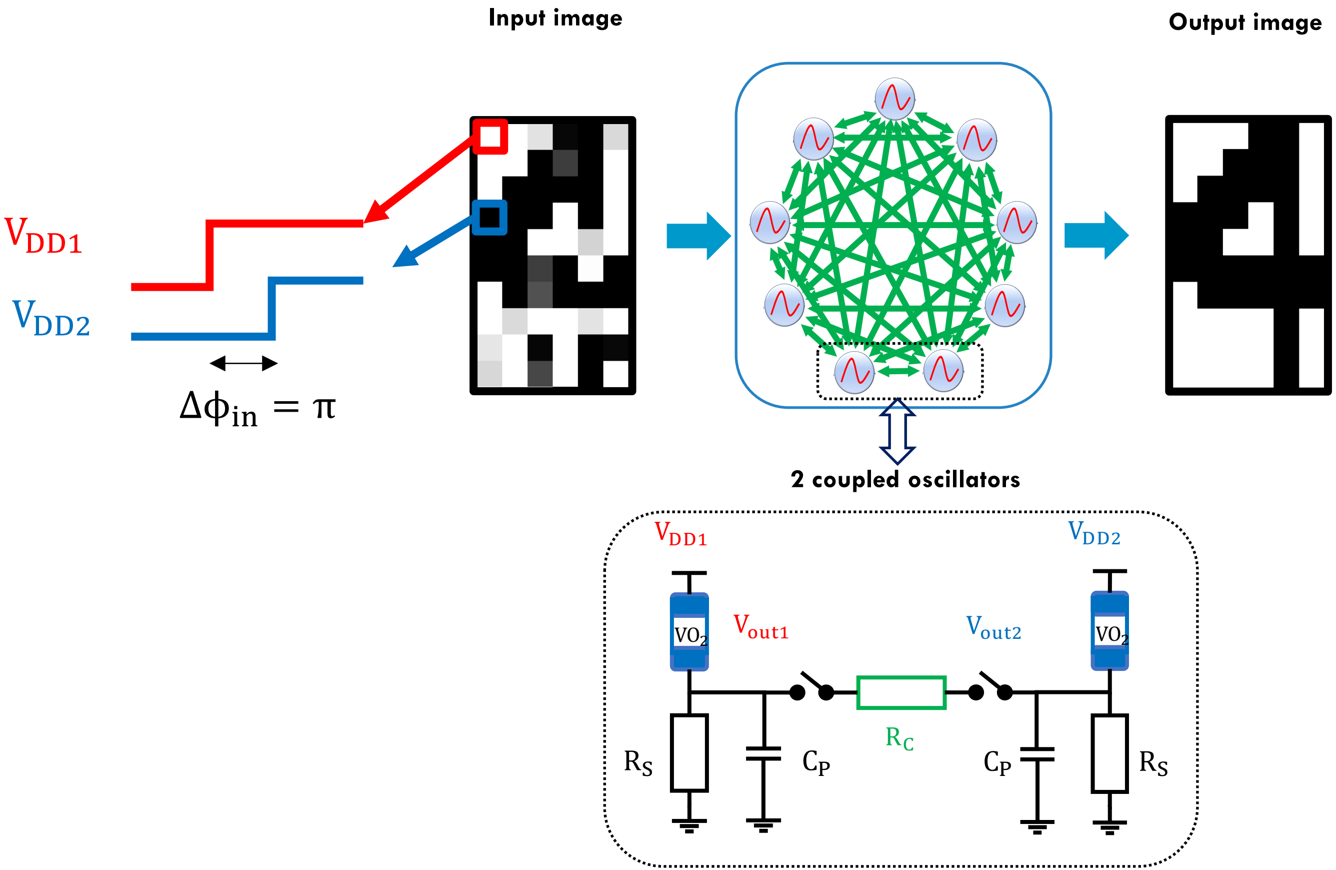


2x2 ONN

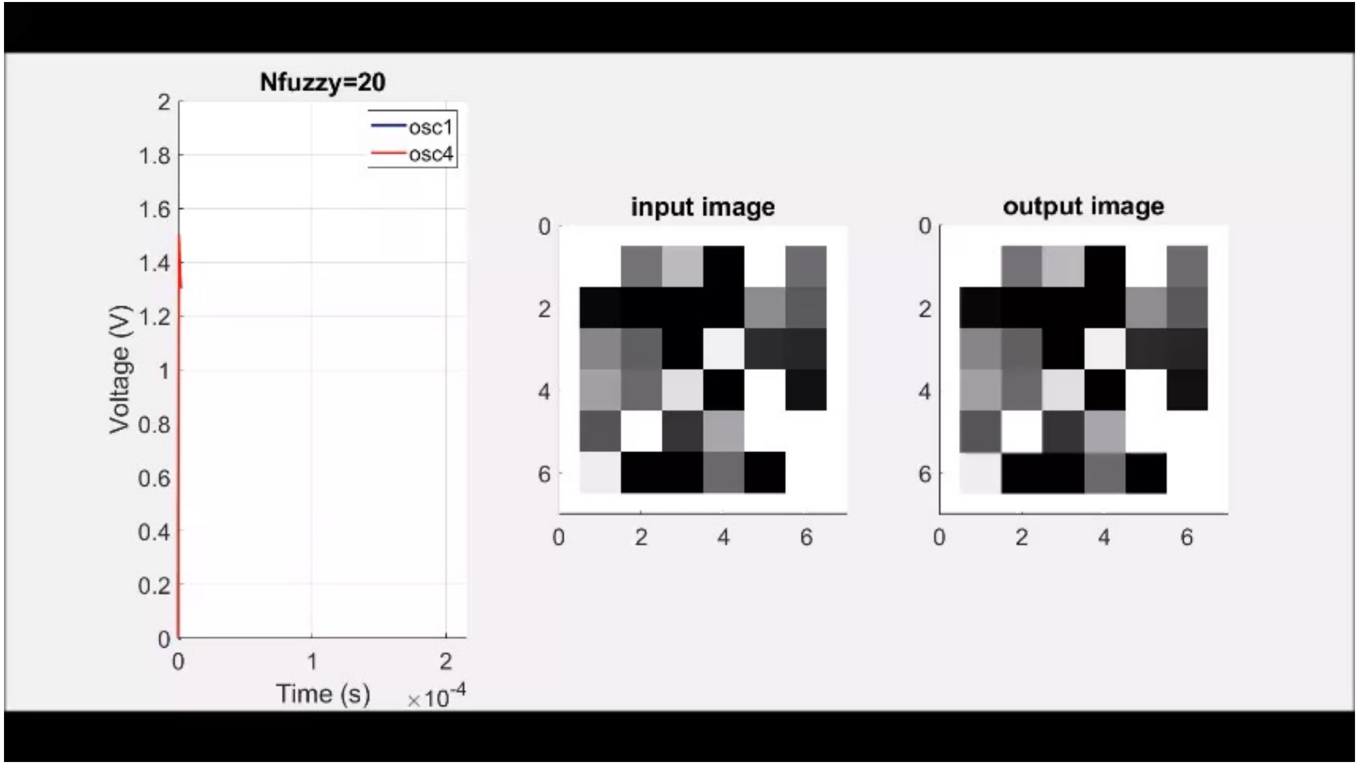


3x3 ONN

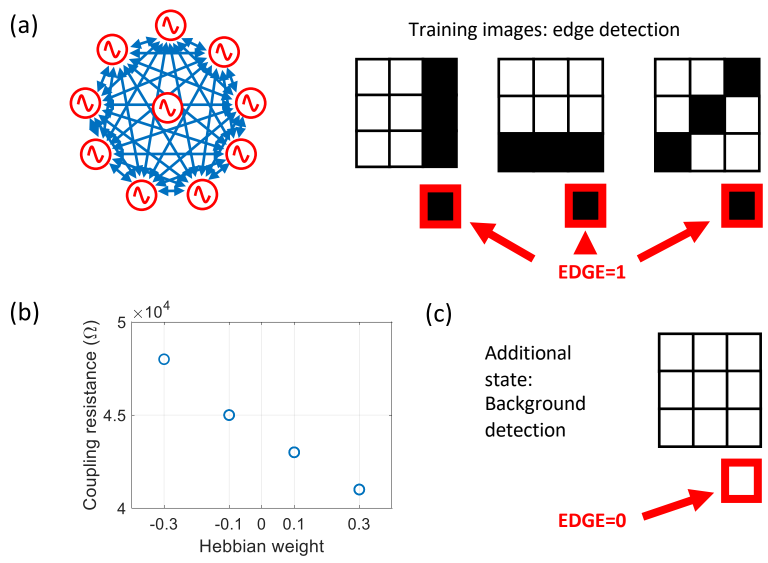
Oscillatory Neural Network

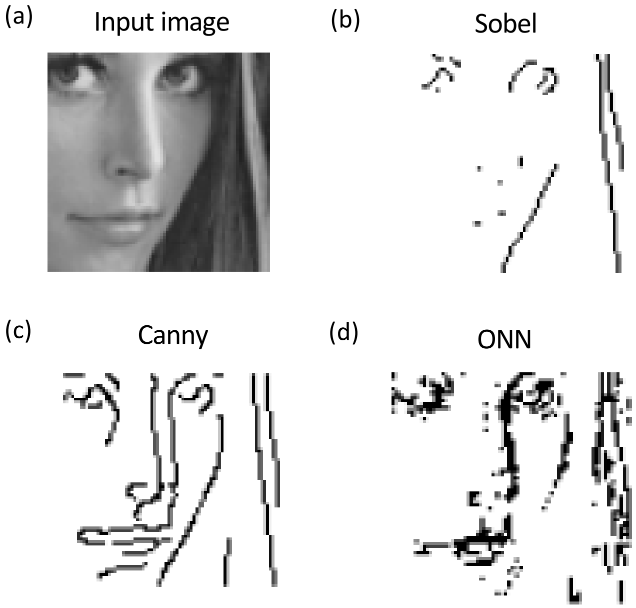


ONN phase dynamics



