STEAM

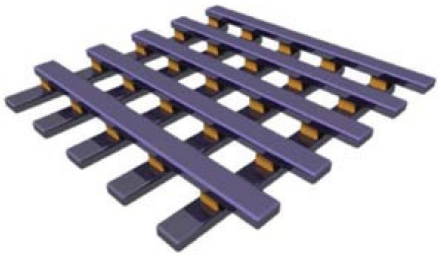
“Spline-based Tables for Efficient and Accurate Device-Modelling”

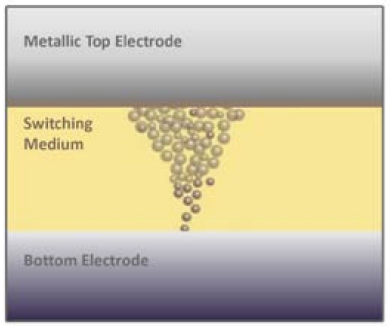
Archit Gupta, Tianshi Wang, Ahmet Gokcen Mahmutoglu, Jaijeet Roychowdhury
EECS Department, UC Berkeley

STEAM is about devices and models

STEAM is about devices and models

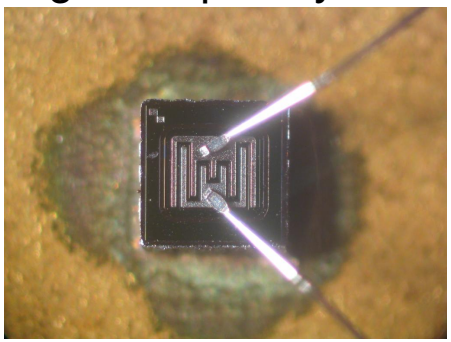
RRAM Crossbar



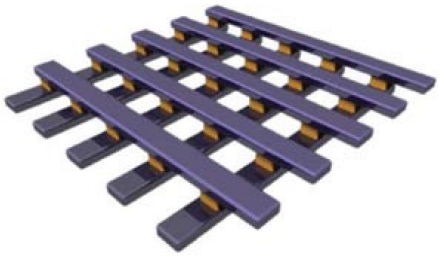


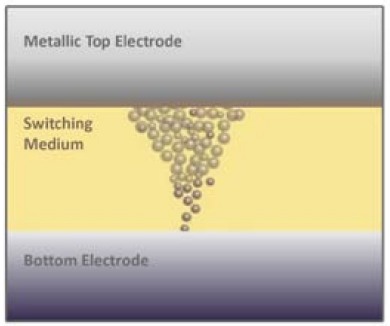
STEAM is about devices and models

High Frequency NPN



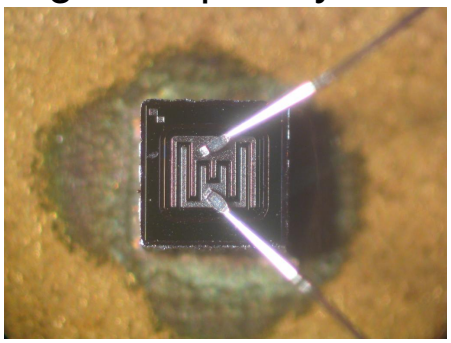
RRAM Crossbar



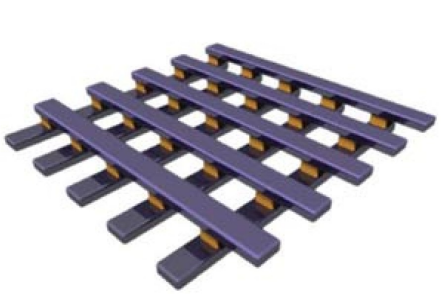


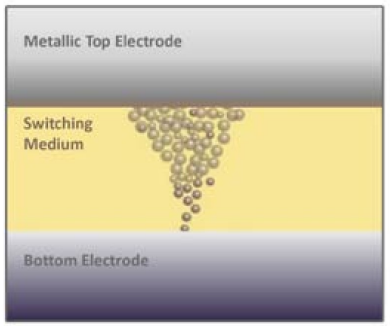
STEAM is about devices and models

High Frequency NPN

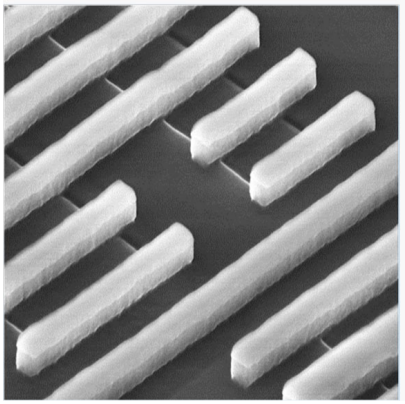


RRAM Crossbar





32nm planar FET



22nm tri-gate FET

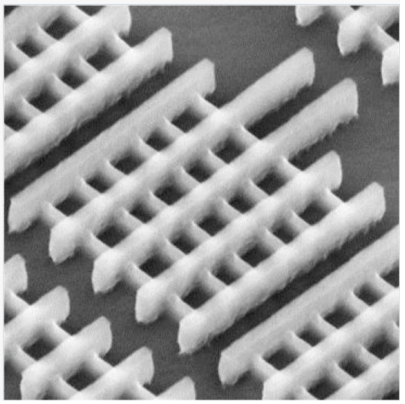
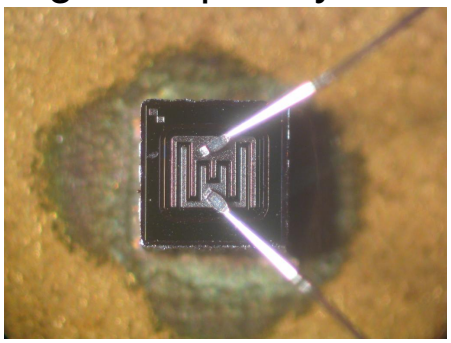


Image Source: Wikipedia

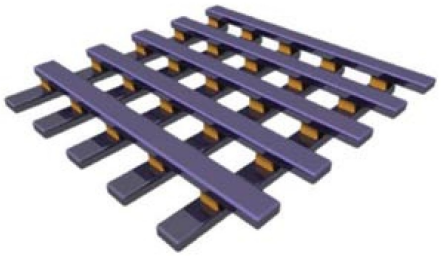
STEAM is about devices and models

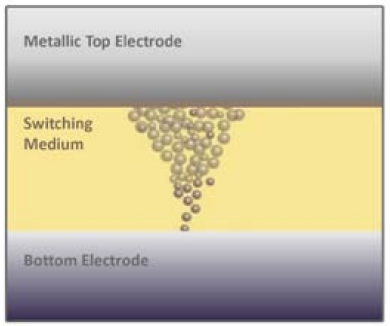
High Frequency NPN



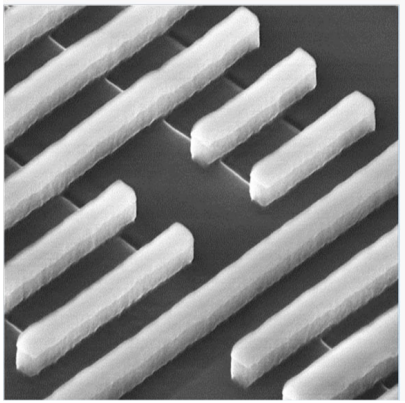
?

RRAM Crossbar





32nm planar FET



22nm tri-gate FET

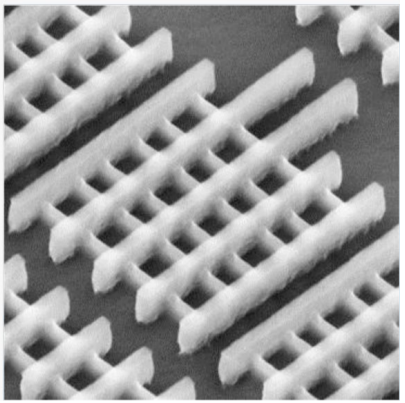
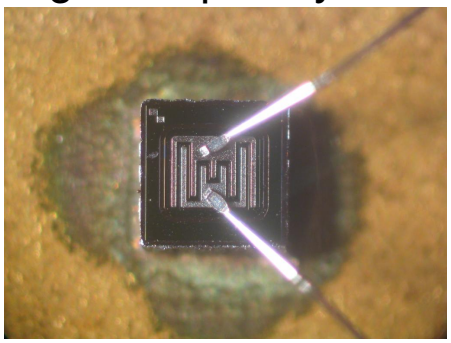


Image Source: Wikipedia

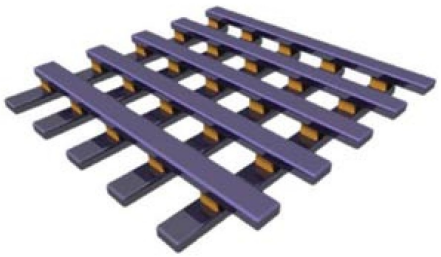
STEAM is about devices and models

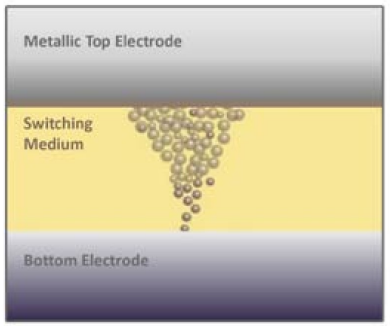
High Frequency NPN



?

RRAM Crossbar

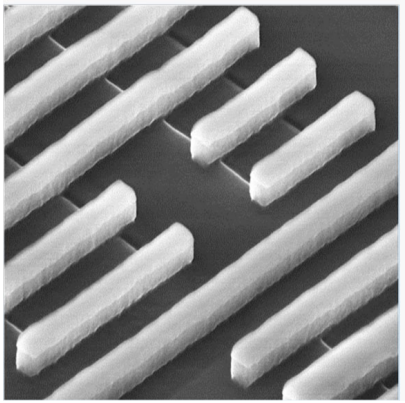





```
  1  function MOD = BSIM3v3_2_4_ModSpec(uniqID)...
1348
1349  %%%%%%%%%%%%%%%%%% NR INITIAL GUESS SUPPORT %%%%%%%%%%%%%%%%%%
1350  function vecLim = initGuess(u, MOD)... % initGuess
1793
1794  %%%%%%%%%%%%%%%%%% NR LIMITING SUPPORT %%%%%%%%%%%%%%%%%%%%%%
1795  function vecLim = limiting...
1796        (vecX,vecY,vecXold,vecYold, u, MOD)... % limiting
2300
2301  %%%%%%%%%%%%%%%%%% CORE DEVICE EVAL FUNCTIONS %%%%%%%%%%%%%%%%
2302  function [fe, qe, fi, qi] = fqei_all(vecX, vecY, vecLim, u, flag, MOD)...
5088
```

```
  1  function MOD = BSIM3v3_2_4d_ModSpec(uniqID)...
458
459  function [fqei_out, J_out] = fqeiJ(vecX, vecY, vecLim, vecU, flag, MOD)...
489
490    function [fe__, qe__, fi__, qi__,...
491              d_fe_d_X__, d_qe_d_X__, d_fi_d_X__, d_qi_d_X__,...
492              d_fe_d_Y__, d_qe_d_Y__, d_fi_d_Y__, d_qi_d_Y__] = ...
493           fqei_dfqeidXYU(vecX__, vecY__, MOD)...
4535
```

