Laboratoire IMS Bordeaux

| IMS Bordeaux

Emerging technology and new computing paradigms: How to reinvent the value chain?

Cristell Maneux[1], Marina Deng[1], Chhandak Mukherjee[1], David Atienza[2], Jens Trommer[3], Oskar Baumgartner[4], Guilhem Larrieu[5] and Ian O'Connor[6]

[1]University of Bordeaux,FR ; [2]École Polytechnique Fédérale de Lausanne (EPFL), CH; [3]Namlab gGmbH, DE; [4]Global TCAD Solutions, AT; [5]LAAS – CNRS, FR; [6]Lyon Institute of Nanotechnology, FR

Outline

- Contexts
- Characterization challenges
- Modelling challenges
- Circuit design challenges

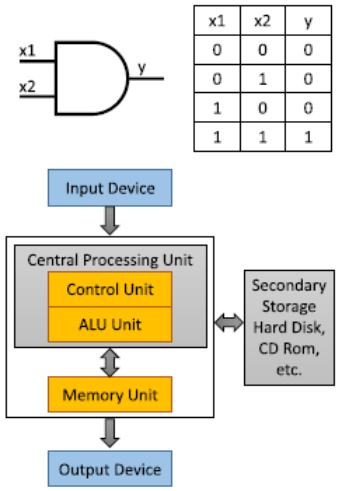
Contexts

- 4th industrial revolution
  - Unprecedented growing demand for neural networks (NNs)
  - **But**, technology solutions still relying on transistors inherited from Moore's Law, optimized for von Neumann machines.
- Energy efficiency of von Neumann based processors
  - limited by data transfer between memory and computing cores implemented in 2D integration schemes.
  - **Crippling limitation** for NNs efficiency.

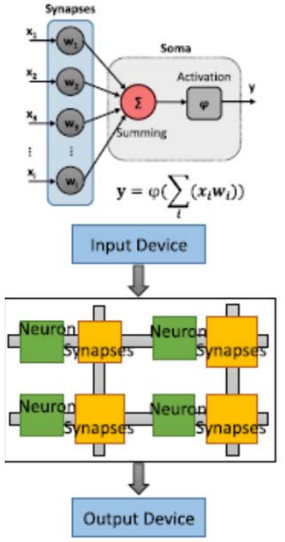
Contexts

- Machine's concept : von Neumann based

Von Neumann architecture

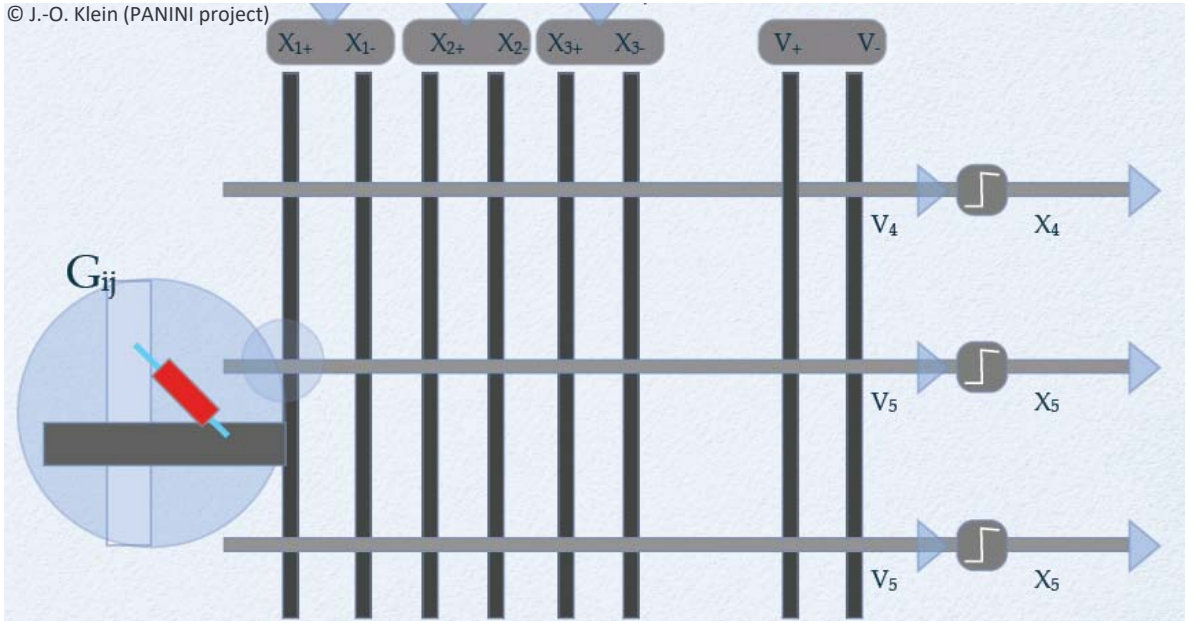


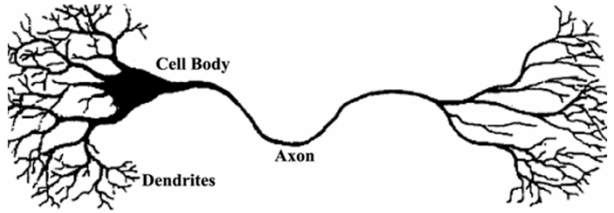
Neuromorphic architecture



Contexts

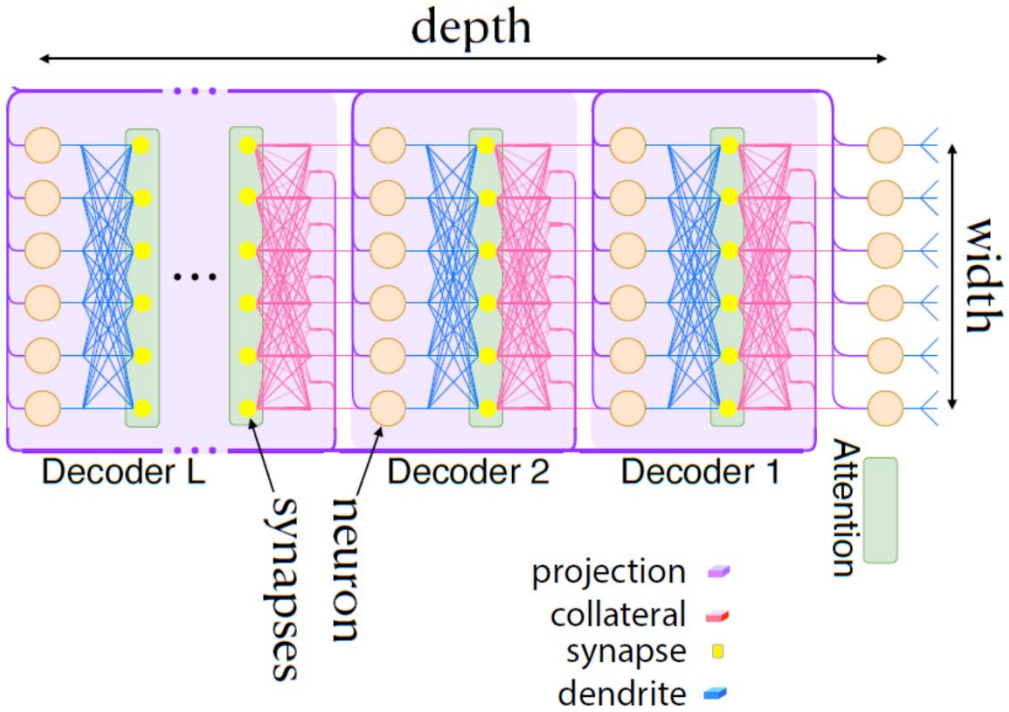
- Machine's concept: Neural Networks





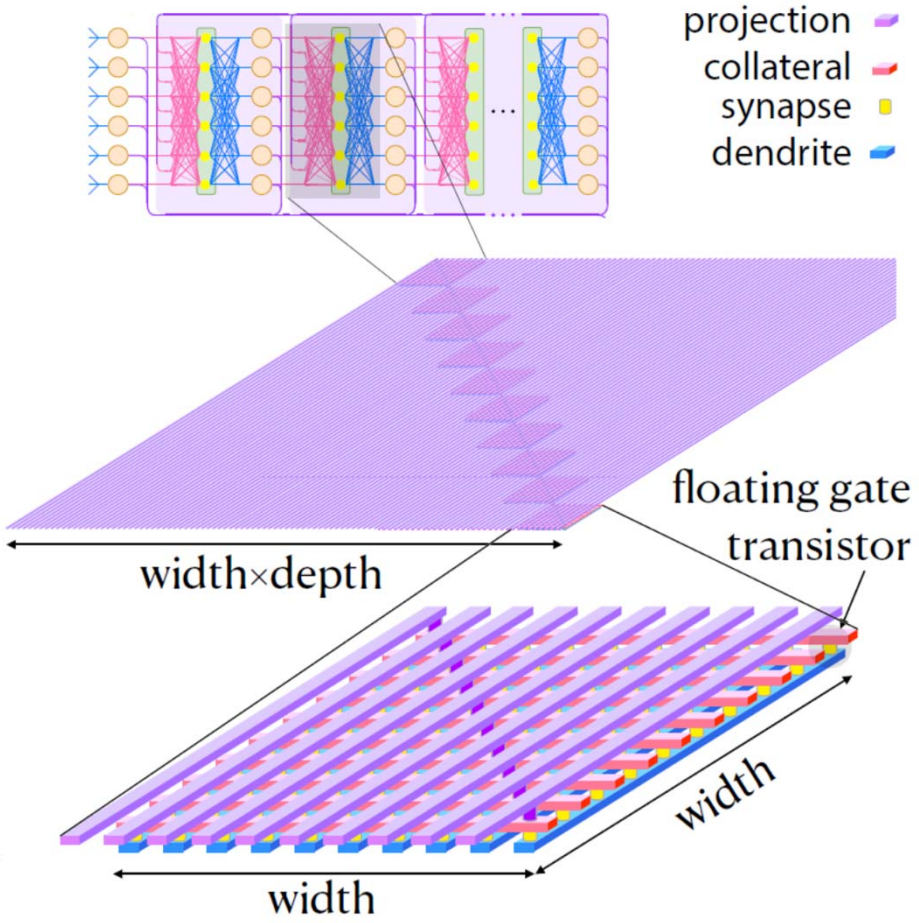
Contexts: Architecture of Neural Networks

- Dense **local connections**
  - feedforward or recurrent
- Sparse **global projections**
  - residual of feedback
- Size = depth × width
  - Depth=Number of stacked layers
  - Width= Number of neurons in a layer



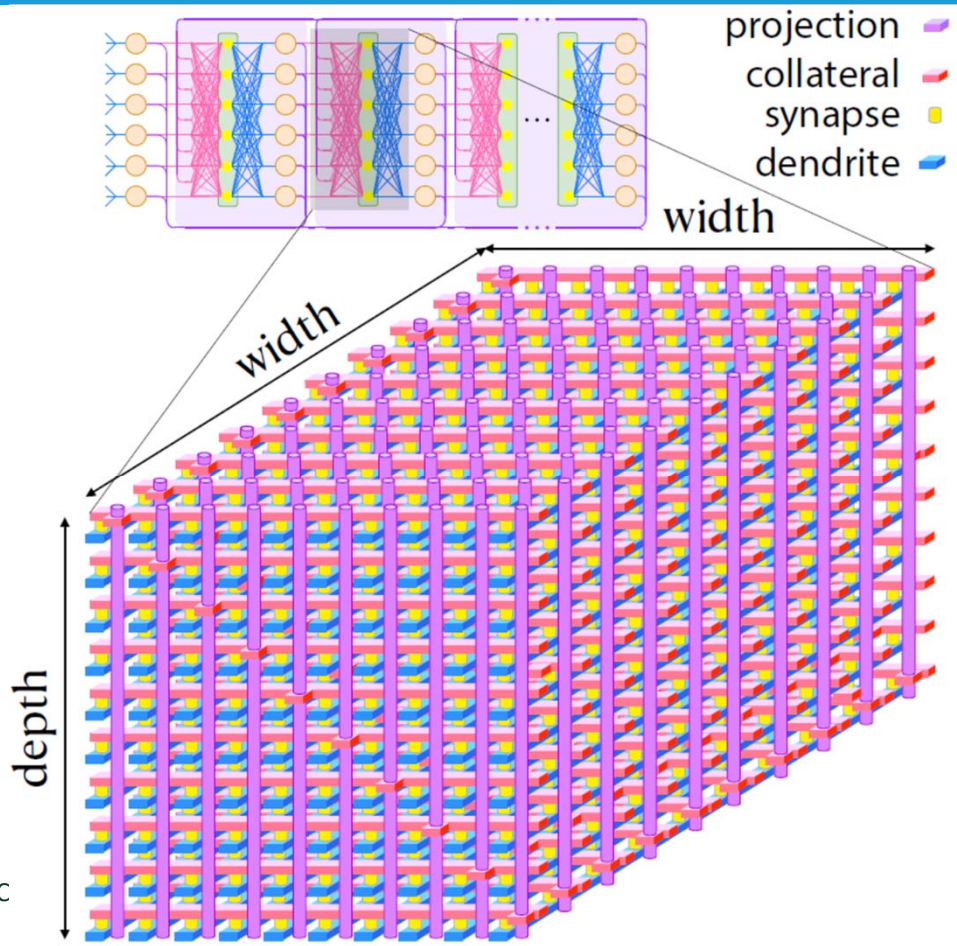
Contexts: 2D Neural Networks

- "Compute-in-memory"
  - work ∝ distance = width × depth
  - signals ∝ size = width × depth
  - energy = work × signals ∝ size^2
  - area ∝ (width × depth)^2 = size^2
  - thermally viable



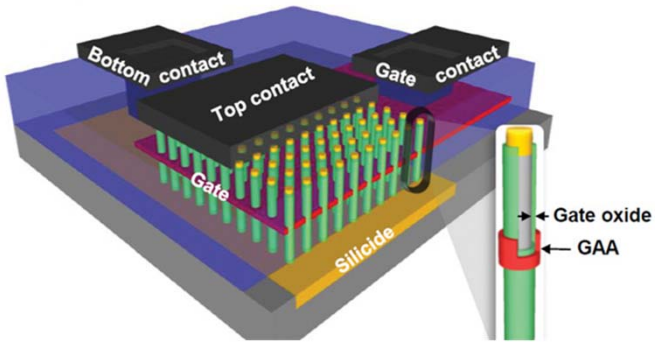
Contexts: 3D Neural Networks

- For width>>depth
  - work ∝ distance ≈ width
  - signals ∝ size = width × depth
  - energy = work × signals
    = $width^2 \times depth$
- And depth ∝ width
  - energy = $size^{1.5}$
  - area ≈ $width^2$ ∝ size
  - Thermally viable?