Edge image detection with analog ONN



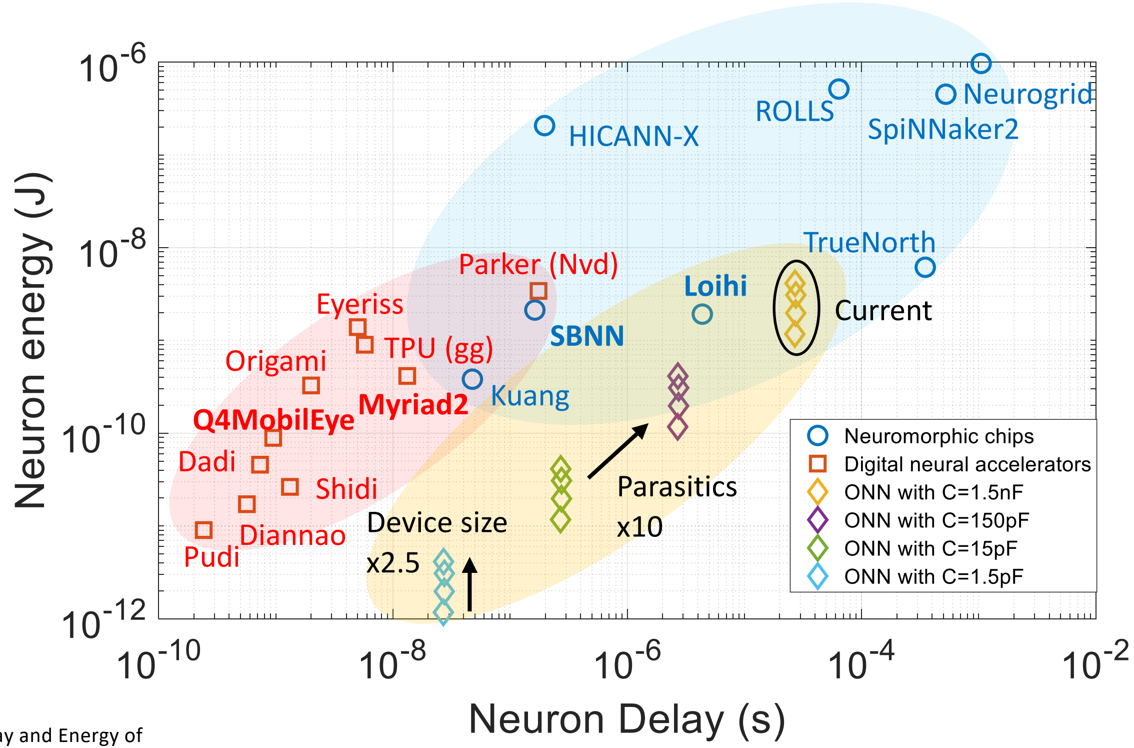


ONN benchmarking

C. Delacour, A. Todri-Sanial, TNNLS 2022

Benchmarking ONN with state of the art chips*

VO$_2$ size: from 2 µm to 5 µm
Vdd: from 0.5 to 0.9 V
Parasitics: from 1.5 pF to 1.5 nF
Oscillator delay= N_{cycles}=4
Optimistic:
  No external circuitry taken into account
Observation: ONN could serve as an accelerator or a general purpose neuromorphic hardware