32nm planar FET



22nm tri-gate FET

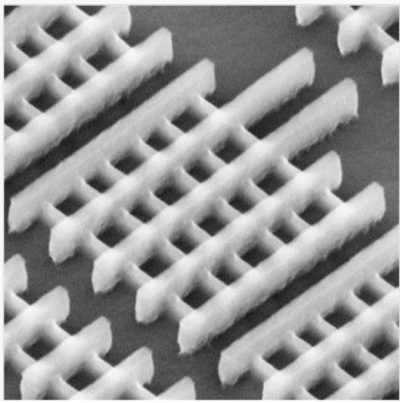
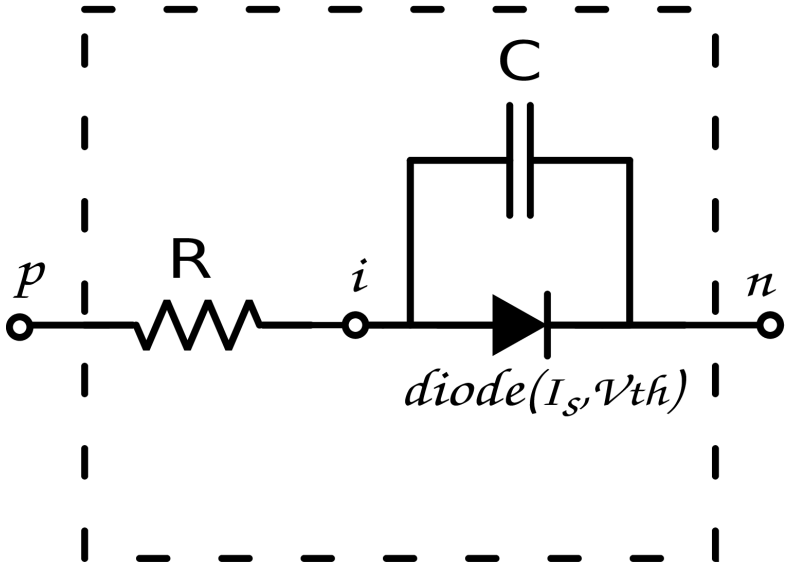


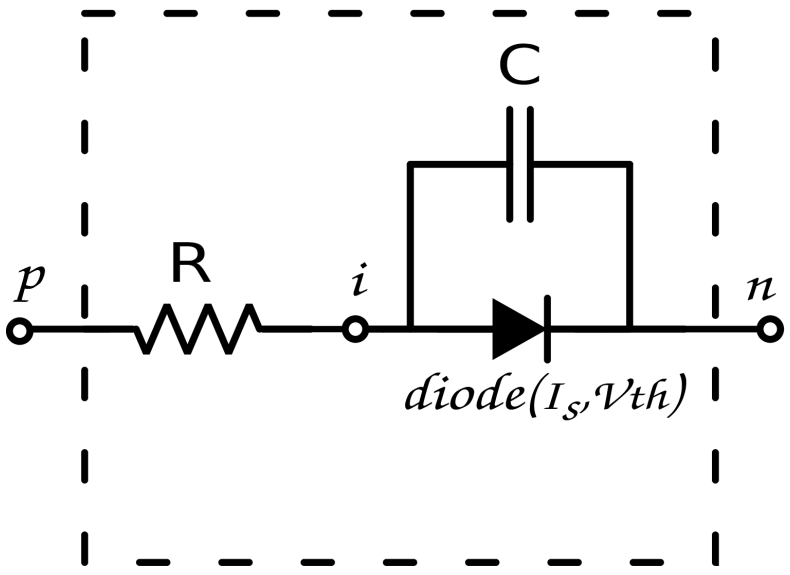
Image Source: Wikipedia

What constitutes a device model?

What constitutes a device model?

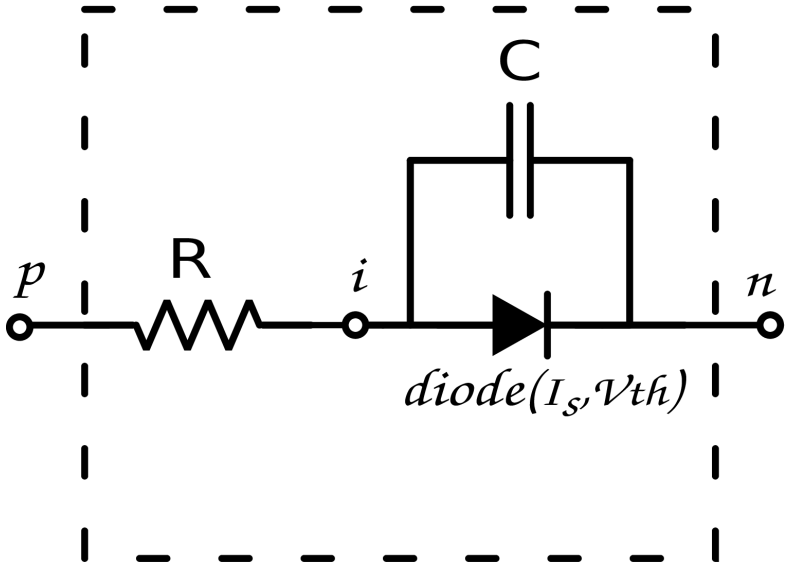


What constitutes a device model?



Simulator (equation engine...) gives v_{pn}, v_{in}

What constitutes a device model?

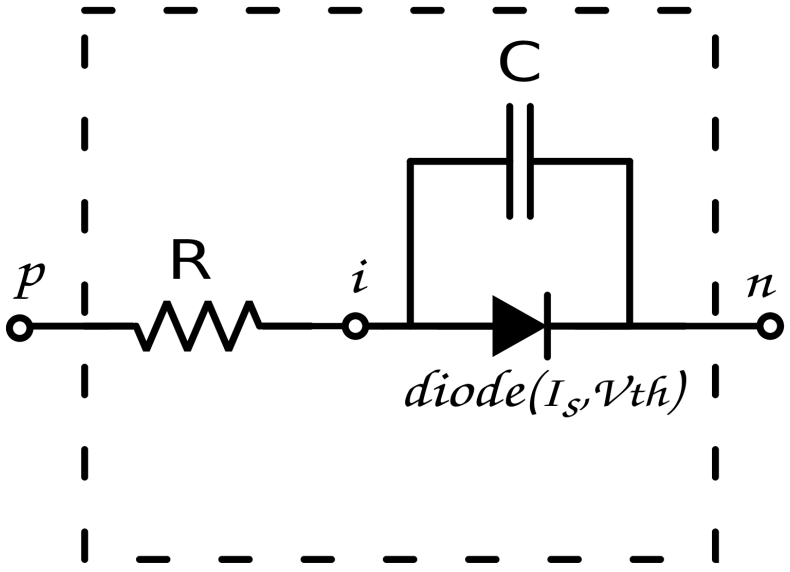


Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

What constitutes a device model?

i_{pn}: “explicit output”

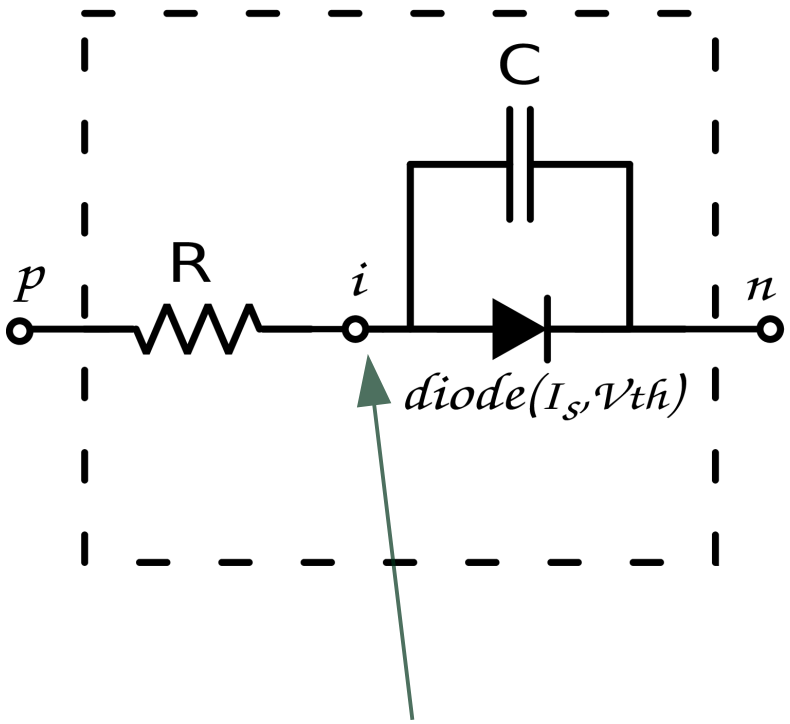


Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

What constitutes a device model?

i_{pn}: “explicit output”

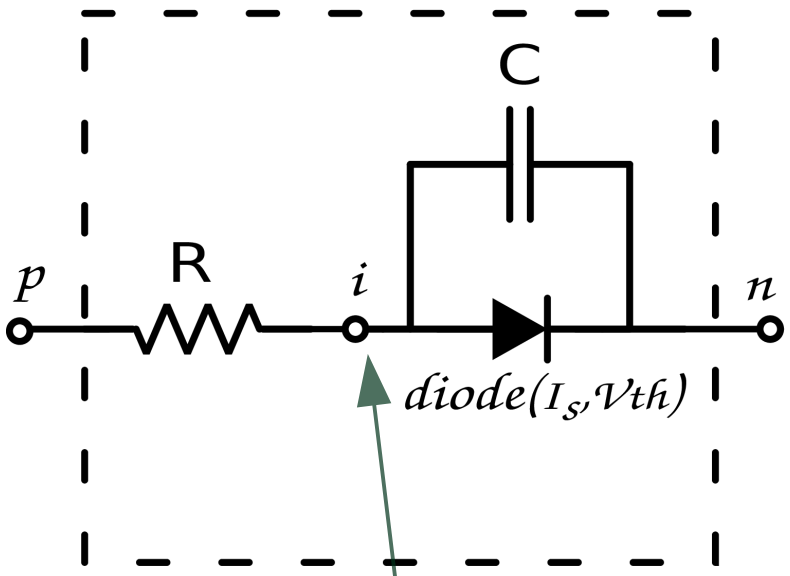


Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

What constitutes a device model?

i_{pn}: “explicit output”



v_{in}: “internal unknown”

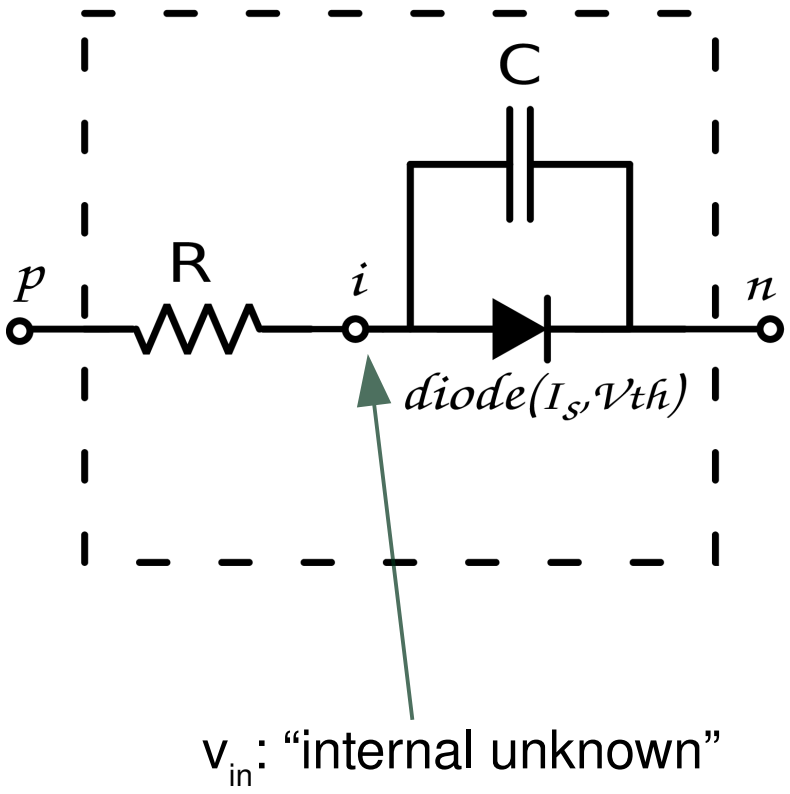
Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