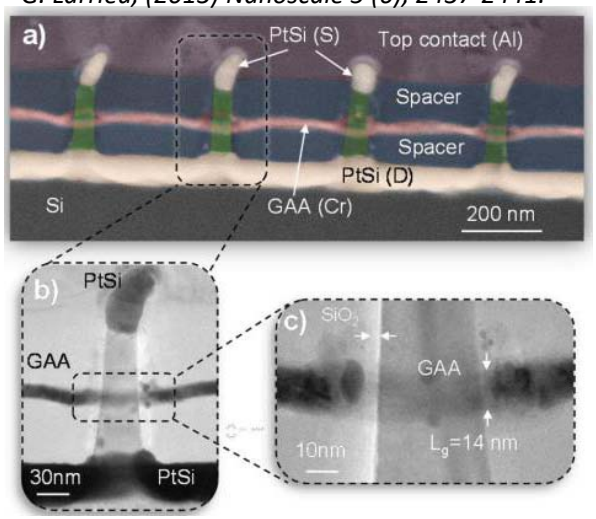


Context: Conclusion

- To save energy, we need :
  - A 3D implementation …
  - Dedicated devices for 3D stacked NNs
- How could we do this?



G. Larrieu, (2013) Nanoscale 5 (6), 2437-2441.



Key features of the technology

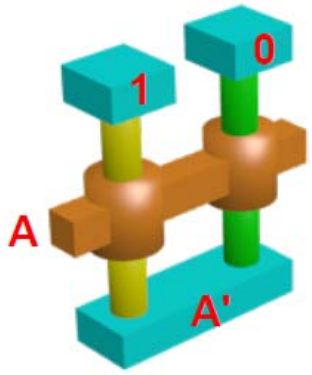
- Vertical Gate-all-around device
- Scaled gate length ~14 nm
- Symmetrical S/D silicided contacts that ensure junctionless operation

Outline

- Contexts
- Characterization challenges
- Modelling challenges
- Circuit design challenges

Vertical Nanowire FET: Characterization challenges

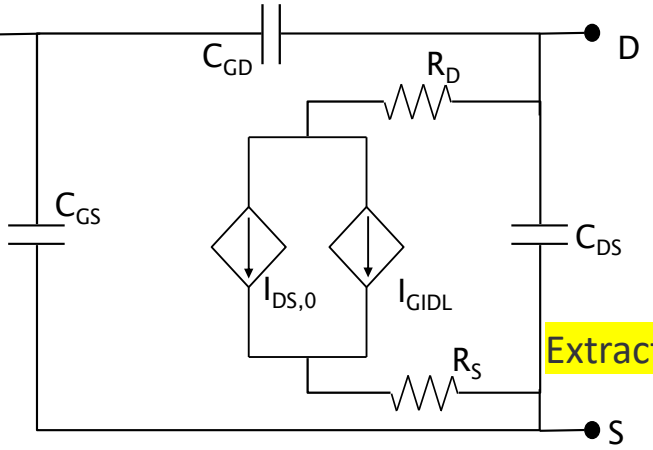
- Measurement accuracy



INV1-SL-NP

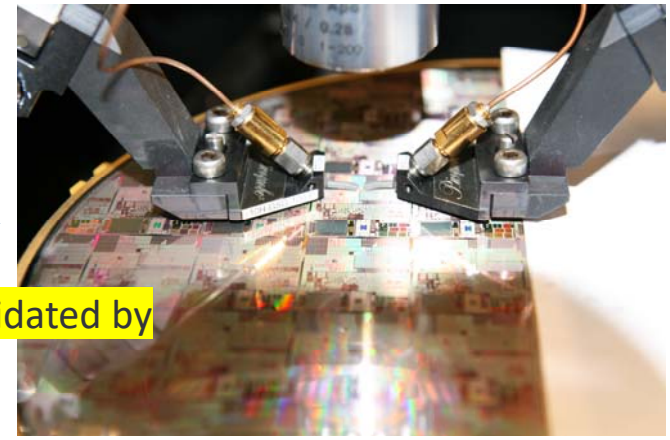
Circuit/system design

Relies on



Accurate compact models

Extracted from/validated by



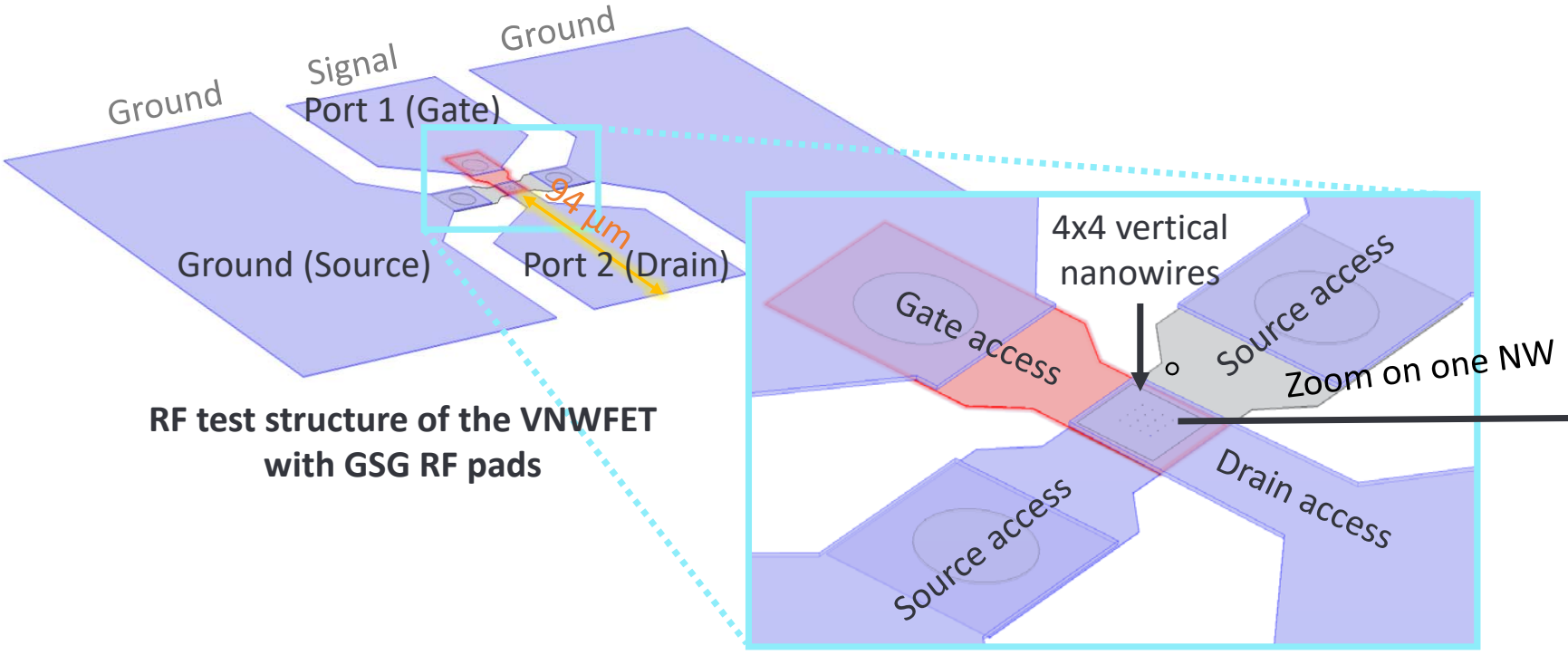
On-wafer device measurements

DC + S-parameters

Accurate S-parameter measurements for nanoscaled devices=big challenge !!!

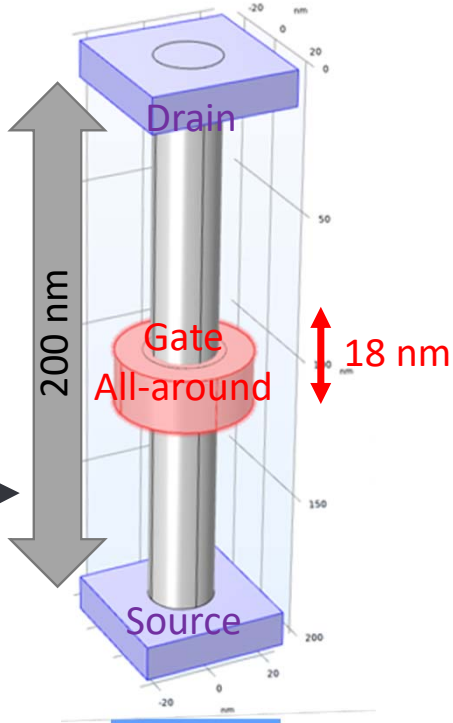
Vertical Nanowire FET: Characterization challenges

- From test structure to intrinsic device



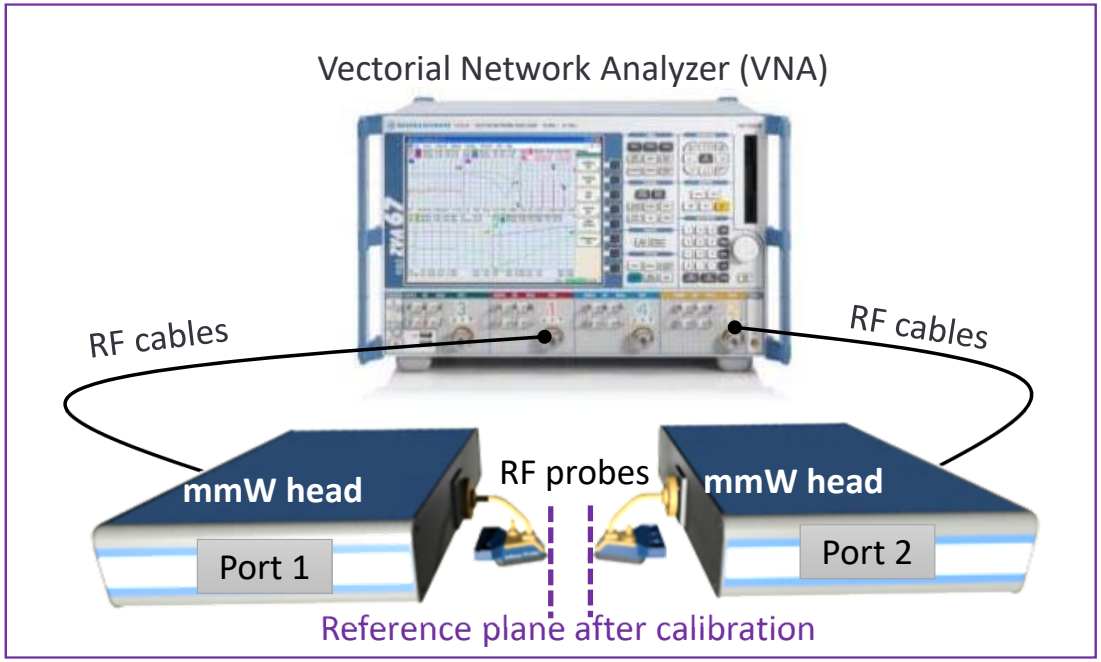
RF test structure of the VNWFET with GSG RF pads

Single nanowire constituting the VNWFET



Vertical Nanowire FET: Characterization challenges

- Calibration & De-embedding Issues
  - Standard procedure=2-step calibration

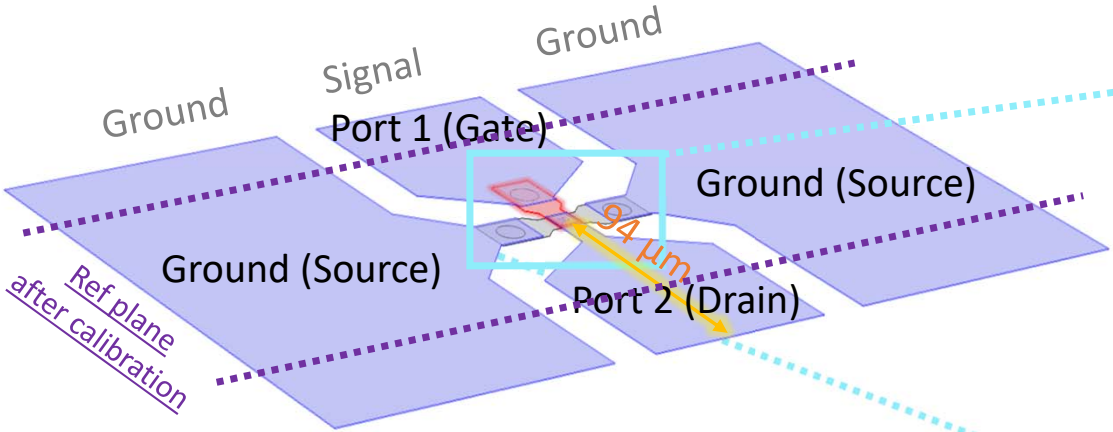


1 → **Off-wafer calibration on ISS**

Impedance Standard Substrate
Source=Cascade Microtech

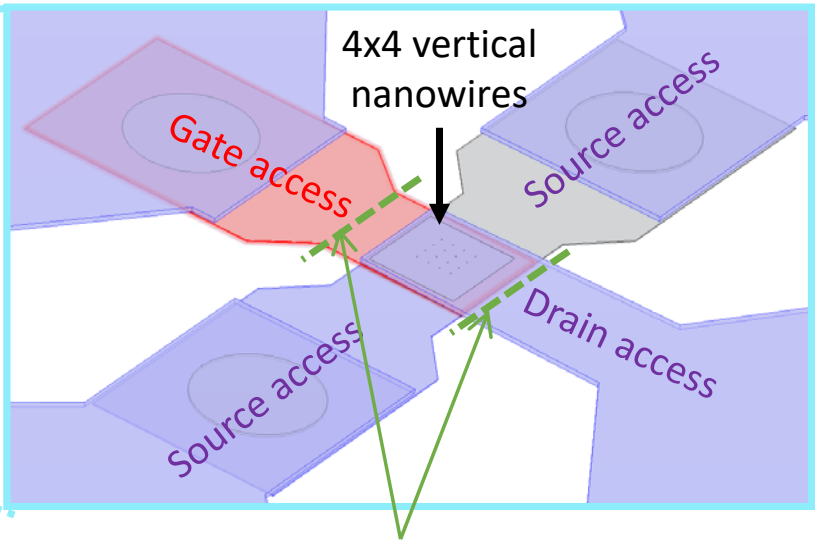
Vertical Nanowire FET: Characterization challenges