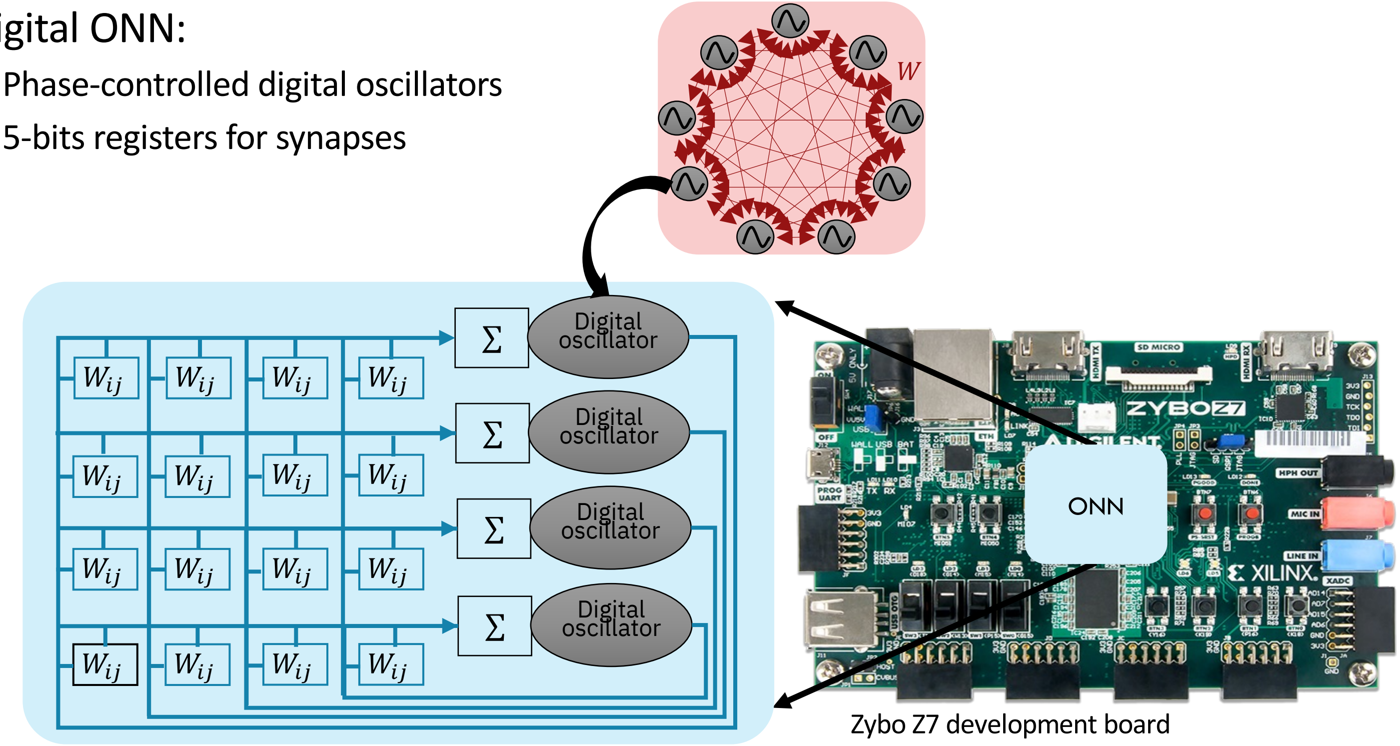


*D. E. Nikonov and I. A. Young, "Benchmarking Delay and Energy of Neural Inference Circuits," in IEEE Journal on Exploratory Solid-State Computational Devices and Circuits, vol. 5, no. 2, pp. 75-84, Dec. 2019

Benchmarking and Applications

Digital Onn on FPGA

- Digital ONN:
  - Phase-controlled digital oscillators
  - 5-bits registers for synapses

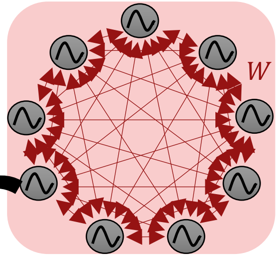


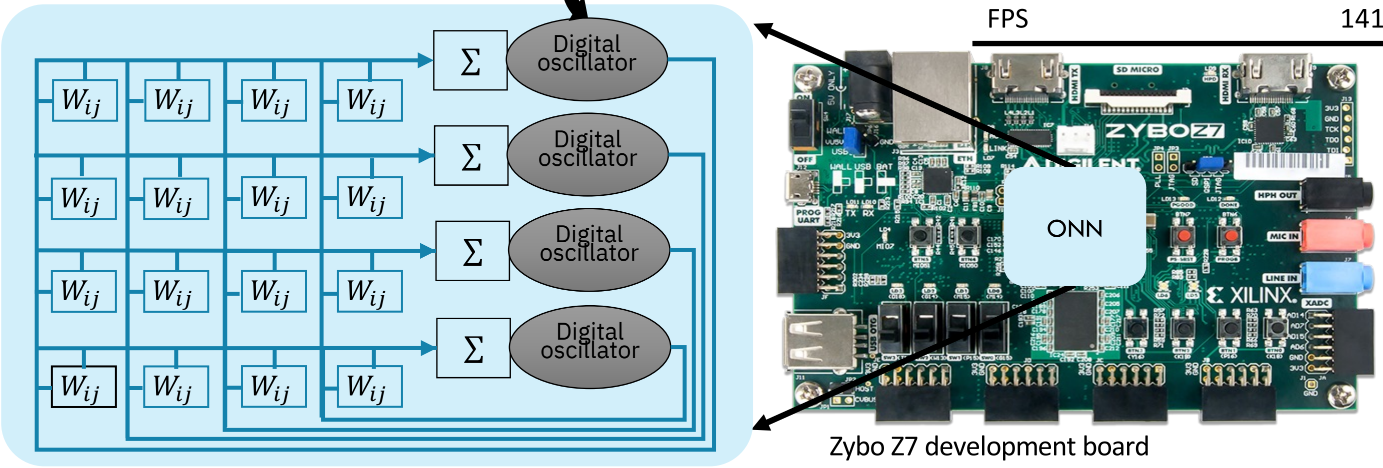
Zybo Z7 development board

M. Abernot *et al.*, “Digital Implementation of Oscillatory Neural Network for Image Recognition Applications,” Front. In Neuroscience, 2021.

Digital ONN on FPGA

- Digital ONN:
  - Phase-controlled digital oscillators
  - 5-bits registers for synapses





Zybo Z7 development board

ONN size	5x3	10x6
Synapses	225	3600
FPGA resources (LUTs)	1,8%	12%
FPGA resources (Flip-flops)	0,68%	2,6%
Init. Time (us)	2	7,8
Comp. Time (us)	5,2	5,4
FPS	141000	75000

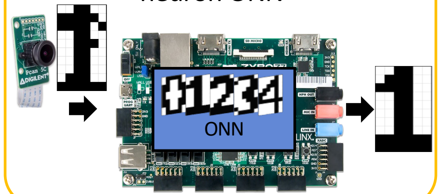
M. Abernot *et al*., "Digital Implementation of Oscillatory Neural Network for Image Recognition Applications," Front. In Neuroscience, 2021.