What constitutes a device model?

i_{pn}: “explicit output”

v_{pn}: “other IO”



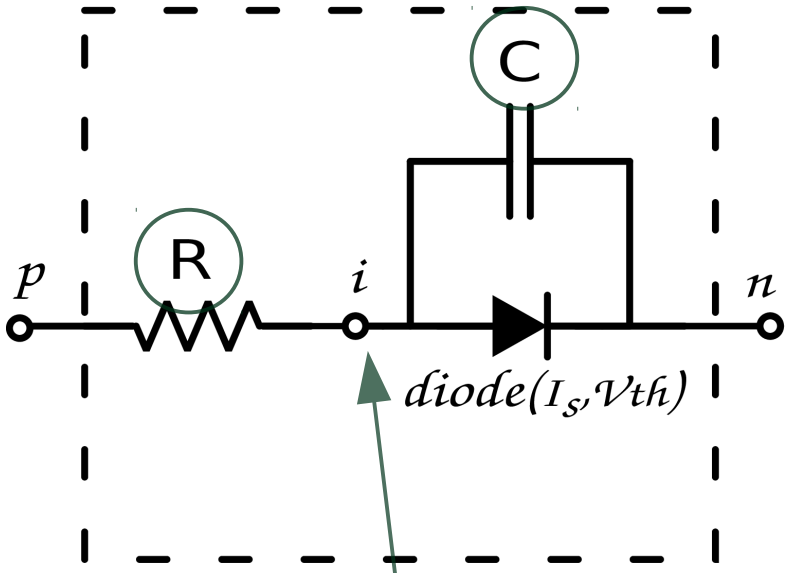
Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

What constitutes a device model?

i_{pn}: “explicit output”

v_{pn}: “other IO”

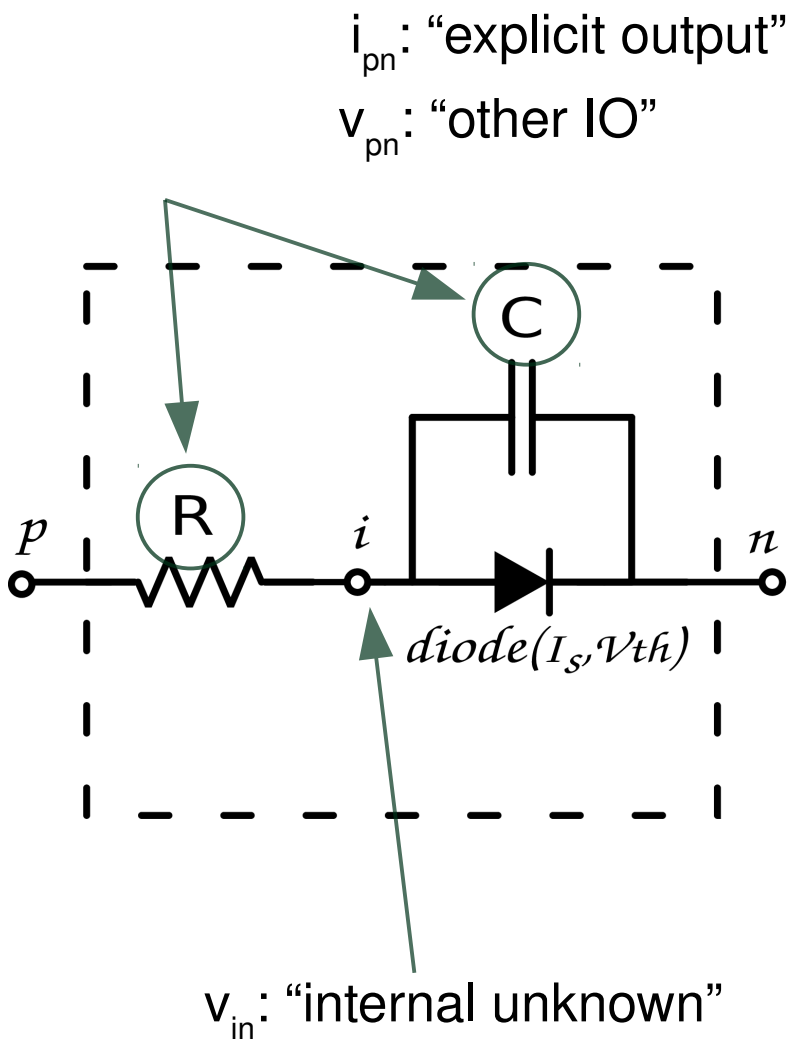


v_{in}: “internal unknown”

Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

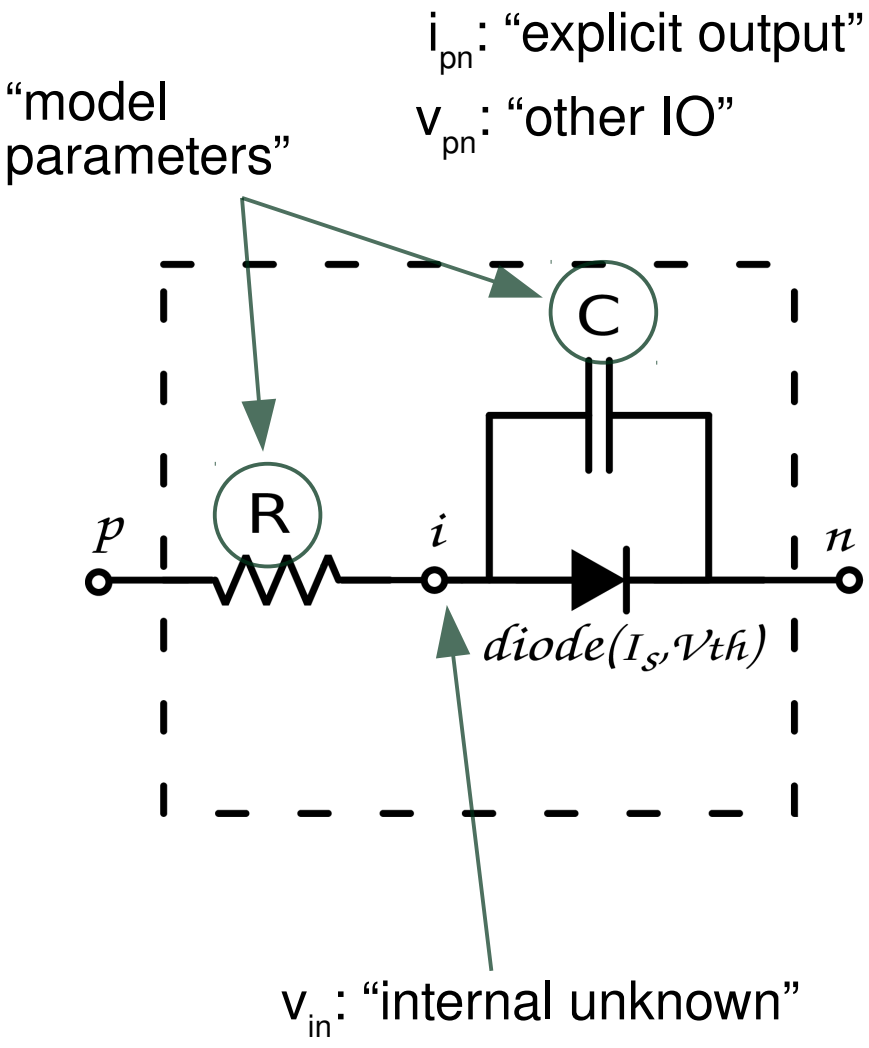
What constitutes a device model?



Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

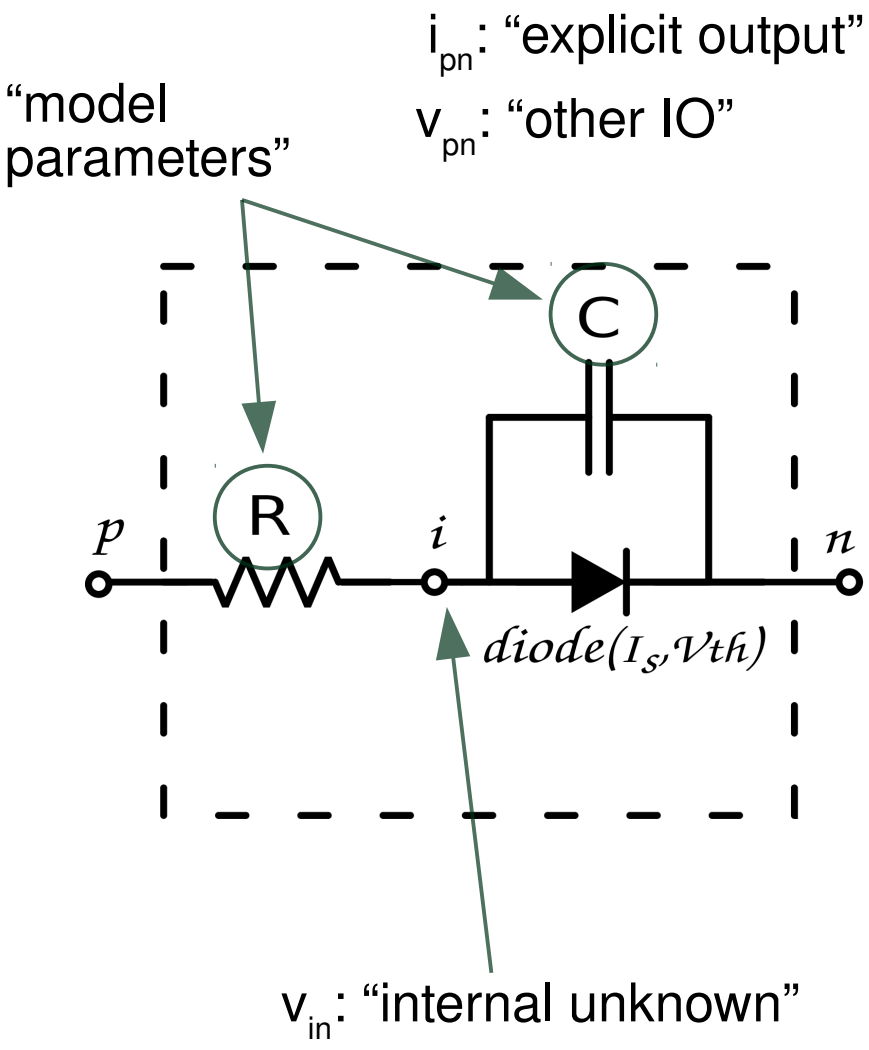
What constitutes a device model?



Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

What constitutes a device model?

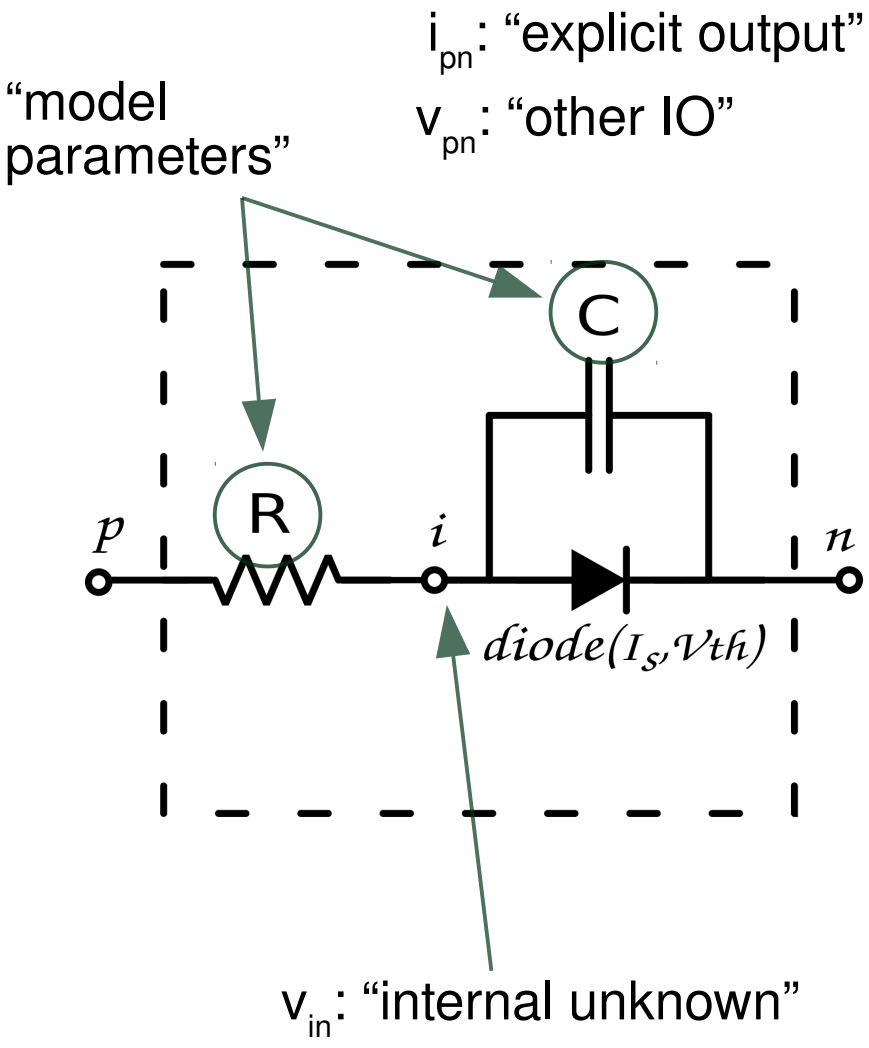


Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

Usually, it also wants some derivatives

What constitutes a device model?



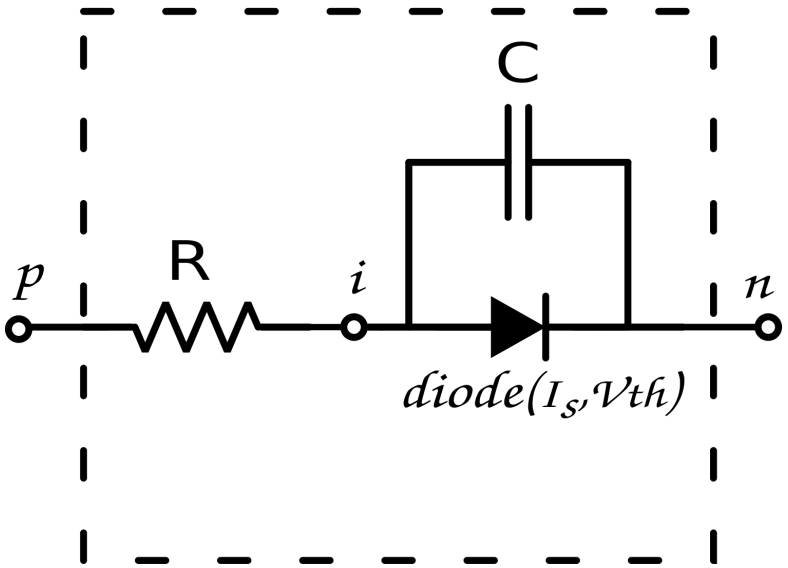
Simulator (equation engine...) gives v_{pn}, v_{in}

It wants i_{pn} to solve a system of equations!

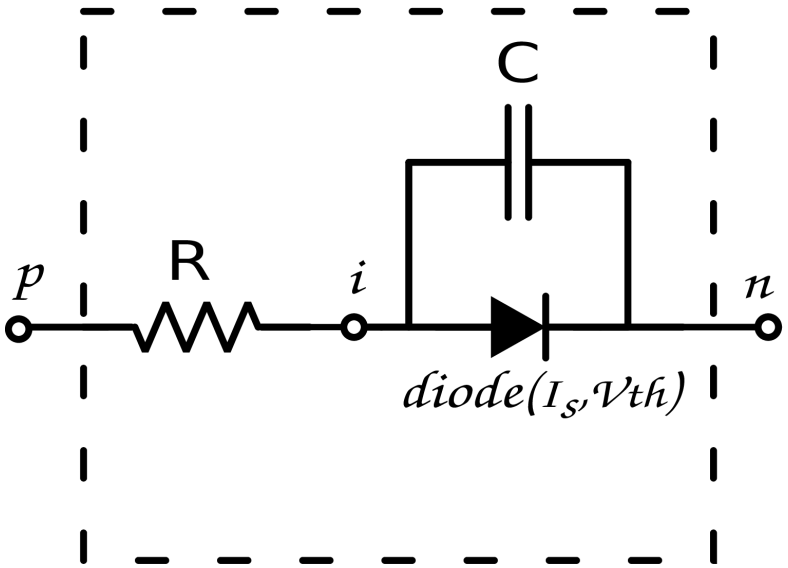
Usually, it also wants some derivatives

We want some sort of Look-Up table to give us i_{pn} and derivatives

Dynamics are governed by Differential Algebraic Equations

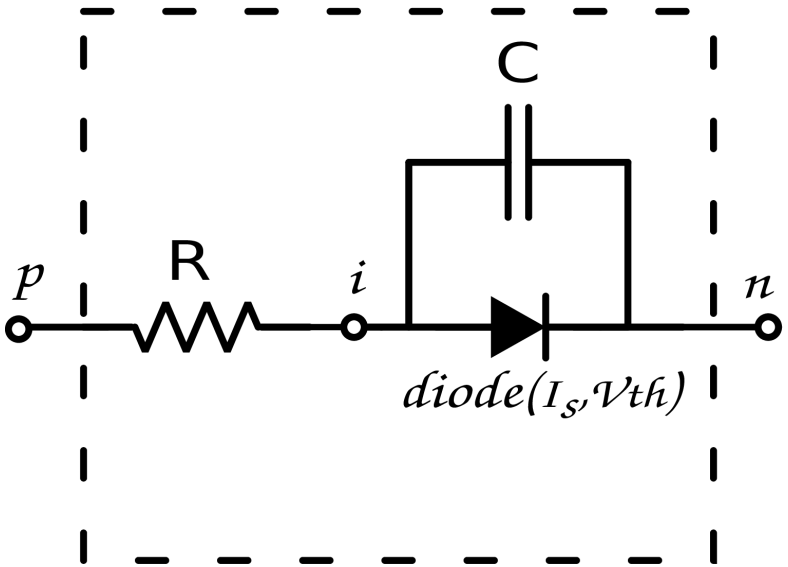


Dynamics are governed by Differential Algebraic Equations



$$\frac{d}{dt}0+\frac{v_{pn}-v_{in}}{R}=i_{pn}$$

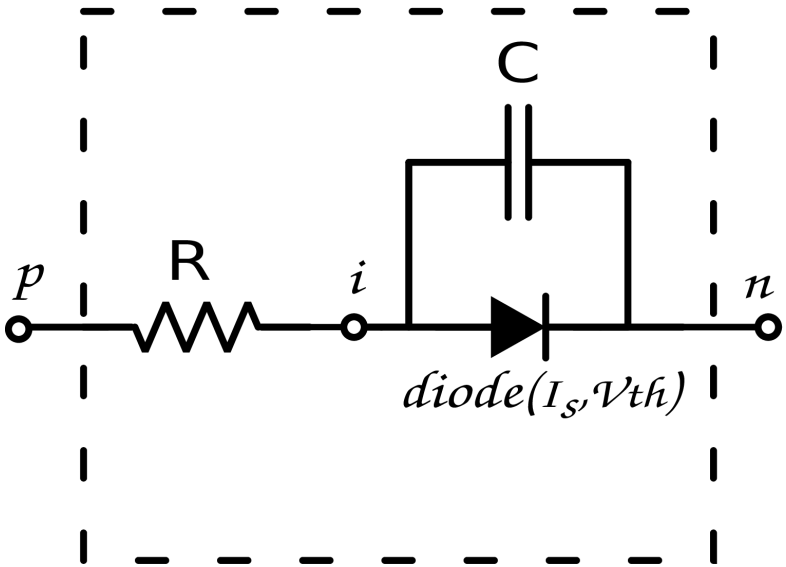
Dynamics are governed by Differential Algebraic Equations



$$\frac{d}{dt}0+\frac{v_{pn}-v_{in}}{R}=i_{pn}$$

$$\frac{d}{dt}Cv_{in}+diode(v_{in})+\frac{v_{in}-v_{pn}}{R}=0$$

Dynamics are governed by Differential Algebraic Equations

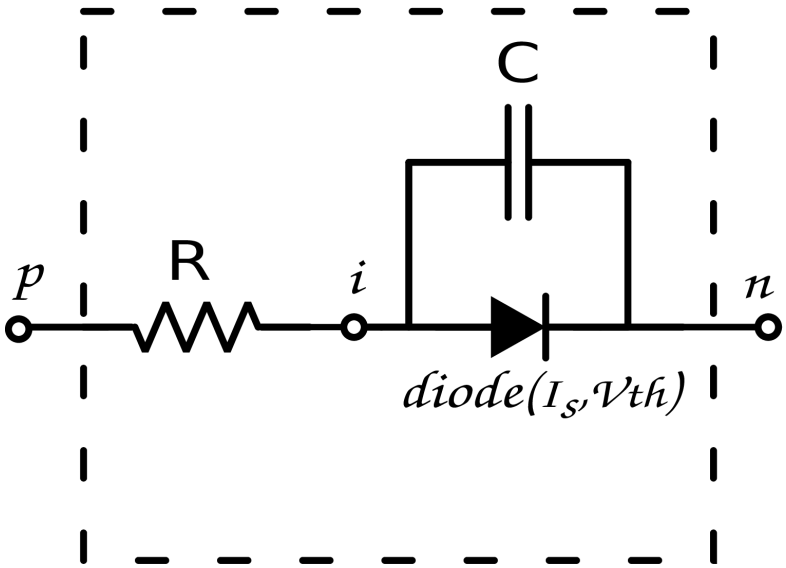


$$\frac{d}{dt}0+\frac{v_{pn}-v_{in}}{R}=i_{pn}$$

$$\frac{d}{dt}Cv_{in}+diode(v_{in})+\frac{v_{in}-v_{pn}}{R}=0$$

This can be generalized to practically any circuit or model that we are interested in.