- Calibration & De-embedding Issues
  - Standard procedure=2-step calibration



RF test structure of the VNWFET with GSG RF pads

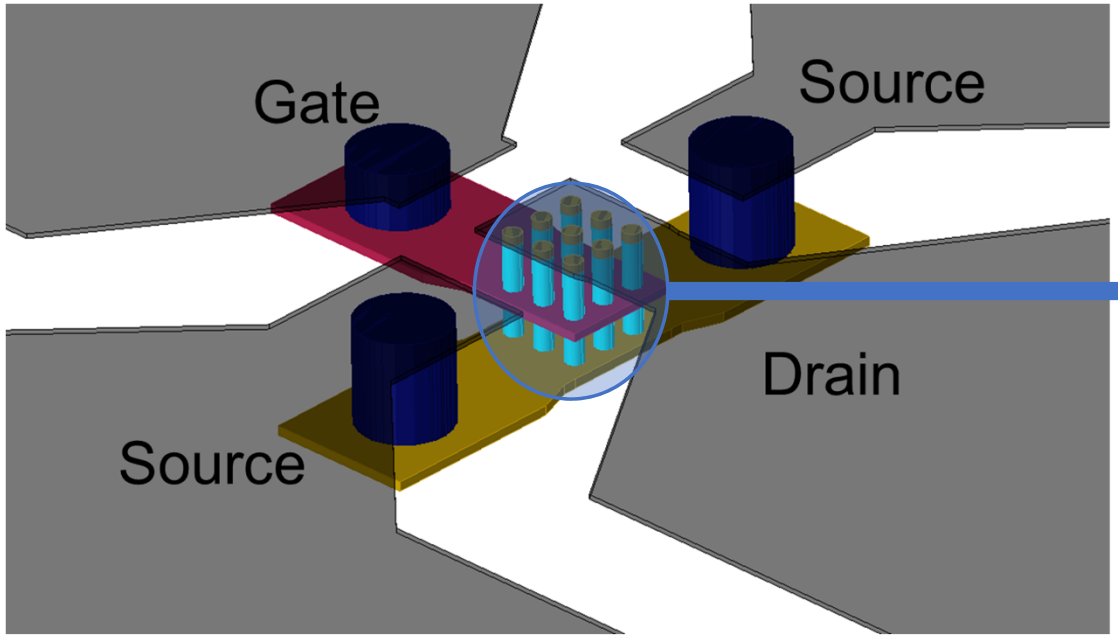
2 → **On-wafer de-embedding**



Reference plane after de-embedding (DUT terminals)

VNWFET = 500x smaller than the accesses
➔ **De-embedding** step is **critical**

Which de-embedding method?



Inside: the intrinsic device,
P type 81 **vertical** nanowires with 17 nm of diameter each

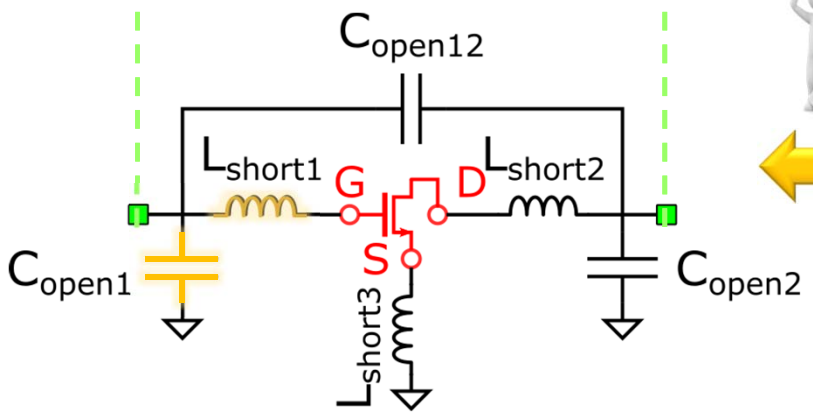
Outside: **3D Interconnects** and pads necessary for electrical contact (DC and RF pads)

Which de-embedding method?

- De-embedding removes errors due to device accesses
  - Based on an equivalent circuit of the transistor accesses

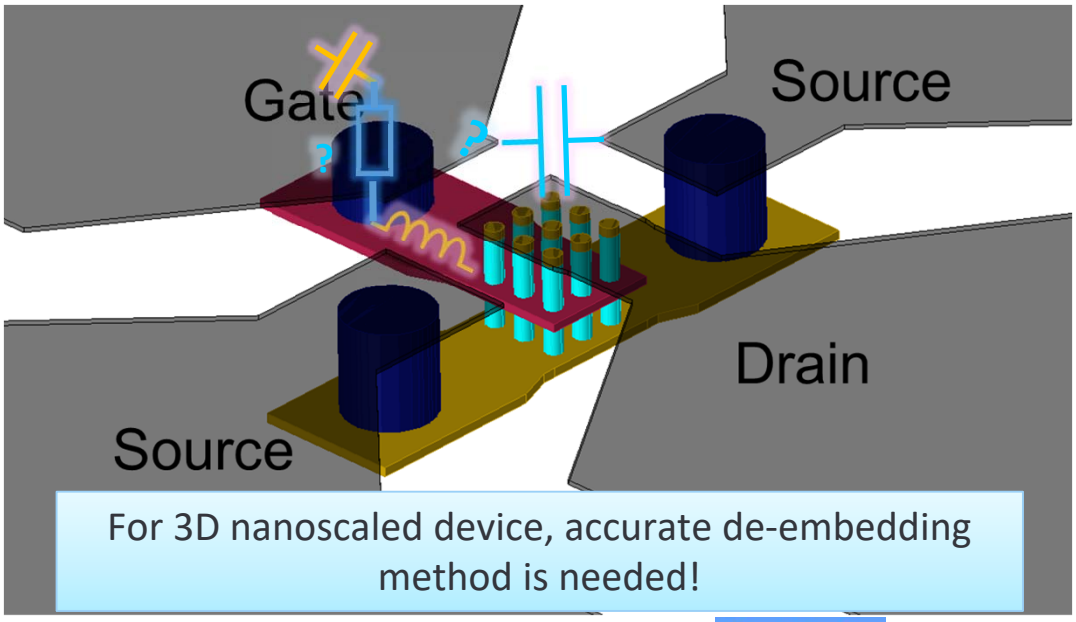
Small Signal Equivalent Circuit: SSEC

Ref. plane after calibration



OPEN-SHORT DE-EMBEDDING

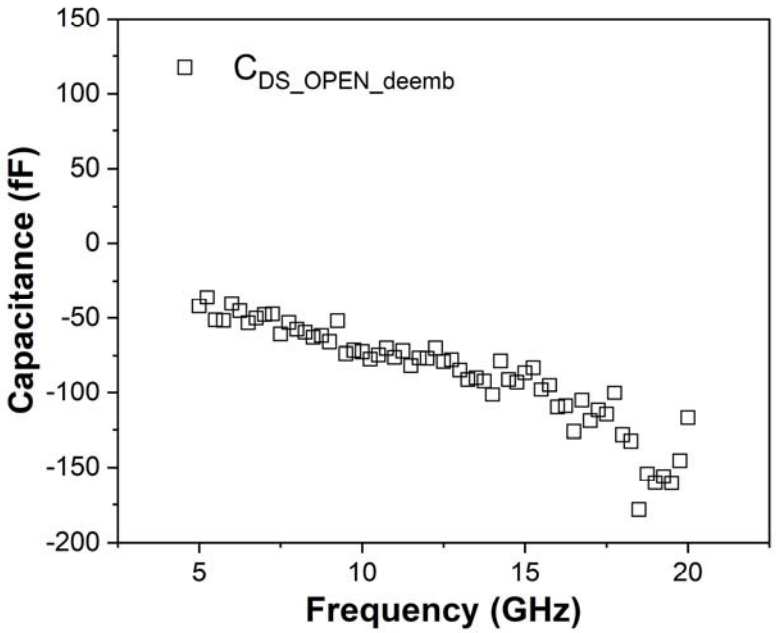
(conventionnal method)



For 3D nanoscaled device, accurate de-embedding method is needed!

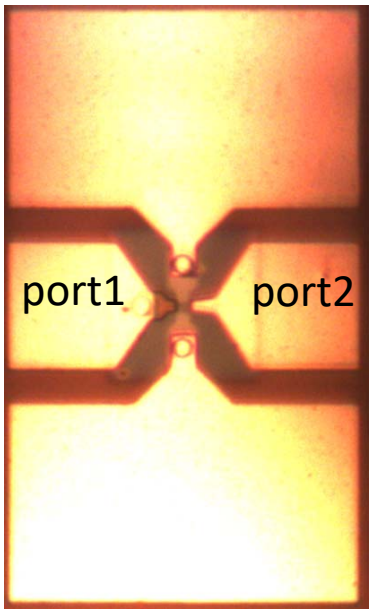
Which de-embedding method?

Open de-embedding



- Unphysical capacitance value
- New de-embedding method necessary

De-embedding method for VNWFETs



OPEN

Measure-ment

- S-parameter measurement up to 40 GHz
- DUT = Open test structure and transistor

EM simulation

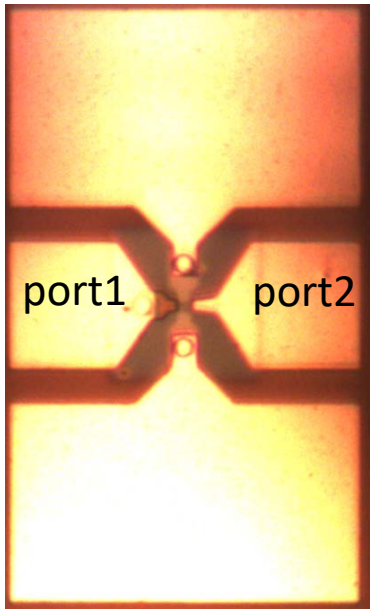
- Simulator = ADS-Momentum
- Definition of the electrical properties for each layer used in the VNWFET process

Electrical simulation

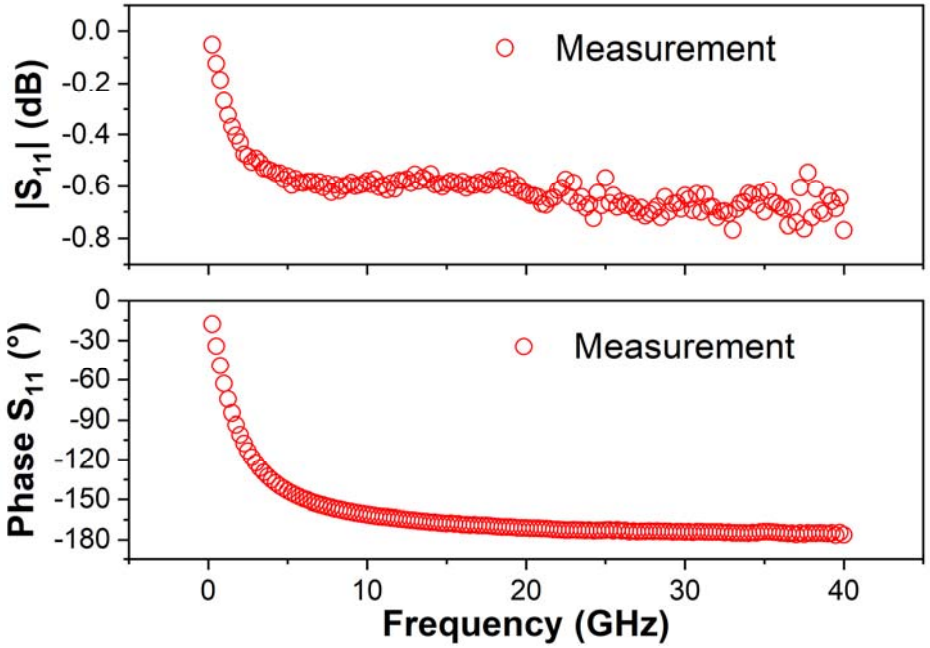
- SSEC of the parasitic network
- De-embedding of VNWFET