Onn applications

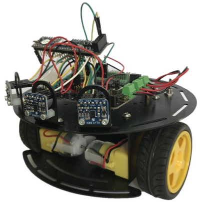
Implemented on the digital ONN

Simulated with the digital ONN target

Digits recognition with 60-neuron ONN

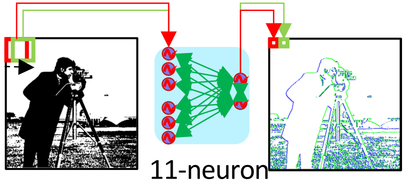


Obstacle avoidance



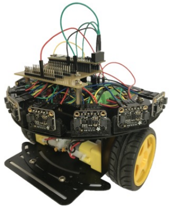
15-neuron ONN

Edge detection



11-neuron ONN-HAM

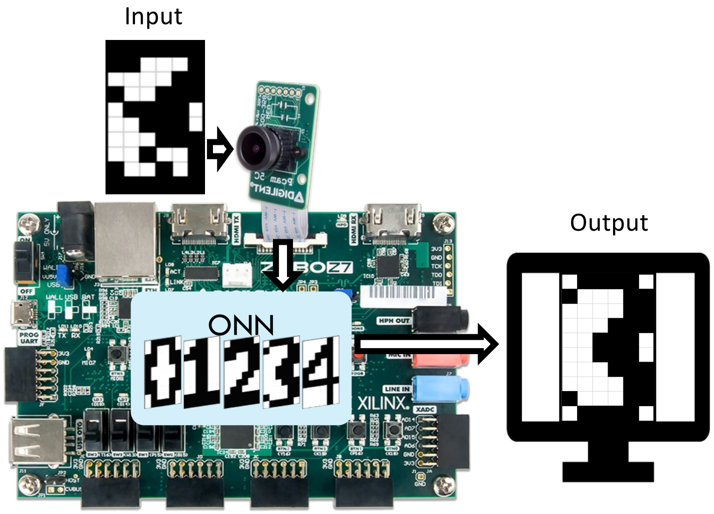
Obstacle avoidance



2 cascaded ONNs:

- 40-neuron ONN
- 8-neuron ONN

ONN for image recognition



10x6 ONN:

5 training patterns: digits from 0 to 4

Training: Unsupervized learning rule (ex: Diederich Opper I)

Input image from a camera stream

Noisy images with flipped pixels (10, 20, 30)

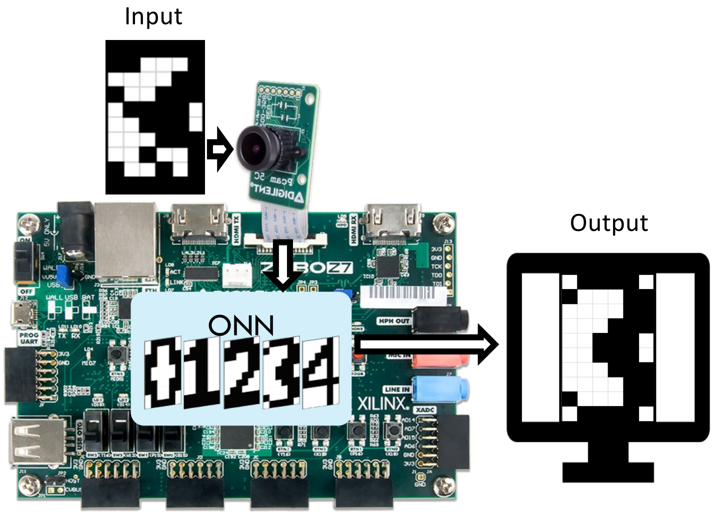
Output image printed on a screen

Table: Accuracy for different input noise level

Input Image noise	10 Flips	20 Flips	30 Flips
Hebbian	86 %	59%	21%
Diederich Opper I	100 %	93,6%	41,8%

M. Abernot *et al.*, "Digital Implementation of Oscillatory Neural Network for Image Recognition Applications," Front. In Neuroscience, 2021.