Dynamics are governed by Differential Algebraic Equations



$$\frac{d}{dt}0+\frac{v_{pn}-v_{in}}{R}=i_{pn}$$

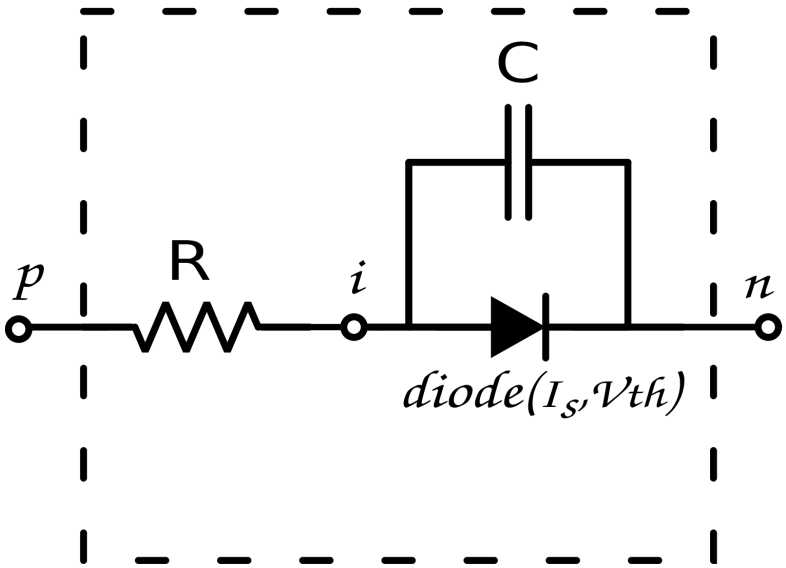
$$\frac{d}{dt}Cv_{in}+diode(v_{in})+\frac{v_{in}-v_{pn}}{R}=0$$

This can be generalized to practically any circuit or model that we are interested in.

$$\frac{d}{dt}q_e(\ldots)+f_e(\ldots)=z(t)$$

$$\frac{d}{dt}q_i(\ldots)+f_i(\ldots)=0$$

Dynamics are governed by Differential Algebraic Equations



$$\frac{d}{dt}0+\frac{v_{pn}-v_{in}}{R}=i_{pn}$$

$$\frac{d}{dt}Cv_{in}+diode(v_{in})+\frac{v_{in}-v_{pn}}{R}=0$$

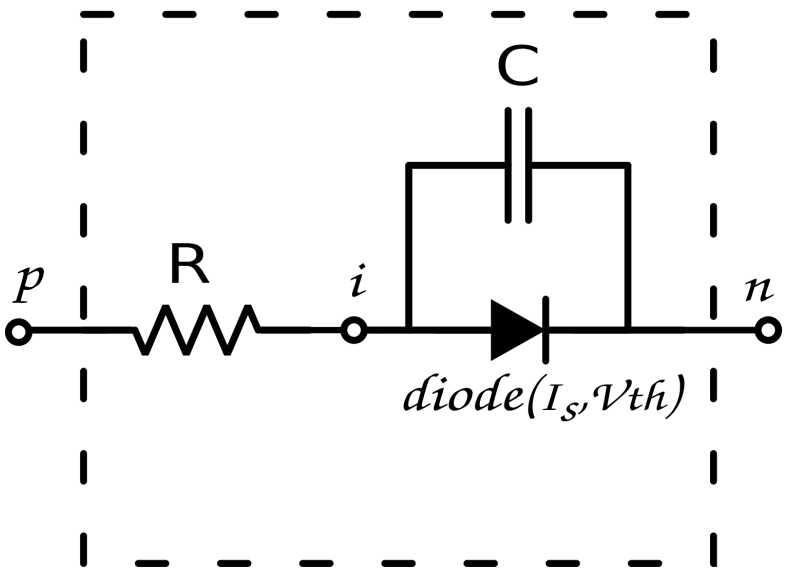
This can be generalized to practically any circuit or model that we are interested in.

$$\frac{d}{dt}q_e(\ldots)+f_e(\ldots)=z(t)$$

$$\frac{d}{dt}q_i(\ldots)+f_i(\ldots)=0$$

i_{pn}: “explicit output”

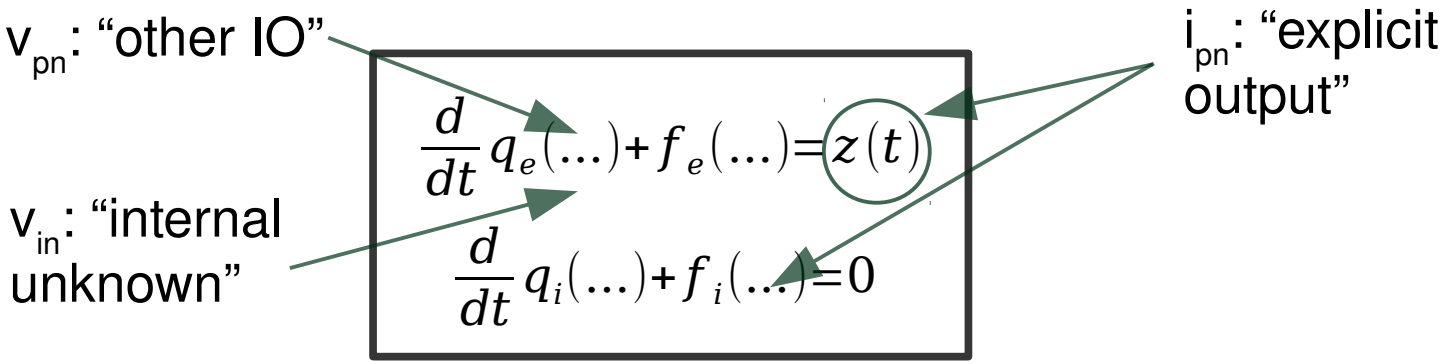
Dynamics are governed by Differential Algebraic Equations



$$\frac{d}{dt}0+\frac{v_{pn}-v_{in}}{R}=i_{pn}$$

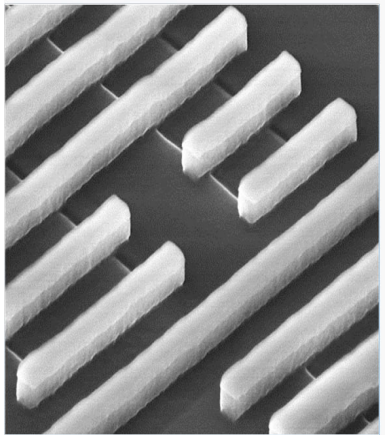
$$\frac{d}{dt}Cv_{in}+diode(v_{in})+\frac{v_{in}-v_{pn}}{R}=0$$

This can be generalized to practically any circuit or model that we are interested in.

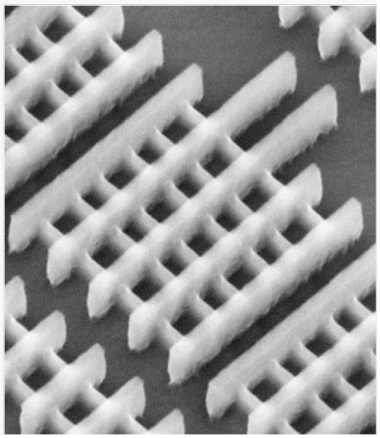


A device is a set of functions

32 nm Planar Transistors

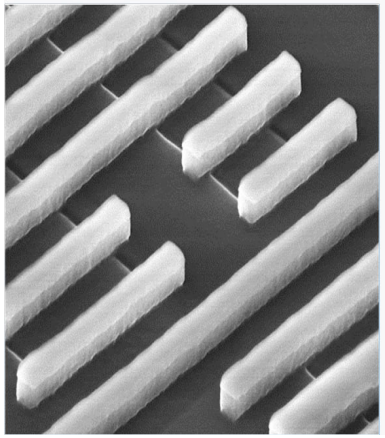


22 nm Tri-Gate Transistors

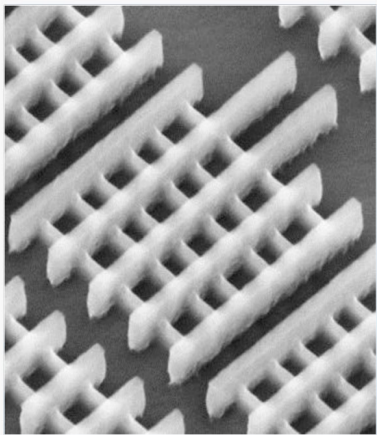


A device is a set of functions

32 nm Planar Transistors



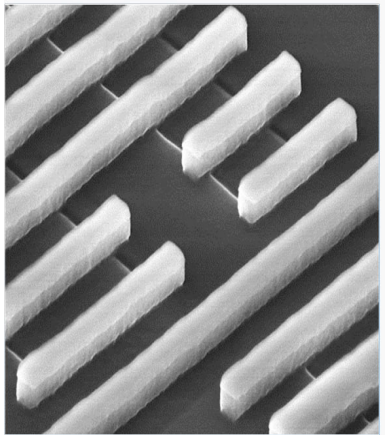
22 nm Tri-Gate Transistors



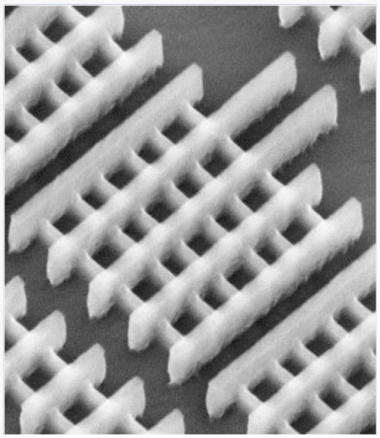
{

A device is a set of functions

32 nm Planar Transistors



22 nm Tri-Gate Transistors

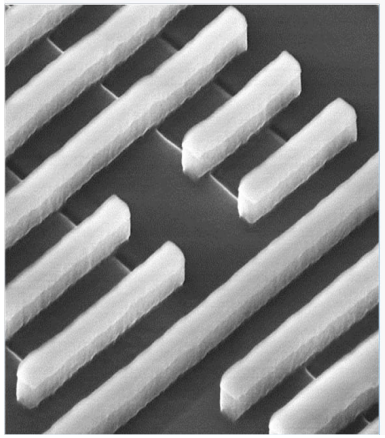


{

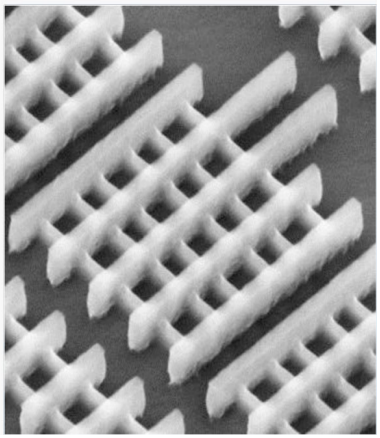
Model parameters,

A device is a set of functions

32 nm Planar Transistors



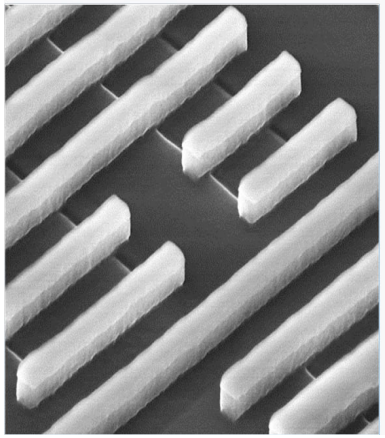
22 nm Tri-Gate Transistors



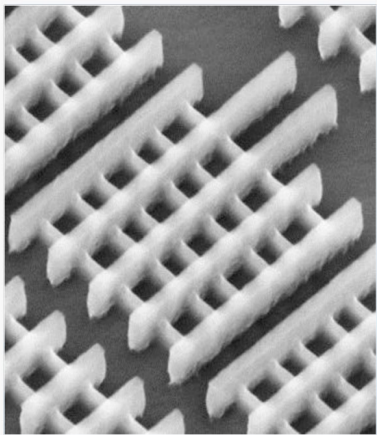
{
Model parameters,
Internal unknown(s),