S-parameters measurements

- S-parameters on-wafer measurements up to 40 GHz of the OPEN structure

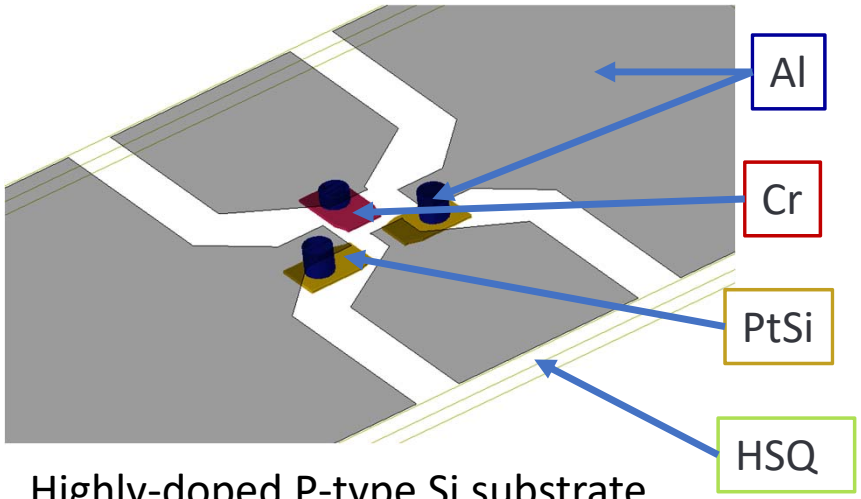


OPEN



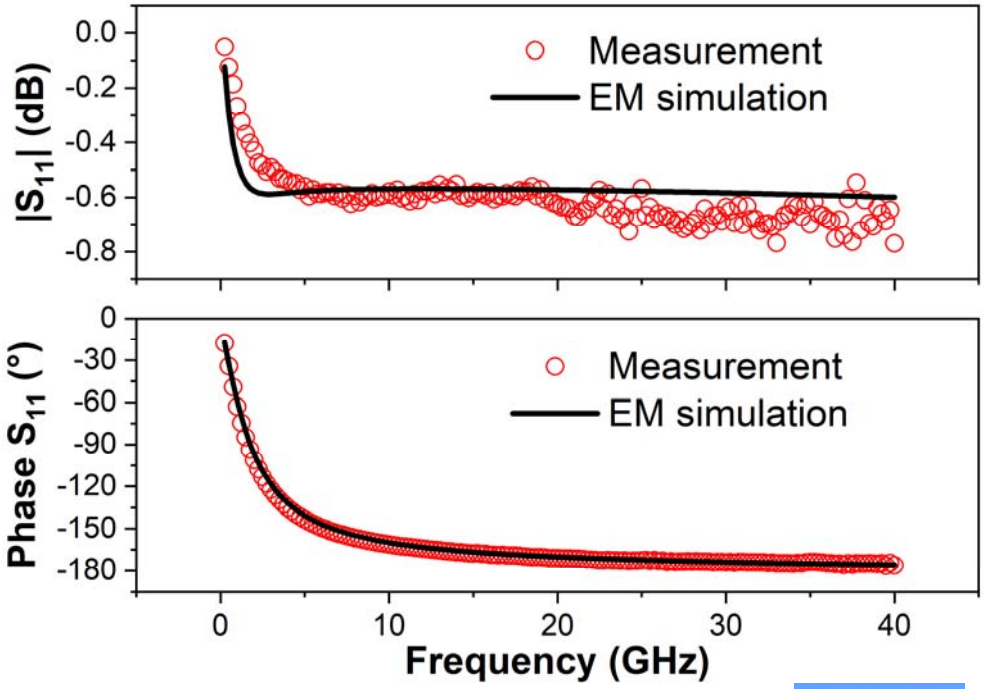
EM simulation

- S-parameters from EM simulation up to 40 GHz of the OPEN structure

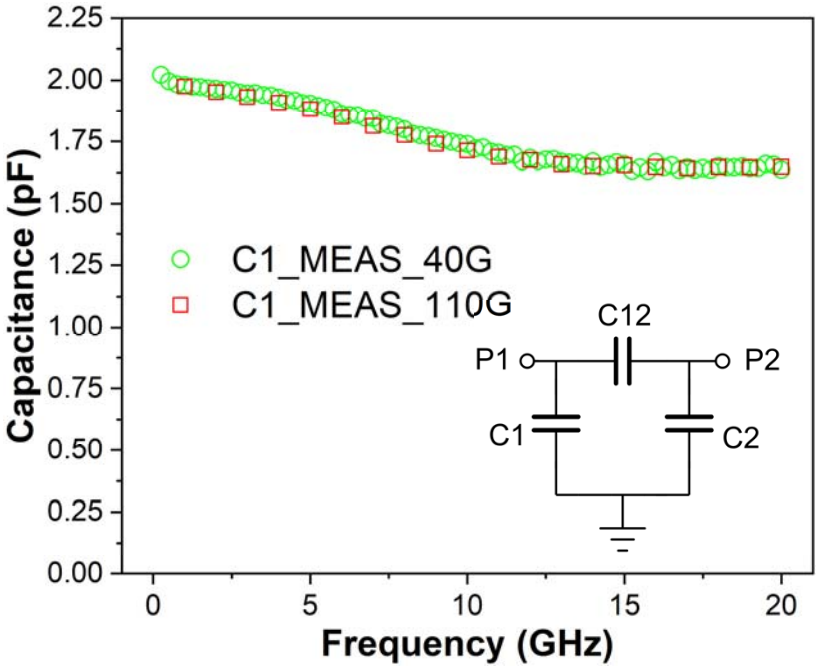


Highly-doped P-type Si substrate

OPEN ADS 3D VIEW



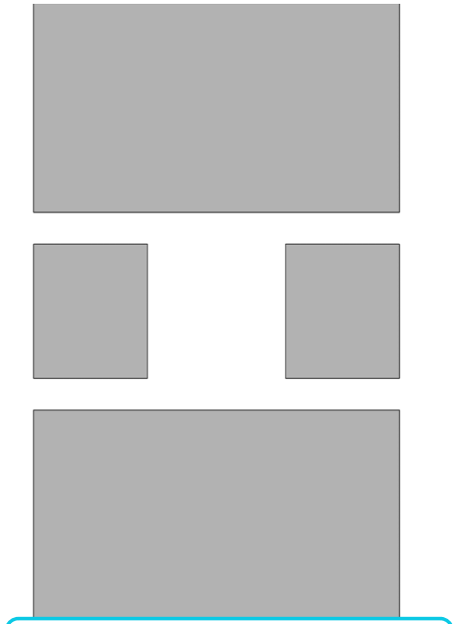
Open Capacitance extraction



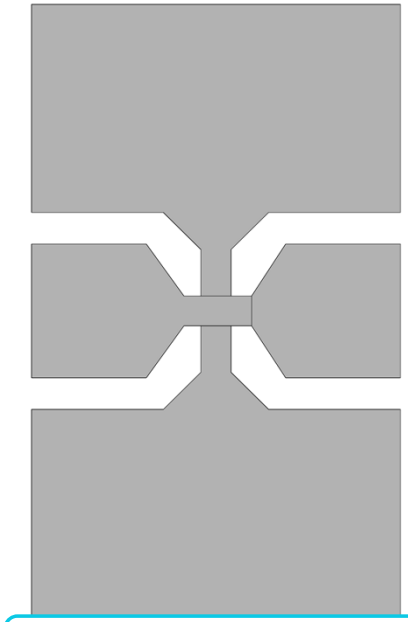
- Open structure as a π-network? Investigate a more complex circuit!