ONN for image recognition *

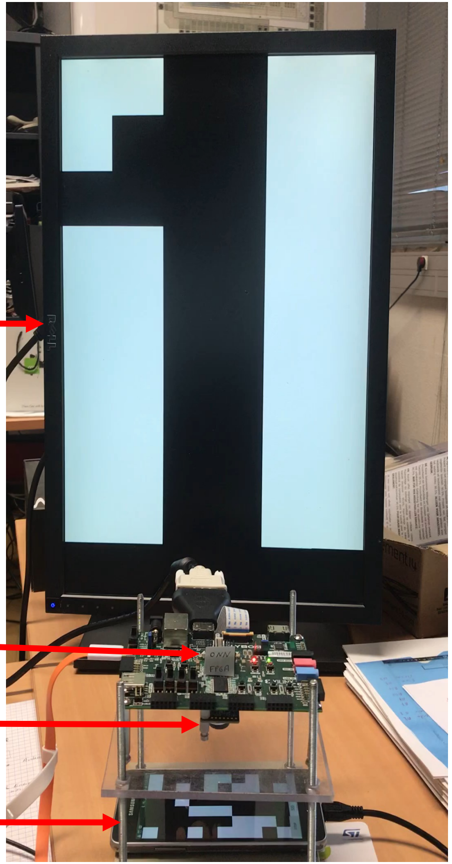


OUTPUT: HDMI screen

ONN-on-FPGA trained with
Diederich Opper

Camera to the FPGA

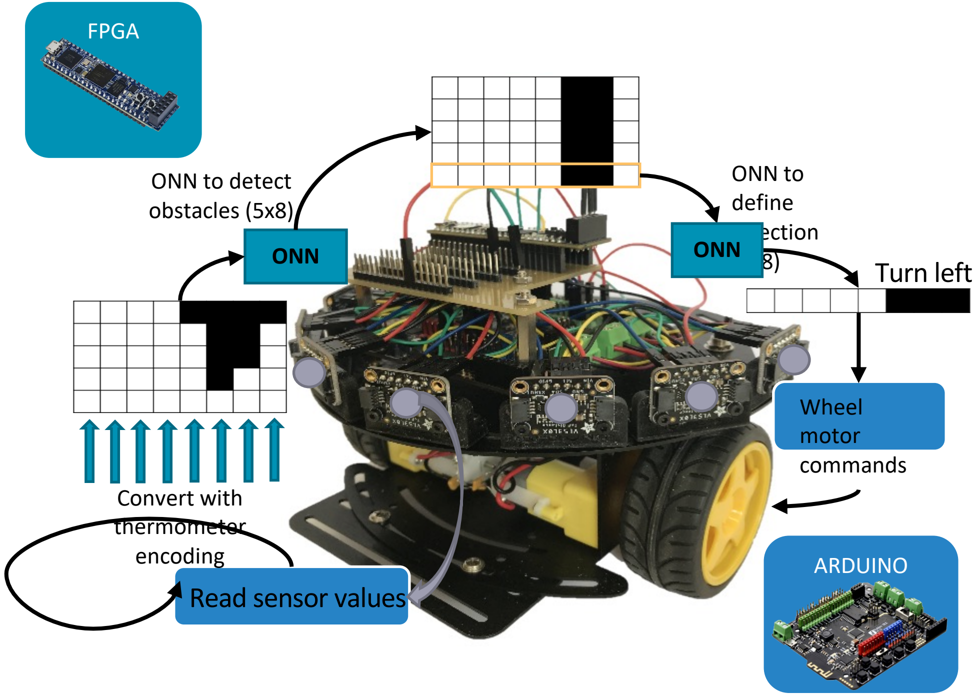
INPUT: Pictures on a phone with 20 flipped pixels noise



*M. Abernot *et al.*, “Digital Implementation of Oscillatory Neural Network for Image Recognition Applications,” Front. In Neuroscience, 2021.

ONN for obstacle avoidance

Arduino mobile robot
8 proximity sensors
2 cascaded ONNs
1 to detect obstacles: 5x8 trained with 256 patterns using Hebbian
1 to define direction: 1x8 trained with 16 patterns using Hebbian
Output of the 2nd ONN associated to robot direction
3 possible directions: left, front, right



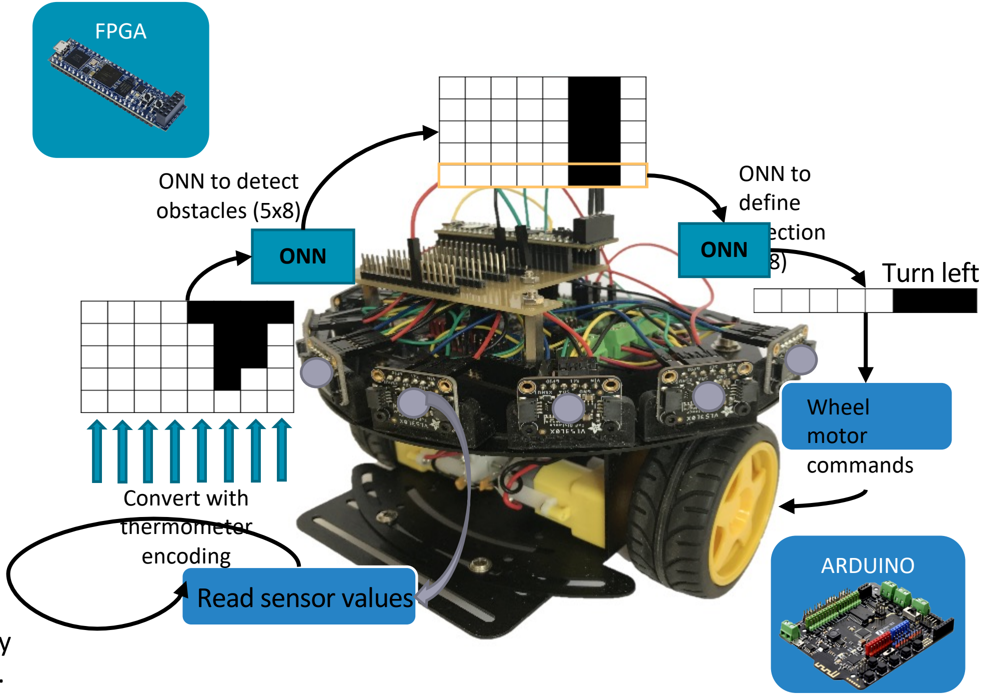
M. Abernot *et al.*, "Mobile Robot Obstacle Avoidance with Oscillatory Neural Networks on FPGA," IBM/IEEE AI Compute Symposium, 2021.