A device is a set of functions

32 nm Planar Transistors



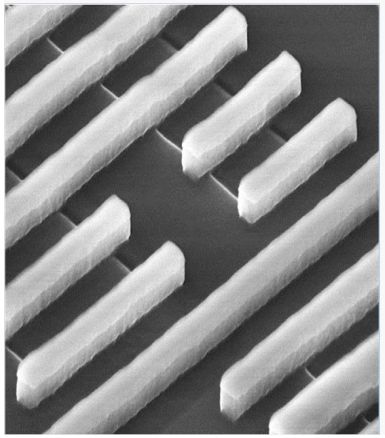
22 nm Tri-Gate Transistors



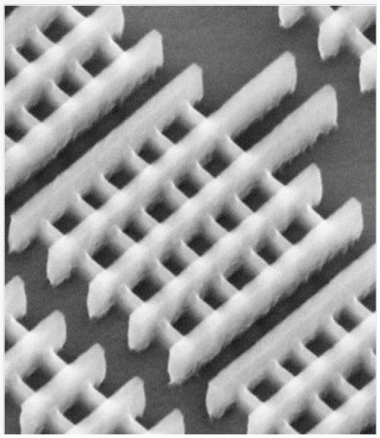
{
Model parameters,
Internal unknown(s),
Other IO(s), …

A device is a set of functions

32 nm Planar Transistors



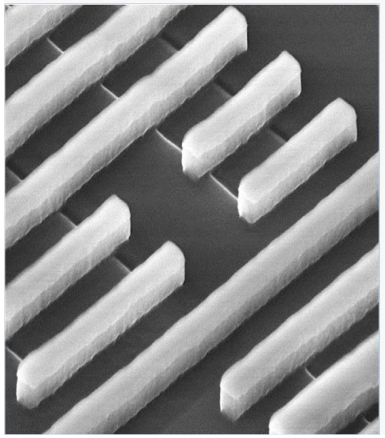
22 nm Tri-Gate Transistors



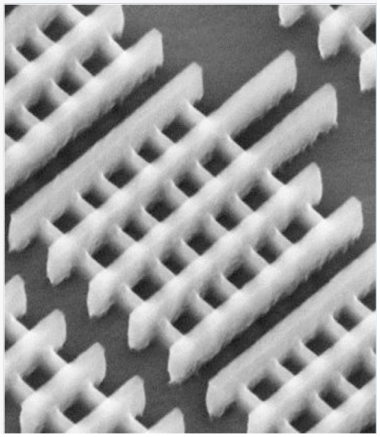
{
Model parameters,
Internal unknown(s),
Other IO(s), …
}

A device is a set of functions

32 nm Planar Transistors



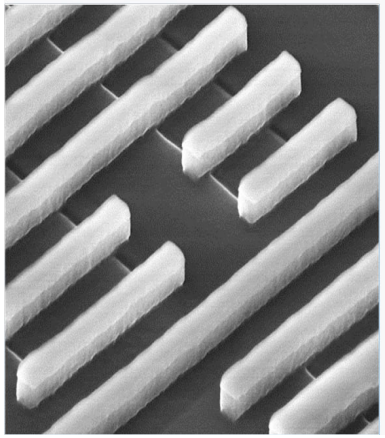
22 nm Tri-Gate Transistors



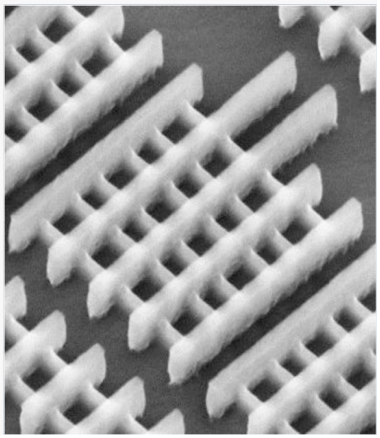
{
Model parameters,
Internal unknown(s),
Other IO(s), …
}

A device is a set of functions

32 nm Planar Transistors



22 nm Tri-Gate Transistors

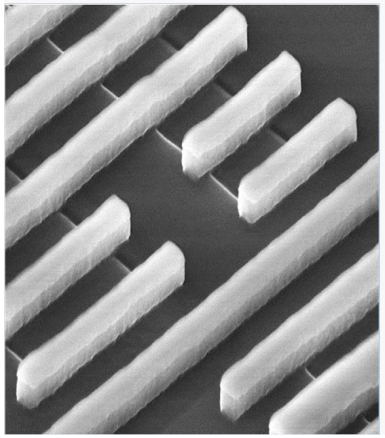


{
Model parameters,
Internal unknown(s),
Other IO(s), …
}

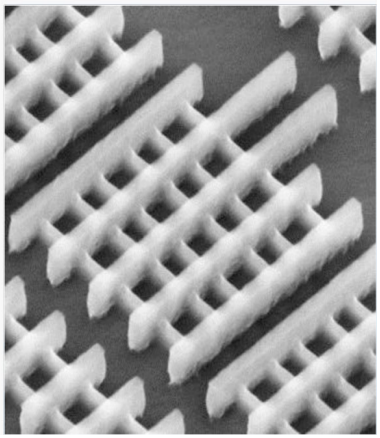
{

A device is a set of functions

32 nm Planar Transistors



22 nm Tri-Gate Transistors

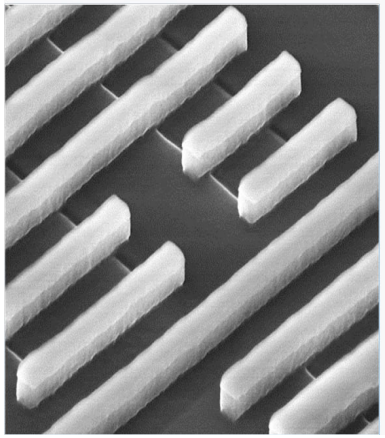


{
Model parameters,
Internal unknown(s),
Other IO(s), …
}

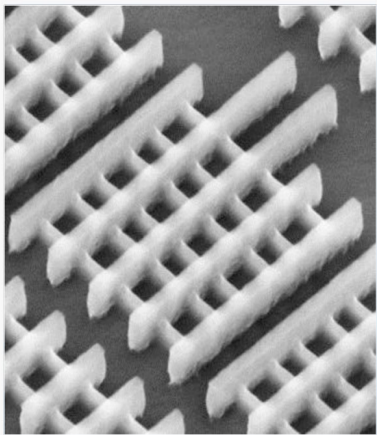
{
Explicit outputs,

A device is a set of functions

32 nm Planar Transistors



22 nm Tri-Gate Transistors

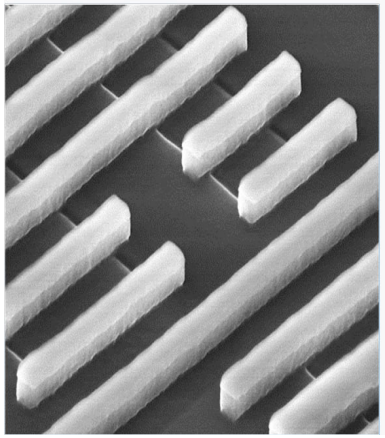


{
Model parameters,
Internal unknown(s),
Other IO(s), …
}

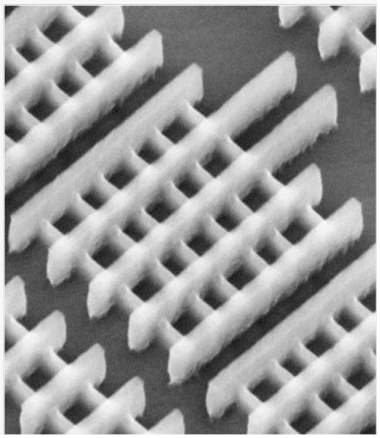
{
Explicit outputs,
Functions, ...

A device is a set of functions

32 nm Planar Transistors



22 nm Tri-Gate Transistors

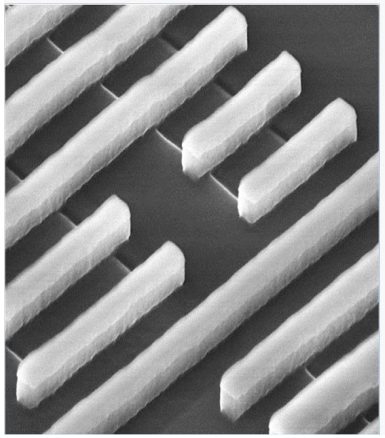


{
Model parameters,
Internal unknown(s),
Other IO(s), …
}

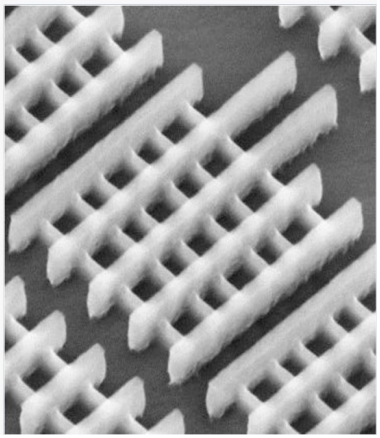
{
Explicit outputs,
Functions, ...
}

A device is a set of functions

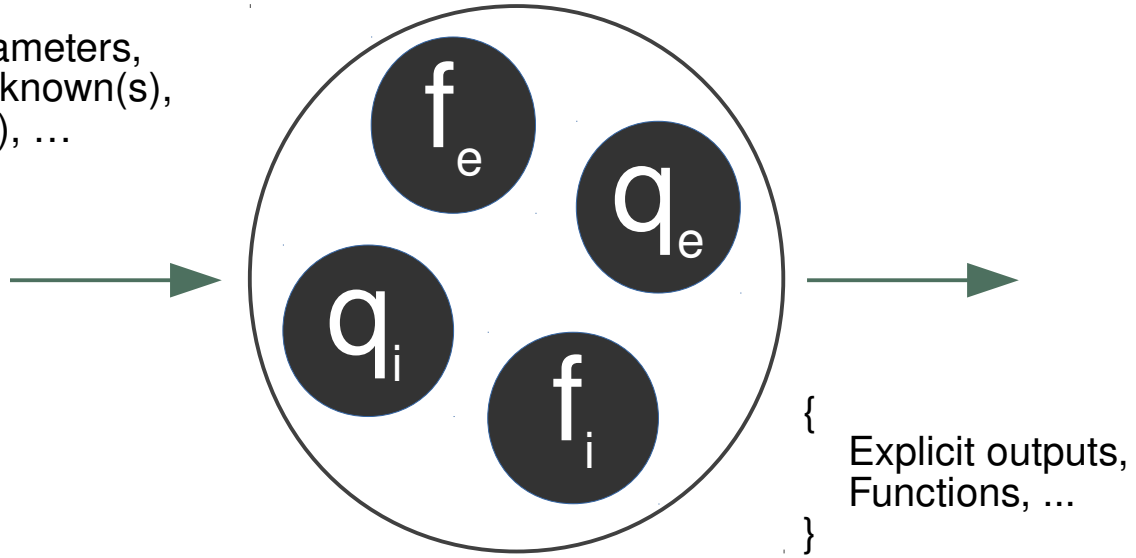
32 nm Planar Transistors



22 nm Tri-Gate Transistors



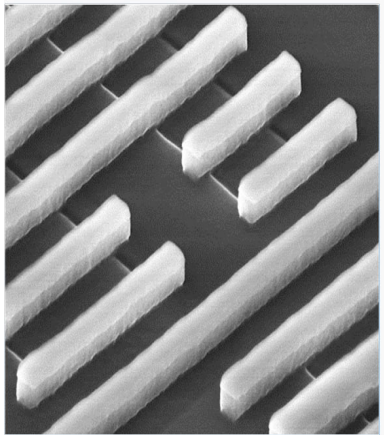
{
Model parameters,
Internal unknown(s),
Other IO(s), …
}