SSEC of the parasitic network

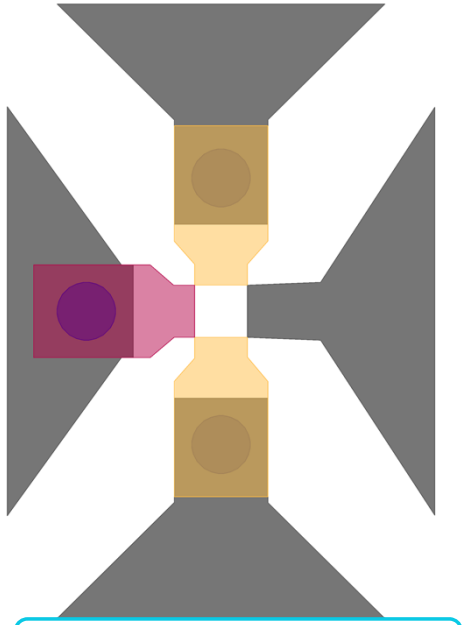
- EM simulation for virtual structures



OPEN PAD



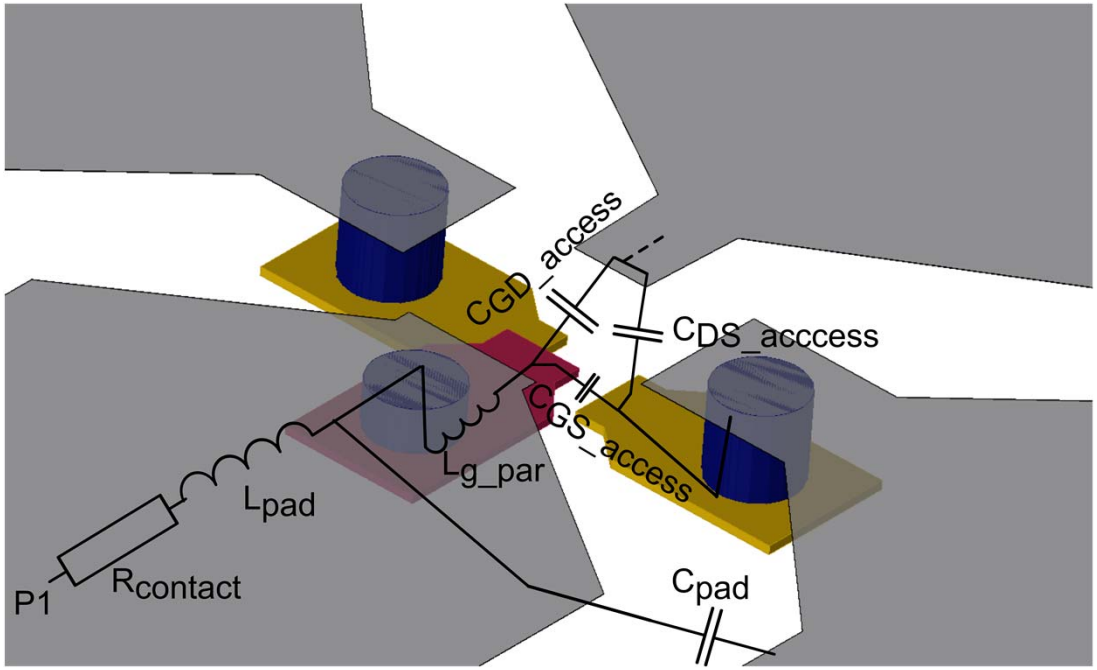
SHORT PAD



OPEN DEVICE

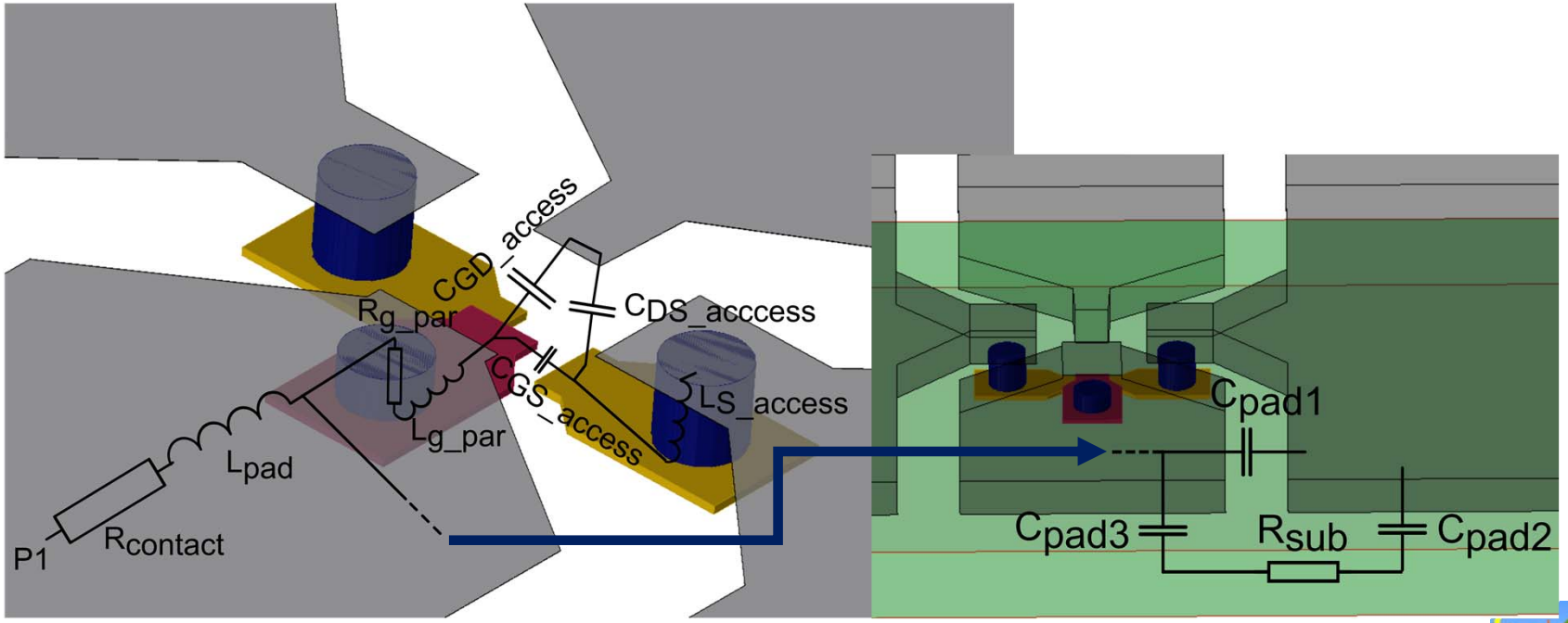
SSEC of the parasitic network

- Distribution of the lumped elements: MODEL1



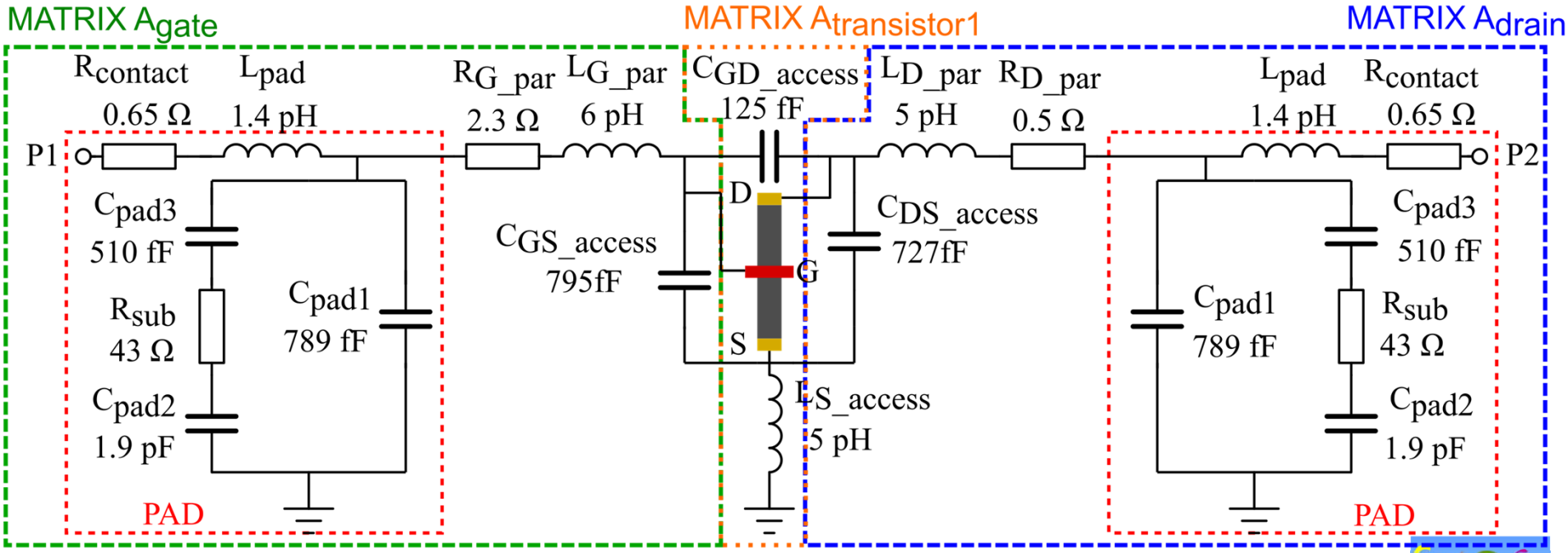
SSEC of the parasitic network

- Distribution of the lumped elements: MODEL2



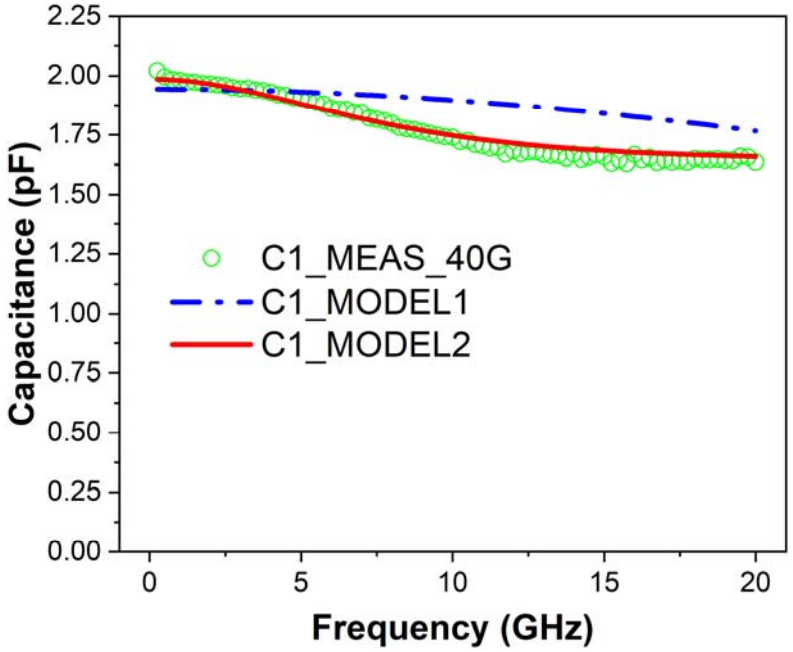
SSEC of the parasitic network

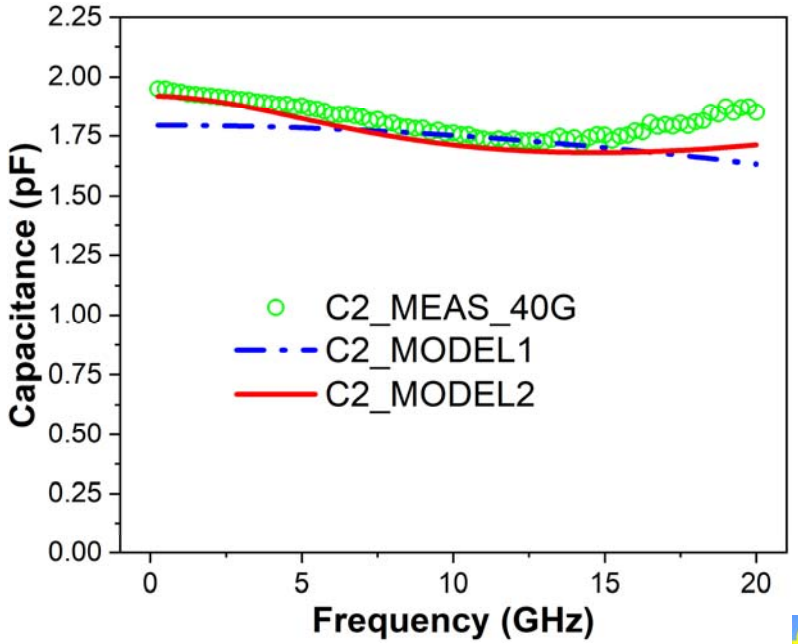
- Final small signal equivalent circuit of the parasitic network: MODEL2



SSEC of parasitic network

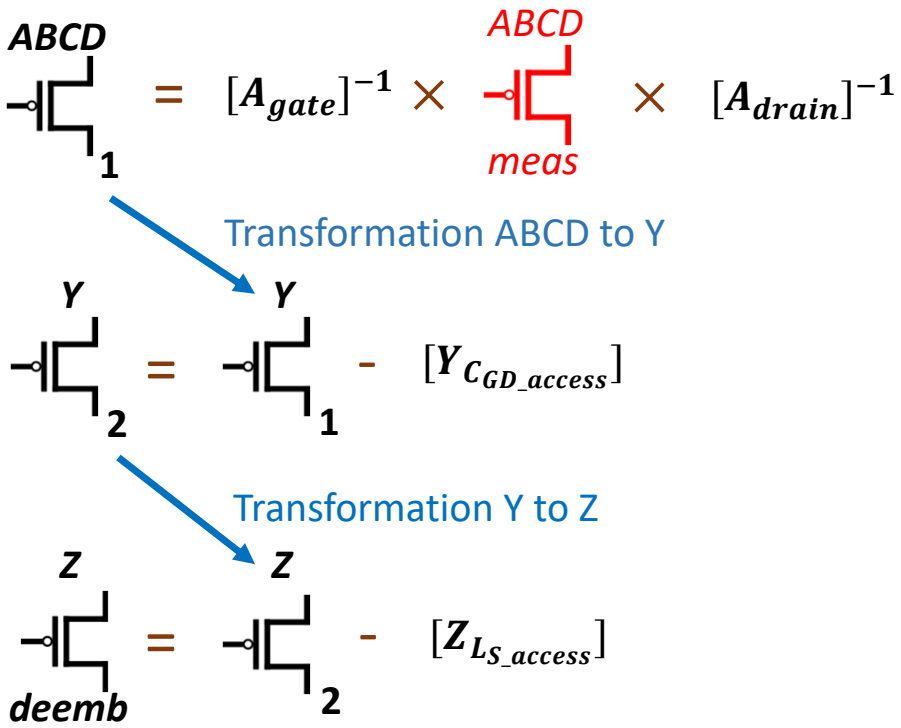
- Capacitance from the small signal equivalent circuit of the open structure up to 20 GHz





Transistor de-embedding: access model

- Using a combination of ABCD matrices and Y and Z parameters

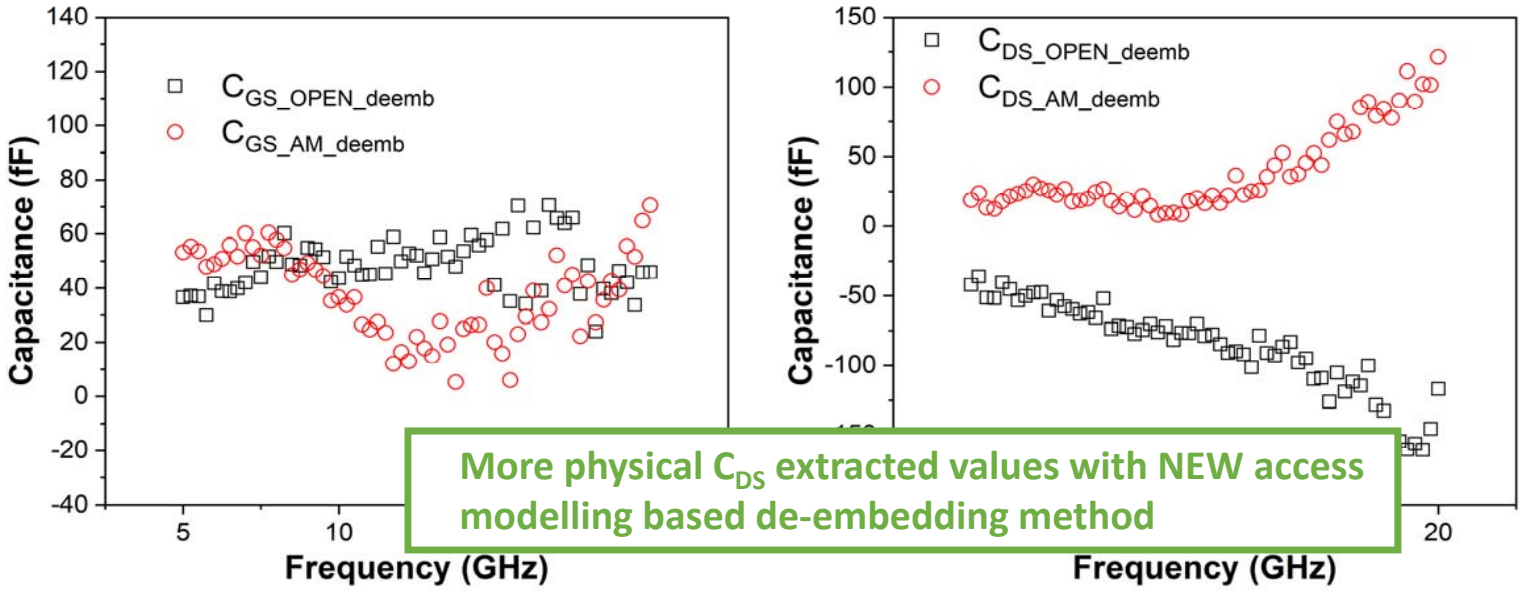


Measurement data

Simulation data

De-embedding results

- Comparison between open and access modelling de-embedding for C_{GS} and C_{DS} extraction

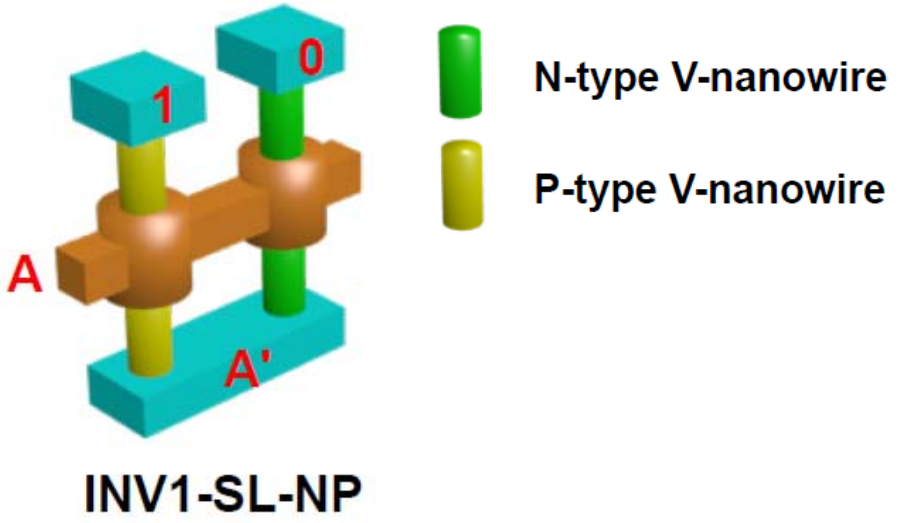


More physical C_{DS} extracted values with NEW access modelling based de-embedding method

VNWFET Characterization

- Reliable measurements for VNWFET model parameter extraction is OK!

➔ Ready for compact modelling!

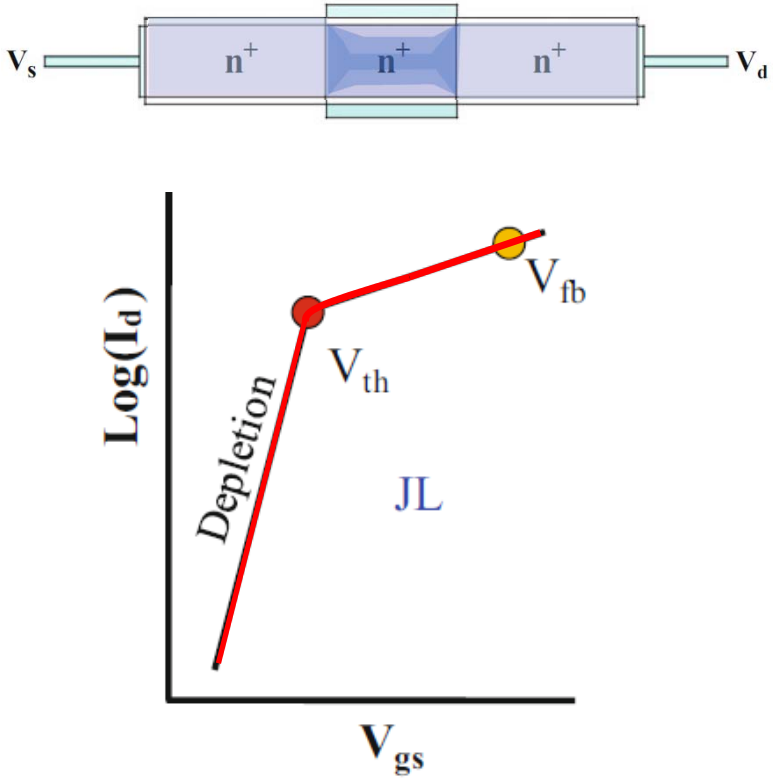


INV1-SL-NP

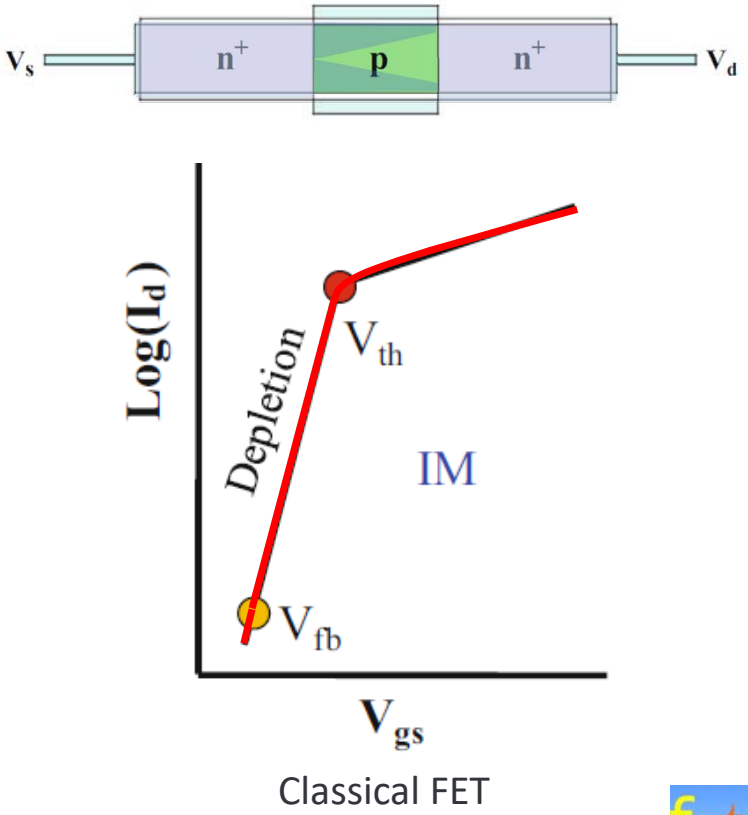
Outline

- Contexts
- Characterization challenges
- Modelling challenges
- Circuit design challenges