ONN for obstacle avoidance

Arduino mobile robot

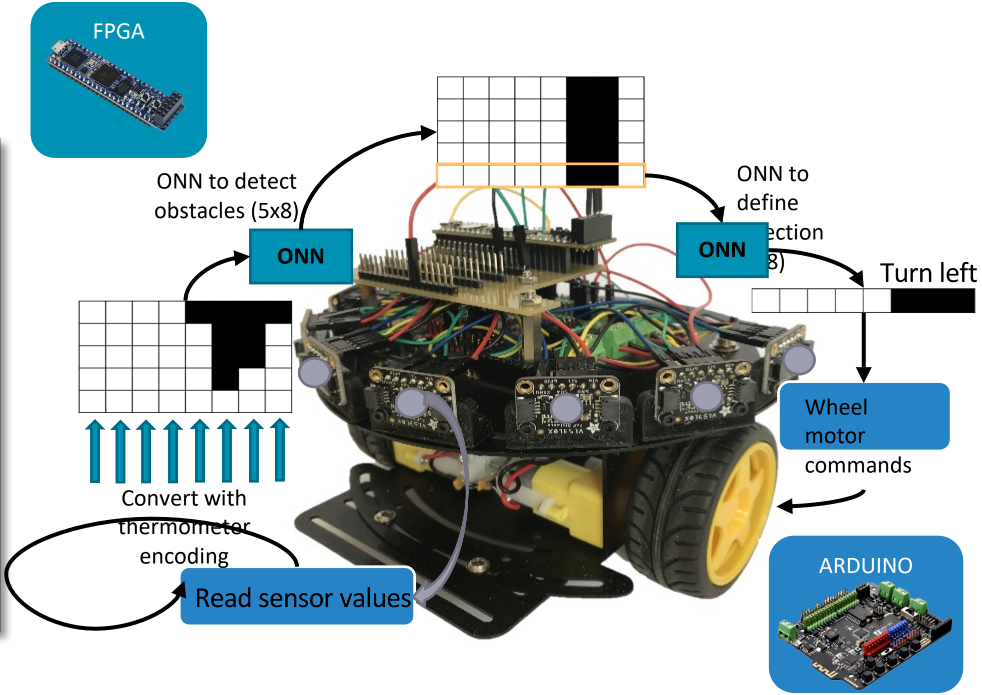
ONNs Results and Performances	ONN 5x8	ONN 1x8
Number of Training Patterns	256	16
LUTs (Artix 7: xc7a35)	11,5%	
Flip-Flops (Artix 7: xc7a35)	5,4%	
Frequency	12 MHz	12 MHz
ONN Computation time	24 us	17 us
Accuracy	100 %	74 %
Full system results and performances		
8-sensor measurement time	18 ms	
FPS	40	
Battery	6V/2850mAh	
Current consumption	700 mA	
Robot life time	4h	

M. Abernot *et al.*, "Mobile Robot Obstacle Avoidance with Oscillatory Neural Networks on FPGA," IBM/IEEE AI Compute Symposium, 2021.