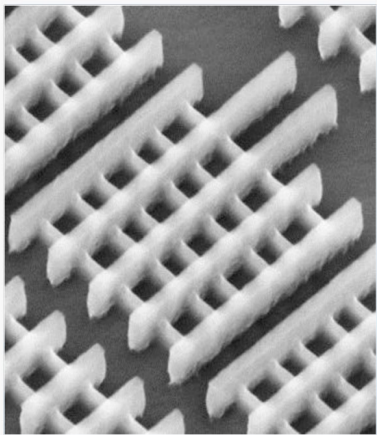
{
Explicit outputs,
Functions, ...
}

A device is a set of functions

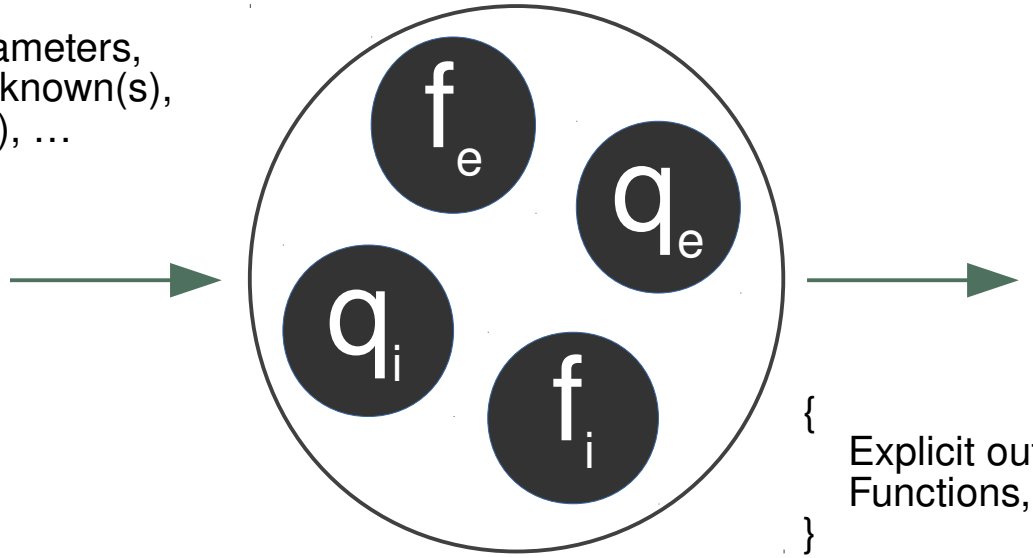
32 nm Planar Transistors



22 nm Tri-Gate Transistors



{
Model parameters,
Internal unknown(s),
Other IO(s), …
}



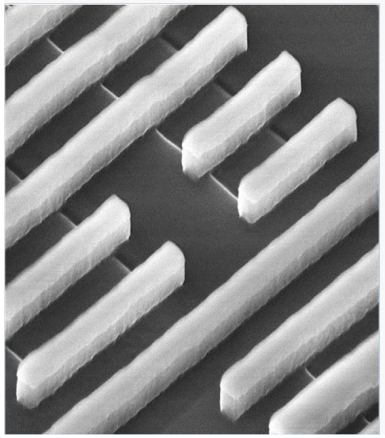
{
Explicit outputs,
Functions, ...
}

$$\frac{d}{dt} q_e(\ldots) + f_e(\ldots) = z(t)$$

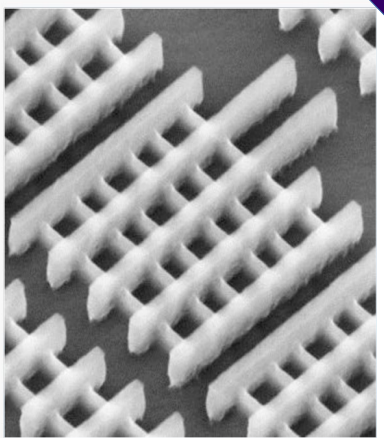
$$\frac{d}{dt} q_i(\ldots) + f_i(\ldots) = 0$$

A device is a set of functions

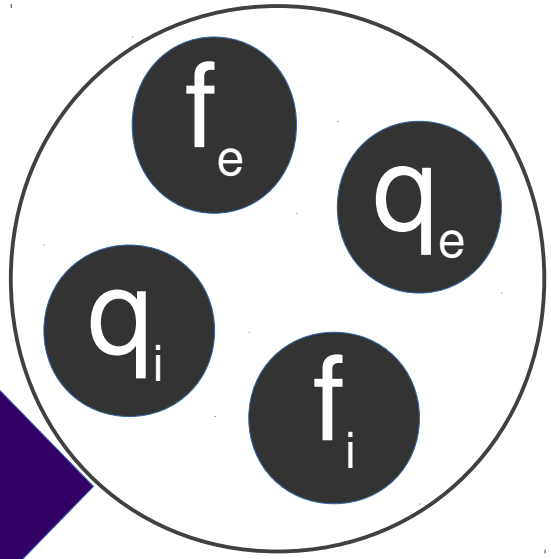
32 nm Planar Transistors



22 nm Tri-Gate Transistors



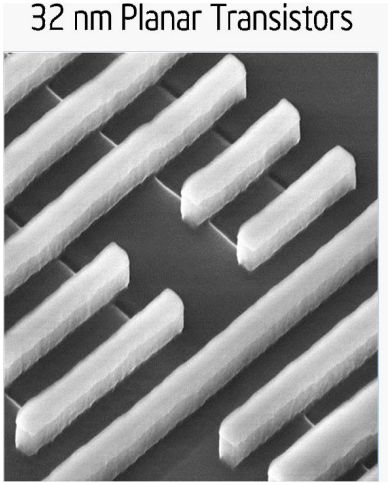
MODSPEC



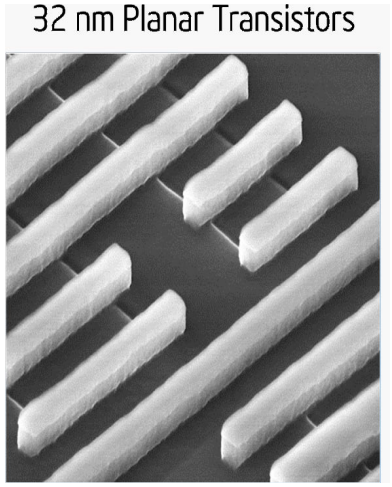
$$\frac{d}{dt} q_e(\ldots) + f_e(\ldots) = z(t)$$

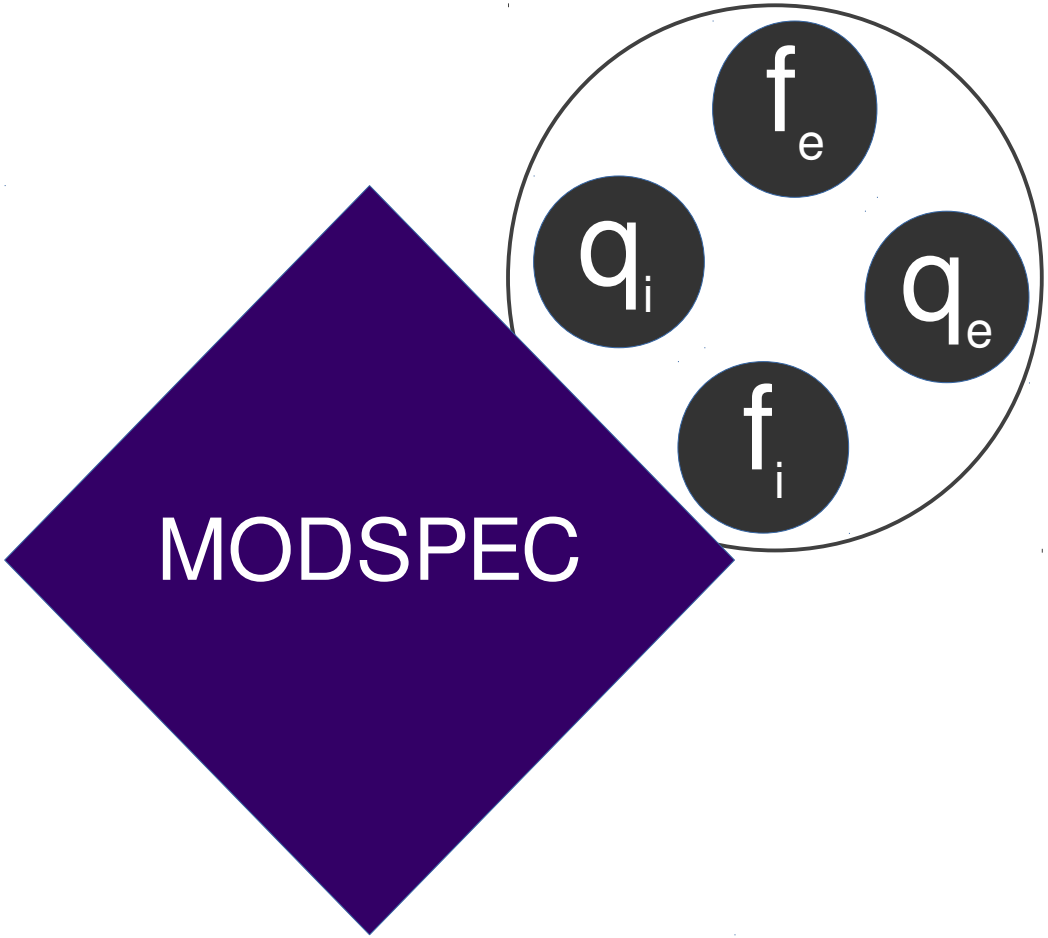
$$\frac{d}{dt} q_i(\ldots) + f_i(\ldots) = 0$$