Junctionless vs. Classical MOSFETs

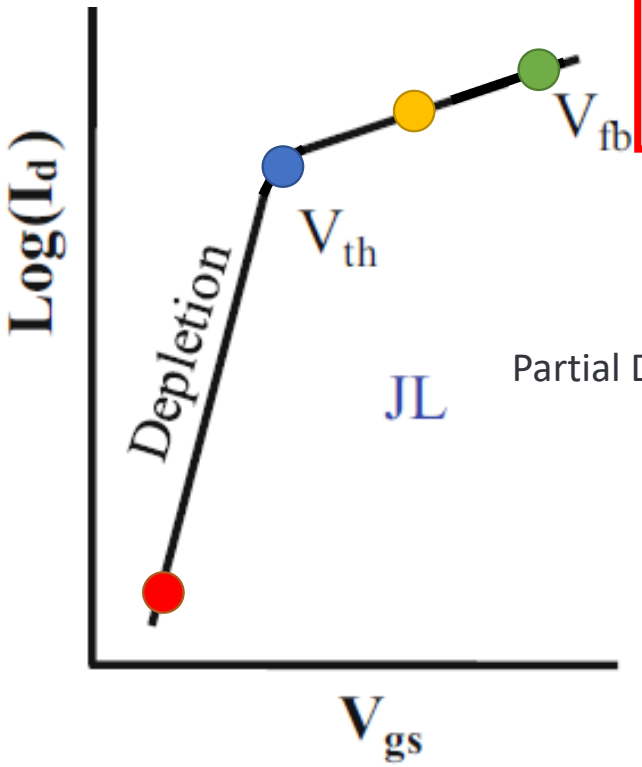


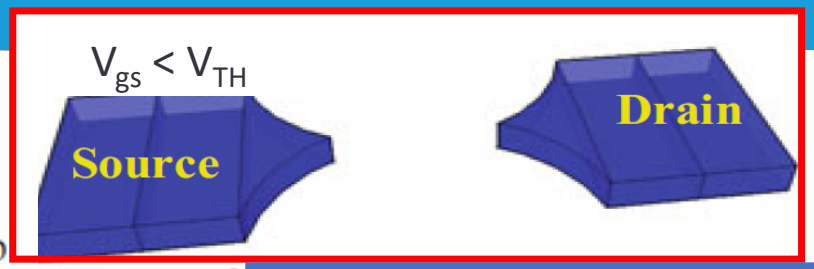
Junctionless FET



Classical FET

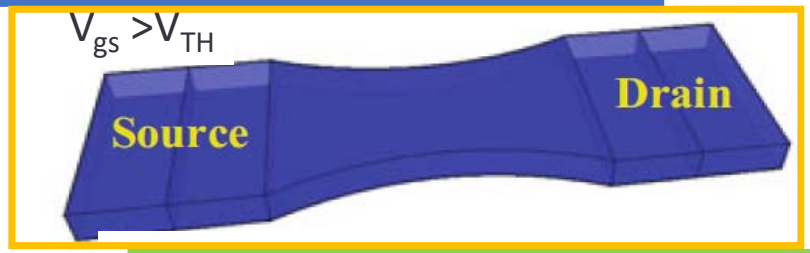
Junctionless FETs: Operation



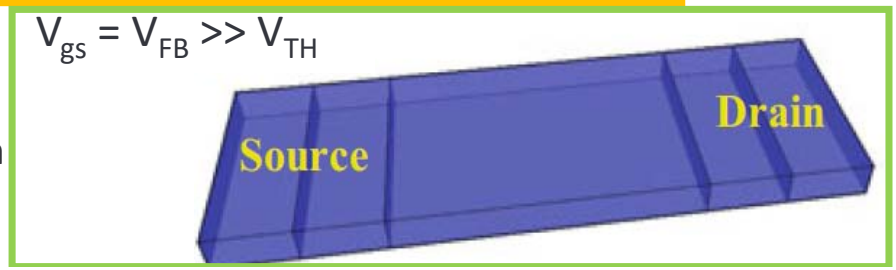


Full Depletion

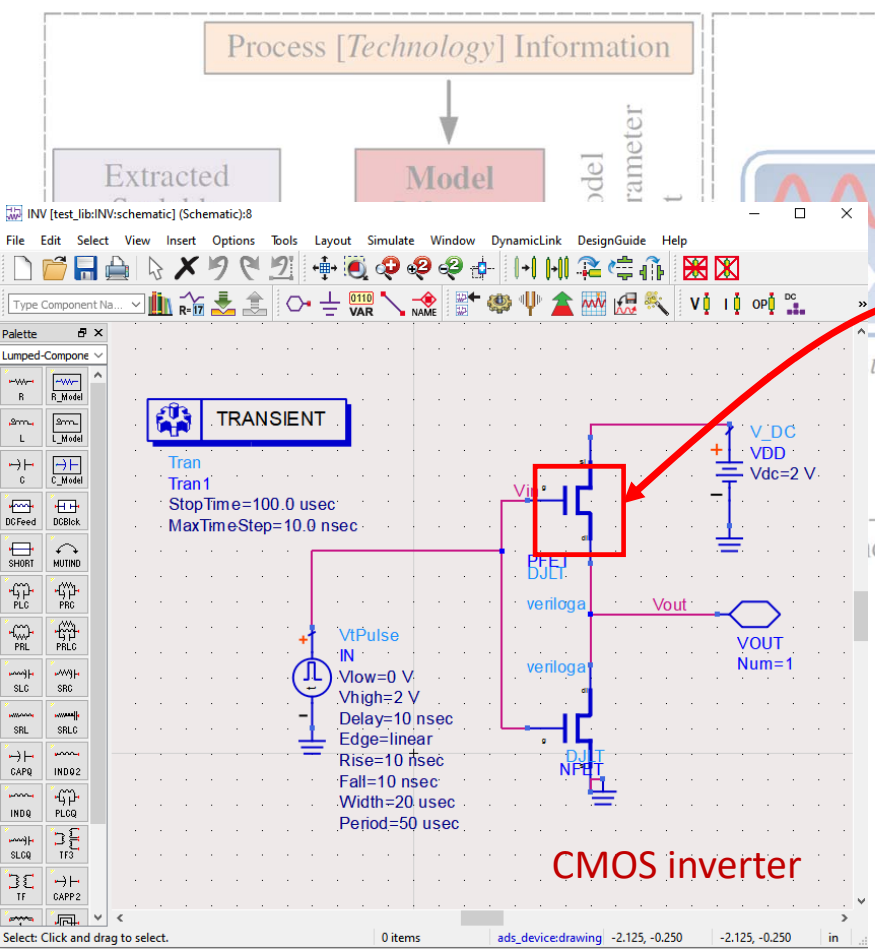
Partial Depletion

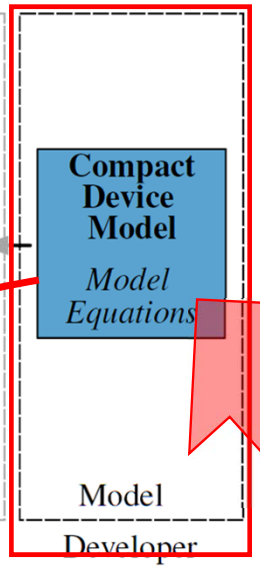


Accumulation

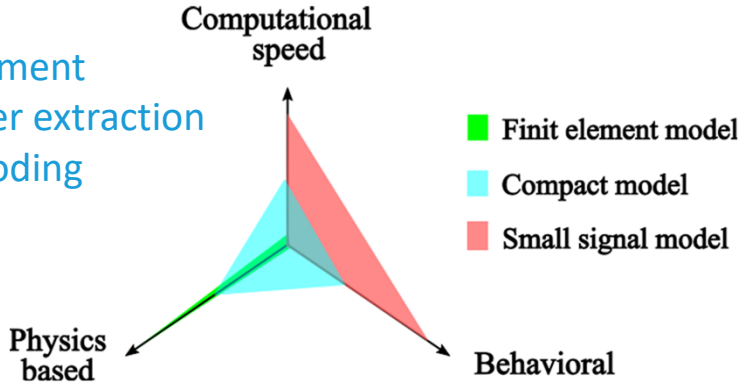


Compact model: A design tool





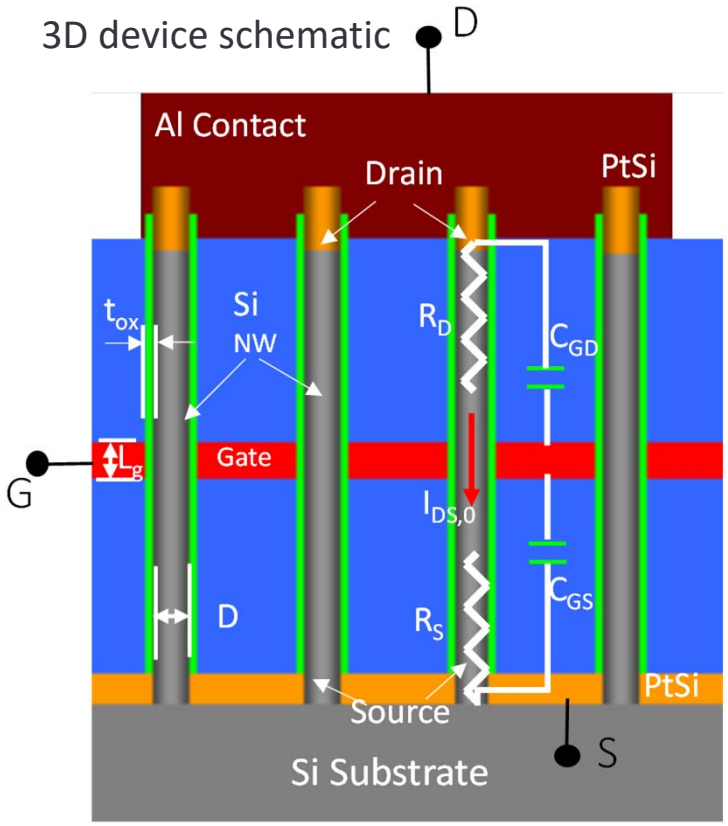
- Measurement
- Parameter extraction
- Model Coding