ONN for obstacle avoidance

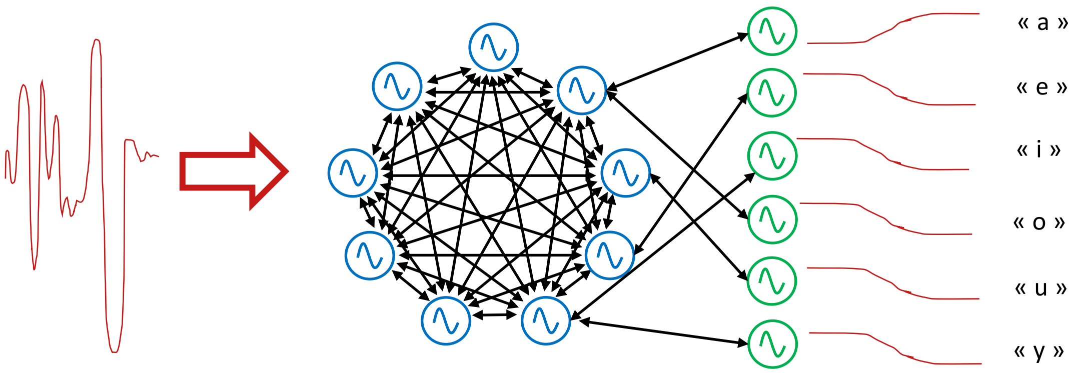
Arduino mobile robot



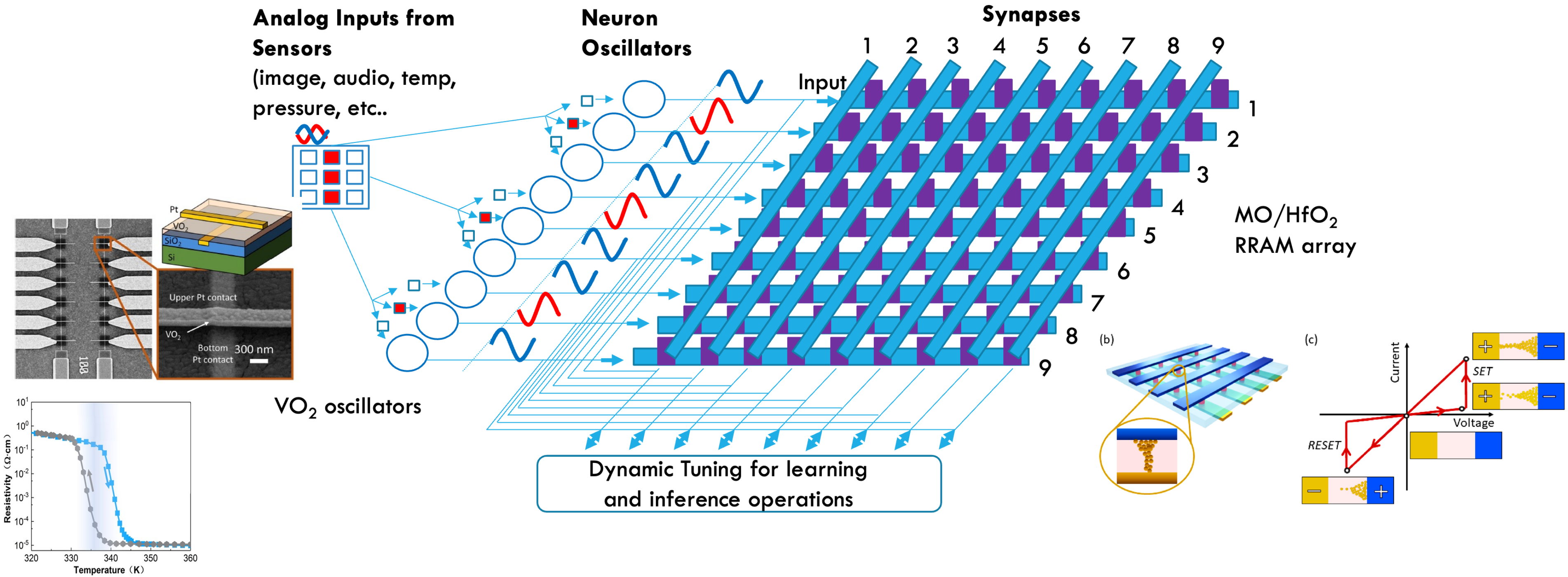
M. Abernot *et al.*, "Mobile Robot Obstacle Avoidance with Oscillatory Neural Networks on FPGA," IBM/IEEE AI Compute Symposium, 2021.

Ongoing and future work:

ONNs are dynamic systems
- How to harness ONN transient dynamics?
- Learning trajectories
- Classification of analog signals (sound, temperature, pressure…)
- Energy efficient and applicable for edge AI computing



Analog ONN for AI Accelerator



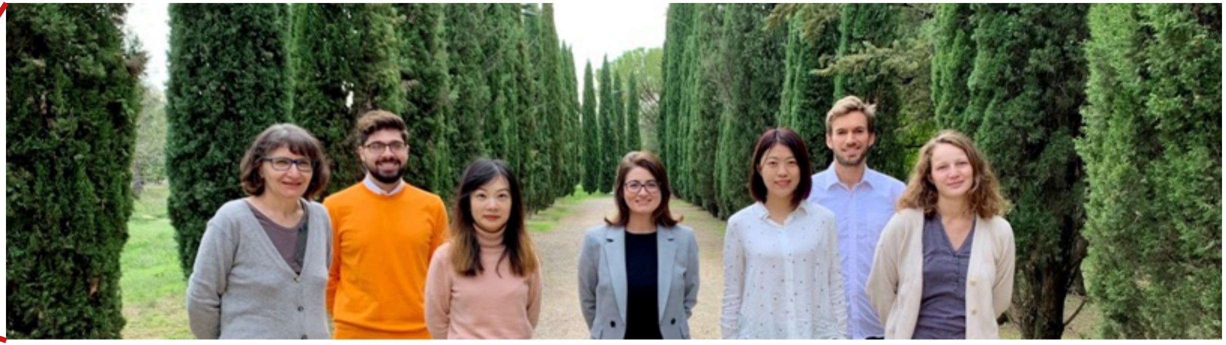
https://www.lirmm.fr/aida-todri-sanial/

PHYDENANO

PHYSICAL DESIGN FOR NANOTECHNOLOGY

- Principal Investigator and Project Coordinator: **Dr. Aida Todri-Sanial**
- Current Research Topics:
  - Beyond CMOS Devices
  - Interconnects
  - Design Methods 3D and Monolithic 3D Integration
  - Neuromorphic Computing – AI at the Edge
  - Quantum Computing

Montpellier, France



Nanocomputing Research Lab at TU/e

Students:

- Thomas Booij
- Edgar Luhulima
- Paul van Geest
- Vincent Cootjans
- Corniels Dekker
- Hamdi Ahmed

Thank you !