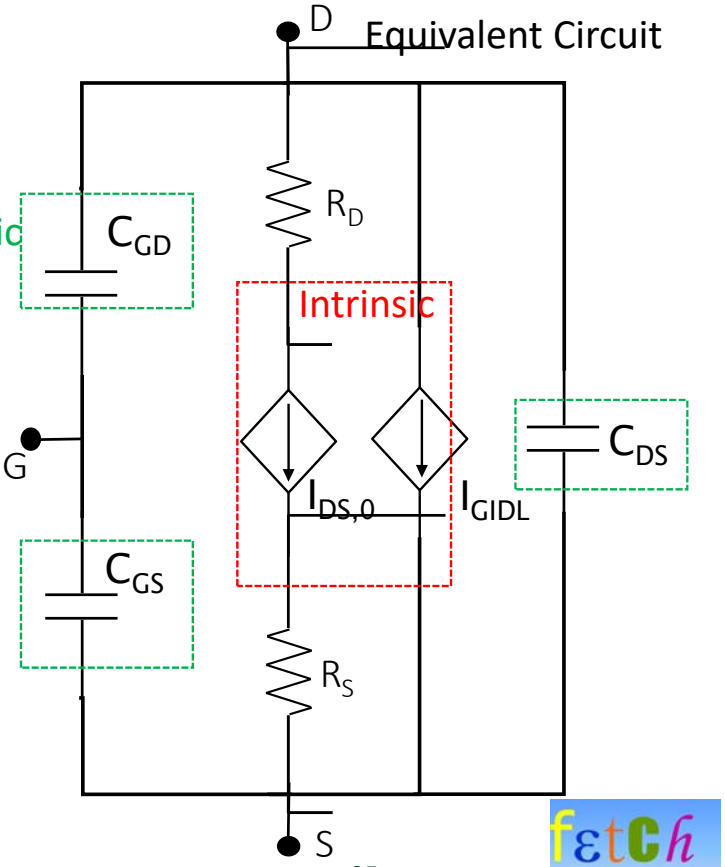


Verilog-A Code

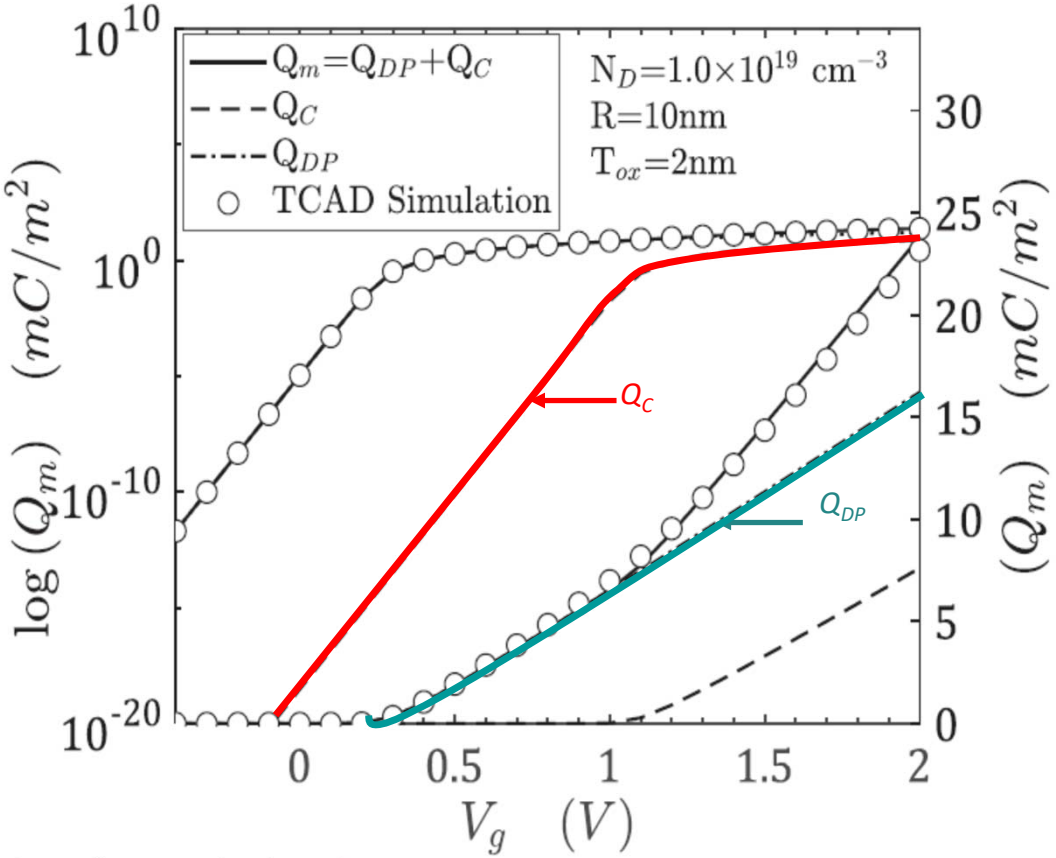
```
//current at source (suffix 0) and drain (suffix L)
IDP0=pow(QDP0,2)/(2*eta2*Cox*uT)+QDP0;
IDPL=pow(QDPL,2)/(2*eta2*Cox*uT)+QDPL;
IC0=pow(QC0,2)/(2*eta2*Cc*uT)+2*QC0;
ICL=pow(QCL,2)/(2*eta2*Cc*uT)+2*QCL;

I0=pow(QDP0,2)/(2*eta2*Cox*uT)+QDP0+pow(QC0,2)/(2*eta2*Cc*uT)+2*QC0;
IL=pow(QDPL,2)/(2*eta2*Cox*uT)+QDPL+pow(QCL,2)/(2*eta2*Cc*uT)+2*QCL;


//short channel effect corrections
VS2=(QDP0+QC0)/Cox+Vmax/(Vmax/Vmin-1);
VDSAT=VS2/(VS2/Vmax+1);
VDEFF=VDSAT-VDSAT*(ln(1+exp(A2*(1-vds/VDSAT)))/ln(1+exp(A2)));
```

Junctionless FET compact model





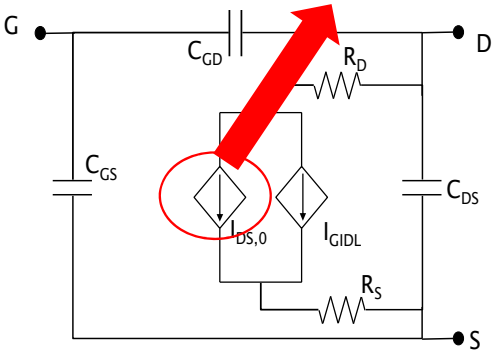
Junctionless FET compact model: I_{DS}

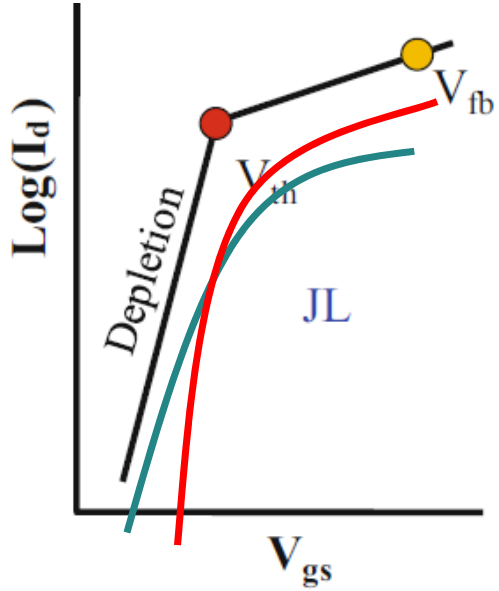


$$I_{DS,0} = \mu_{eff}\frac{2\pi R}{L_{eff}}\left[\int_{Q_{DP0}}^{Q_{DPL}} Q_{DP}\frac{dV}{dQ_{DP}}dQ_{DP} + \int_{Q_{C0}}^{Q_{CL}} Q_C\frac{dV}{dQ_C}dQ_C\right]$$

Depletion charge

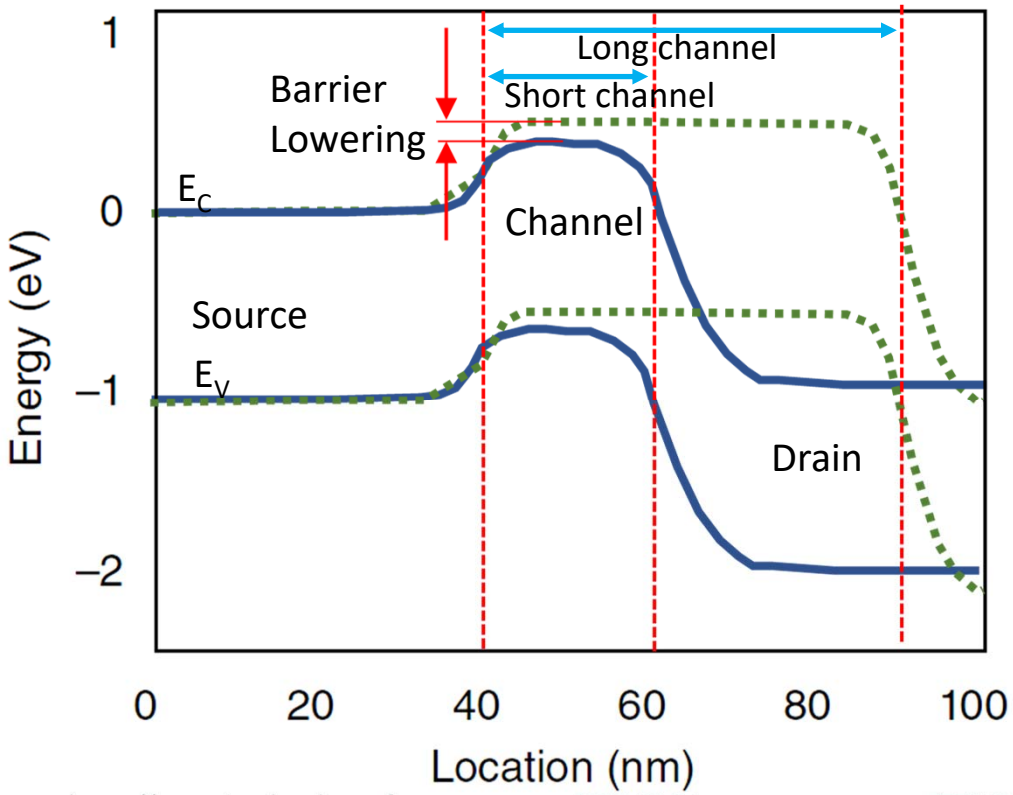
Accumulation charge

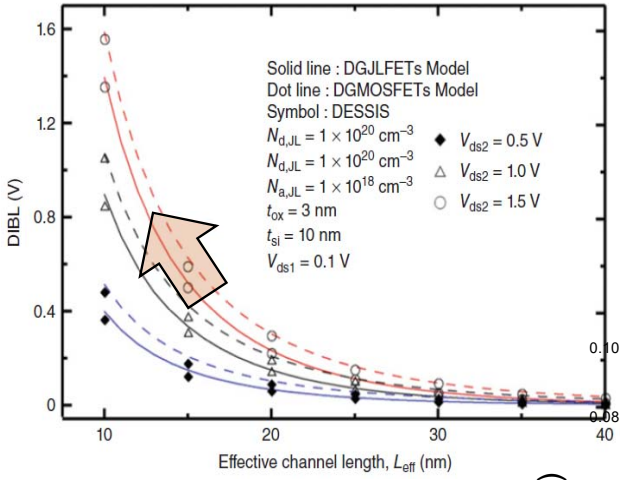


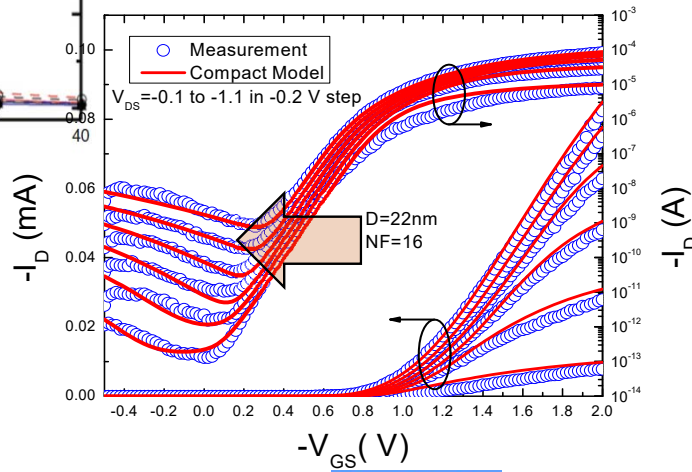


Short Channel Effects: DIBL

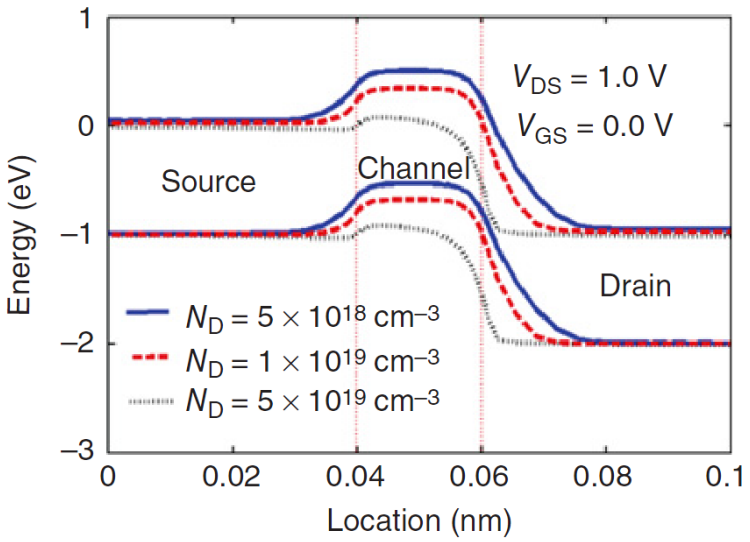
Drain-induced barrier lowering

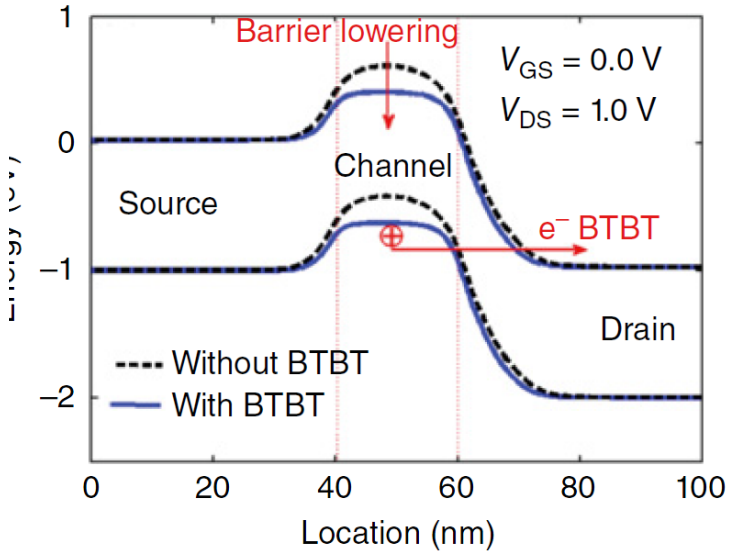


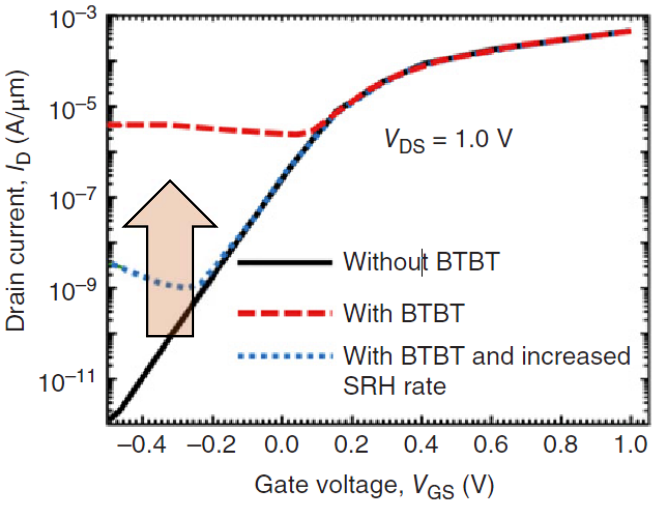




Short Channel Effects: BTBT/GIDL



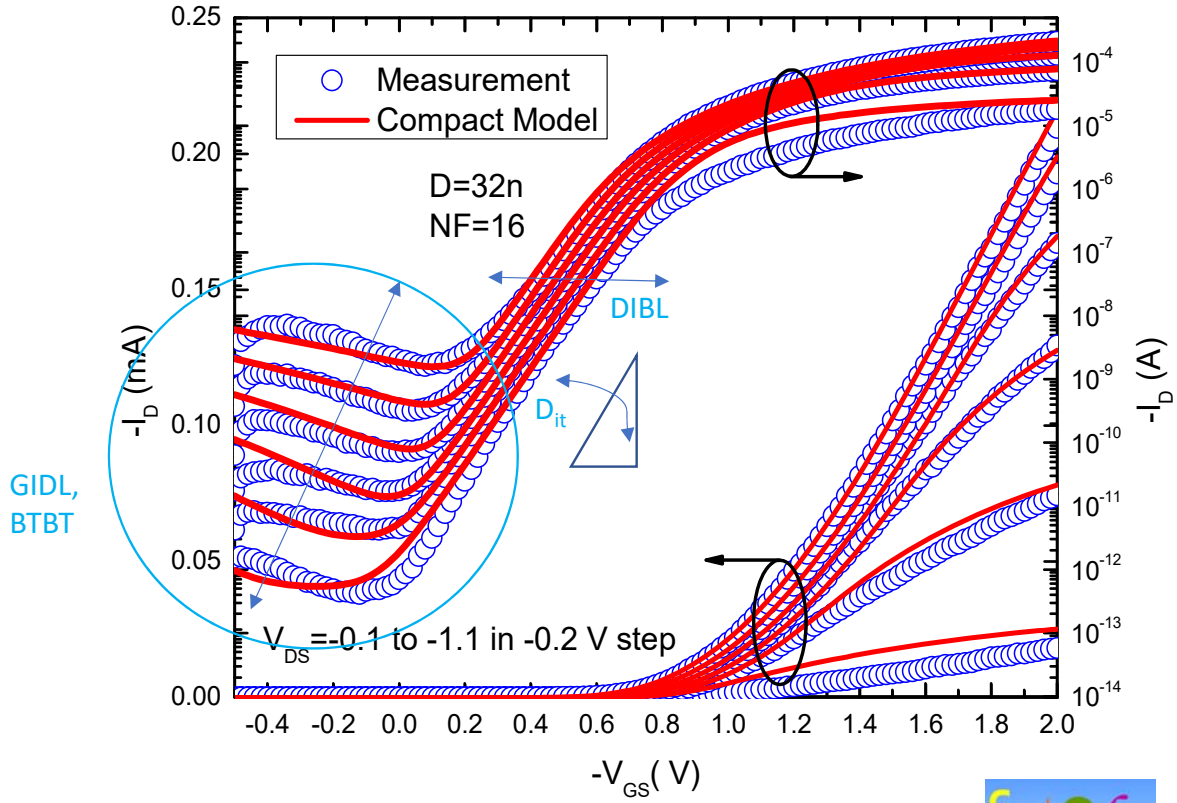




Gate-induced drain leakage through band-to-band tunneling (BTBT)

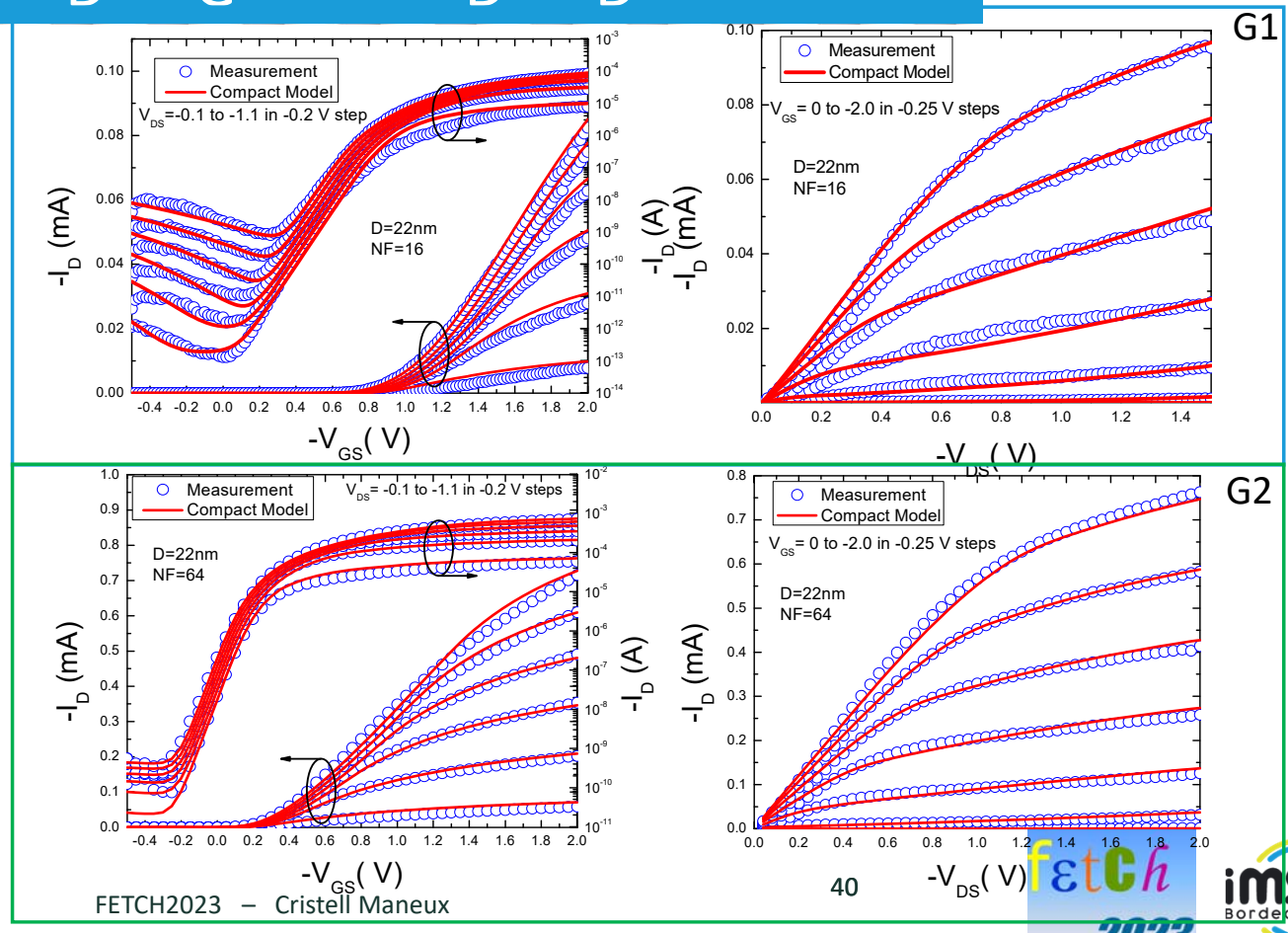
Modeling the DC Characteristics

- Geometrical parameters
  - Diameter (D)
  - Number of nanowires in parallel (NF)
  - Oxide thickness
  - Gate work-function
  - NW doping
- Modeling the Current
  - Mobility
  - Carrier saturation velocity
  - D/S access resistances
- Modeling the Subthreshold:
  - Interface Traps
  - Drain-induced barrier lowering
  - Band-to-band tunneling



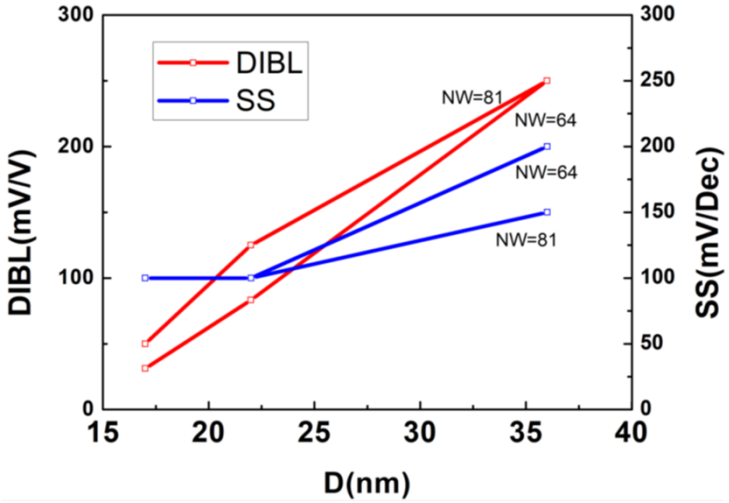
DC Characteristics: I_D-V_G and I_D-V_D

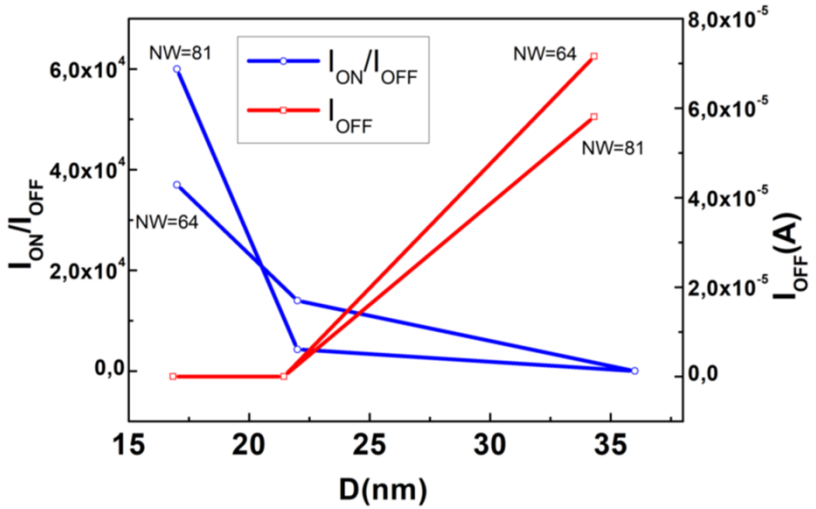
- Two generations of JLNTs
- Improved performance
  - Better subthreshold slope
  - Lower leakage



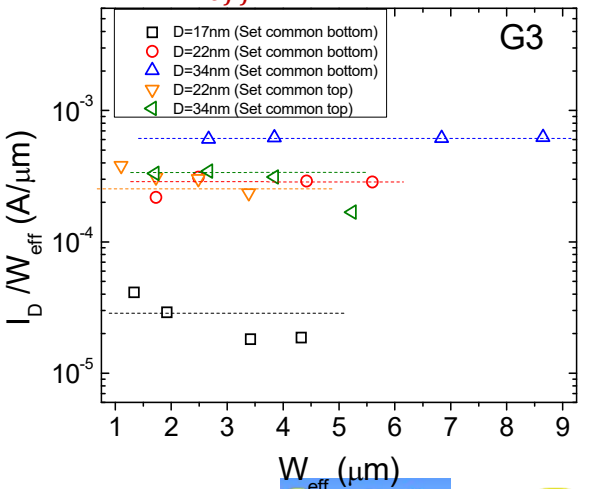
Scaling

- Scalability is crucial for **predictive modeling**
- Increasing NW diameter degrades DIBL, SS, I_{off} due to loss of electrostatic control
- Optimal design criterion: **smaller diameter** for better Immunity against SCE, **increase of NWs in parallel** for higher drive current and better scaling





$$W_{eff} = \pi \times D \times NF$$

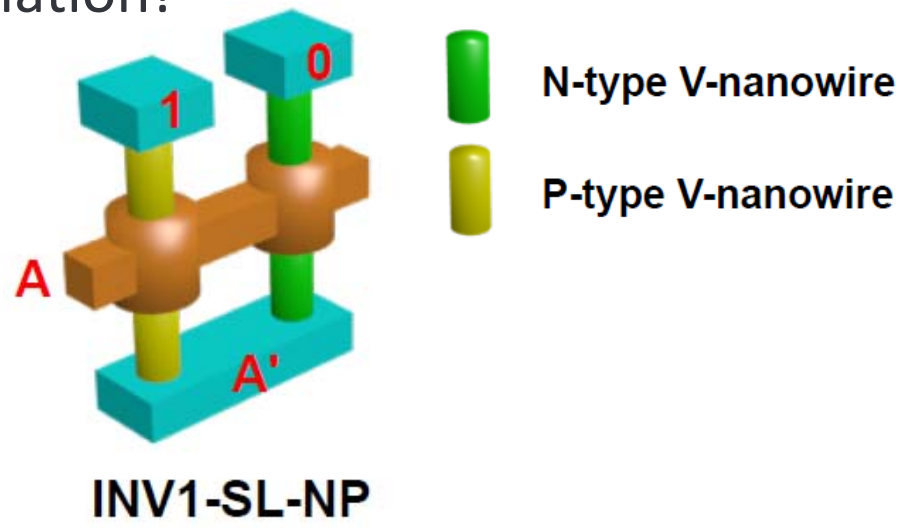


Improvements to Compact Model

- Scaling rules to be improved for **predictive** modeling and circuit design **extrapolation**
- Trap dynamics needs to be studied using **pulse measurements** and modeled for improving drain current behavior at different temperatures
- Electro-thermal effects/**self-heating** and temperature dependence of model parameters (current, threshold voltage , etc.) to be incorporated using an equivalent thermal network.

3D Logic circuits with the VNWFETs

➔ Compact model ready for circuit simulation!



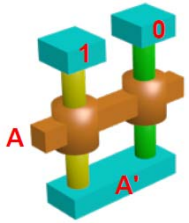
INV1-SL-NP

Outline

- Contexts
- Characterization challenges
- Modelling challenges
- Circuit design challenges

Towards 3D NN

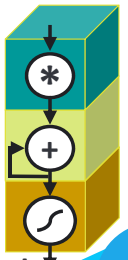
- Dedicated library of 3D logic cells leveraging VNWFET devices
- Versatile VNWFET-logic cell based 3D neural network compute cube (N2C2) for NN-based architecture design
- Scalable and versatile 3D architectural model leveraging reconfigurable 3D interconnect framework
- 3D NN Suitable for exploration of hw/sw co-design



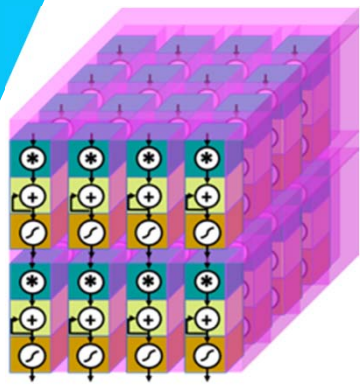
INV1-SL-NP

N-type V-nanowire

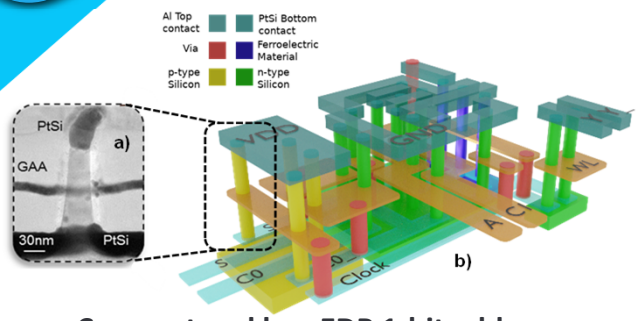
P-type V-nanowire



N^2C^2 concept



3D NN



Compact and low EDP 1-bit adder

https://www.linkedin.com/in/fvllmonti-project/

https://fvllmonti.eu/

Grant n°101016776

FETCH2022 0'Connor

Thanks for listening