From ModSpec to table-based model

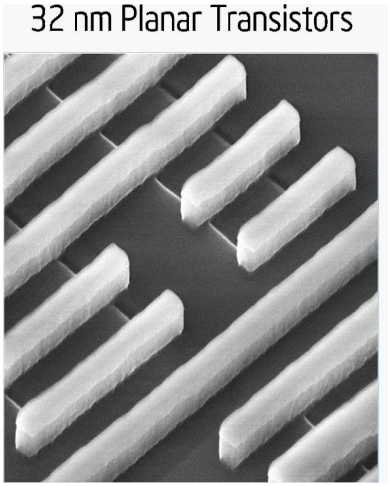


From ModSpec to table-based model

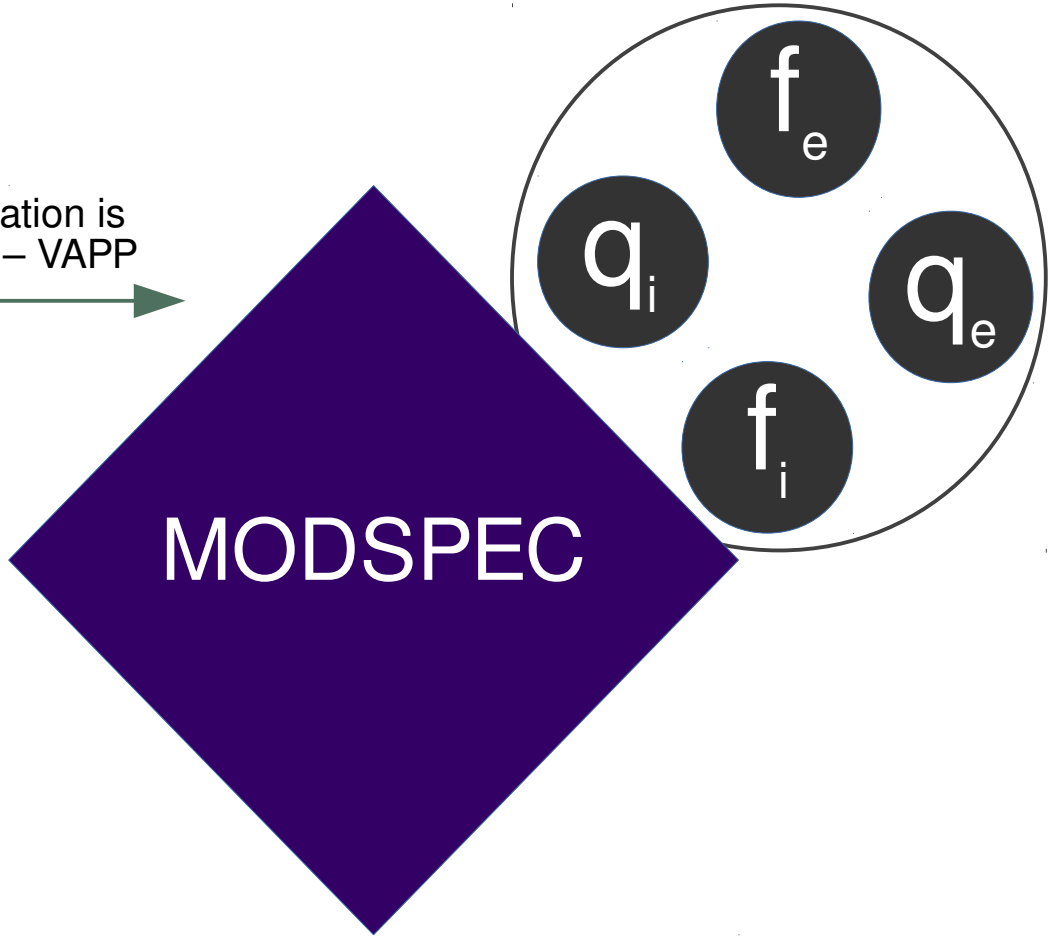




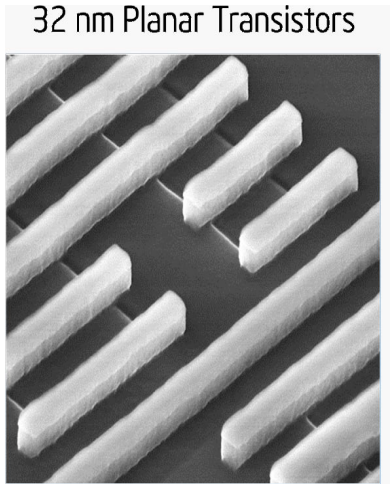
From ModSpec to table-based model



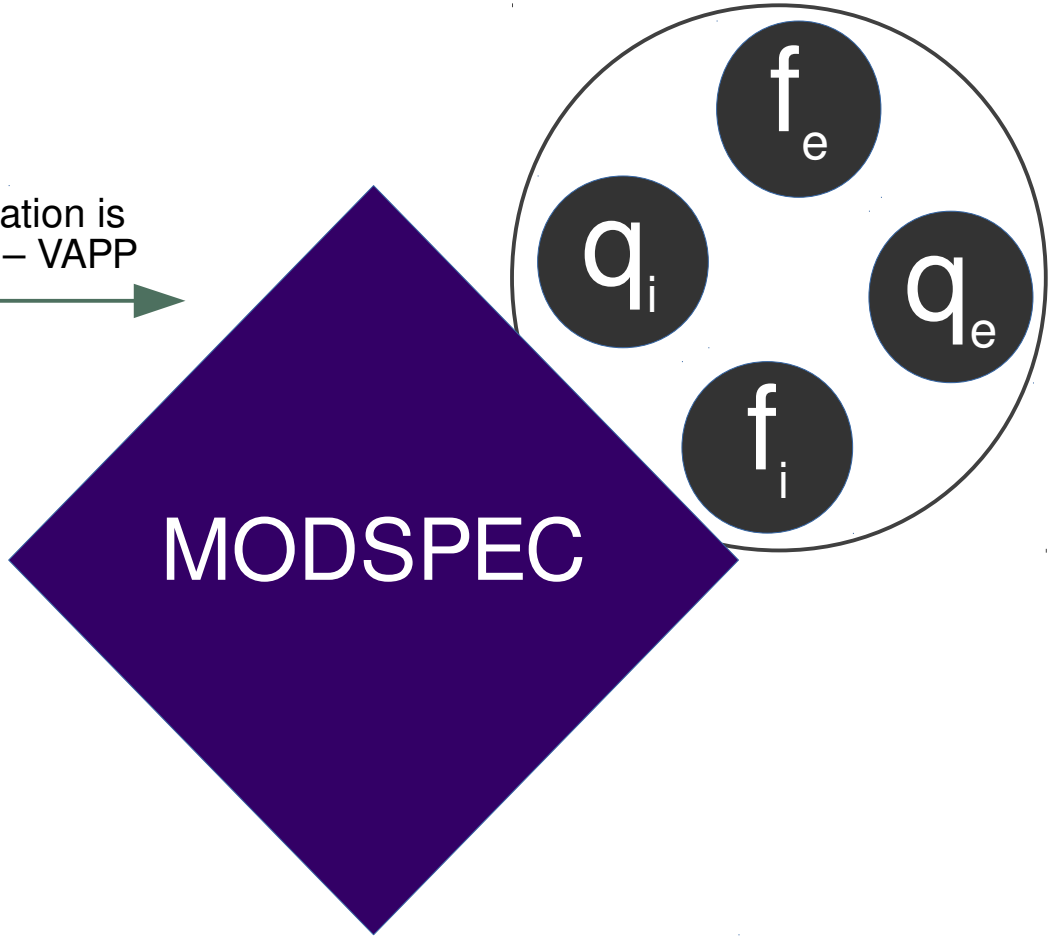
Machine Translation is openly available – VAPP



From ModSpec to table-based model

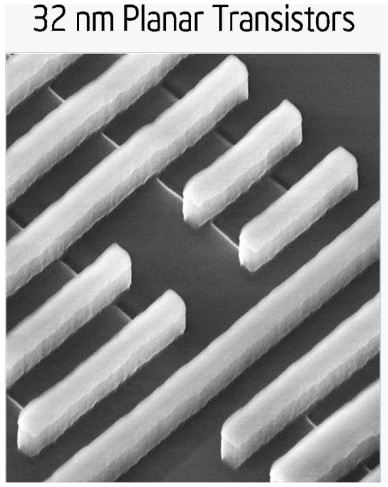


Machine Translation is openly available – VAPP

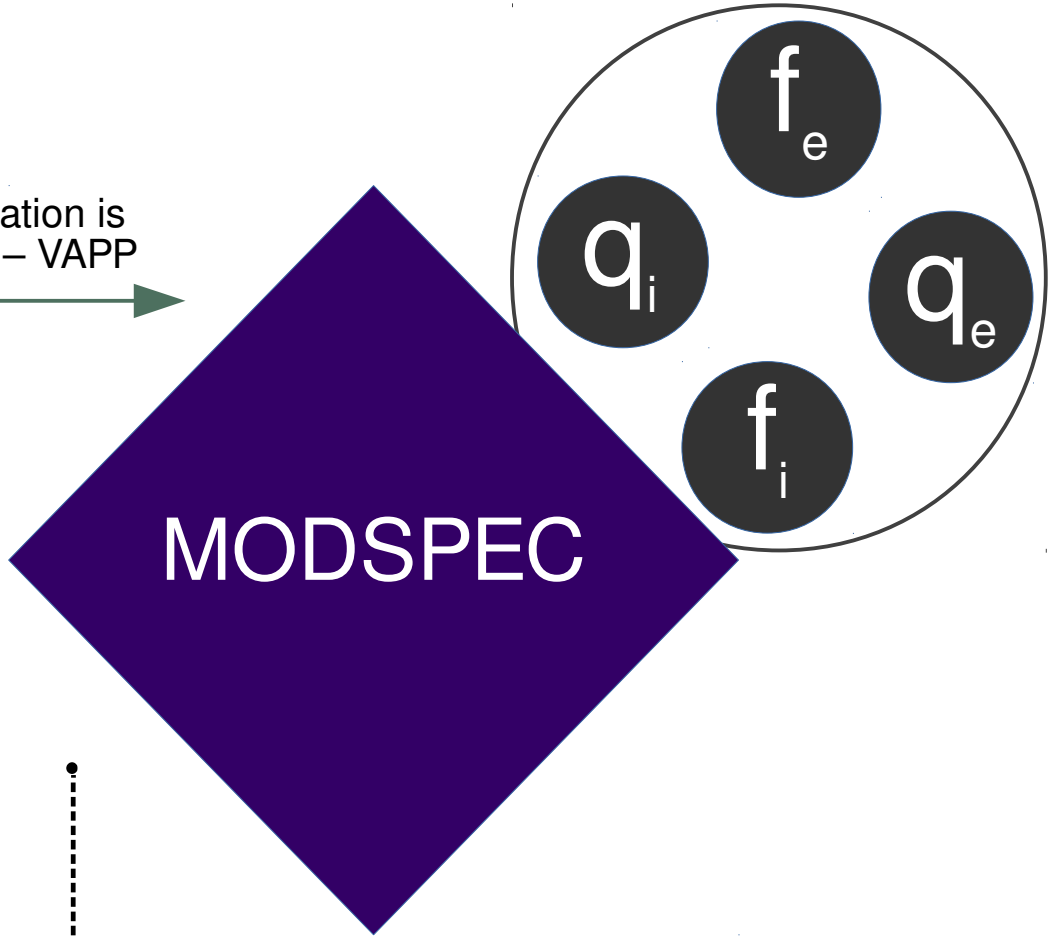


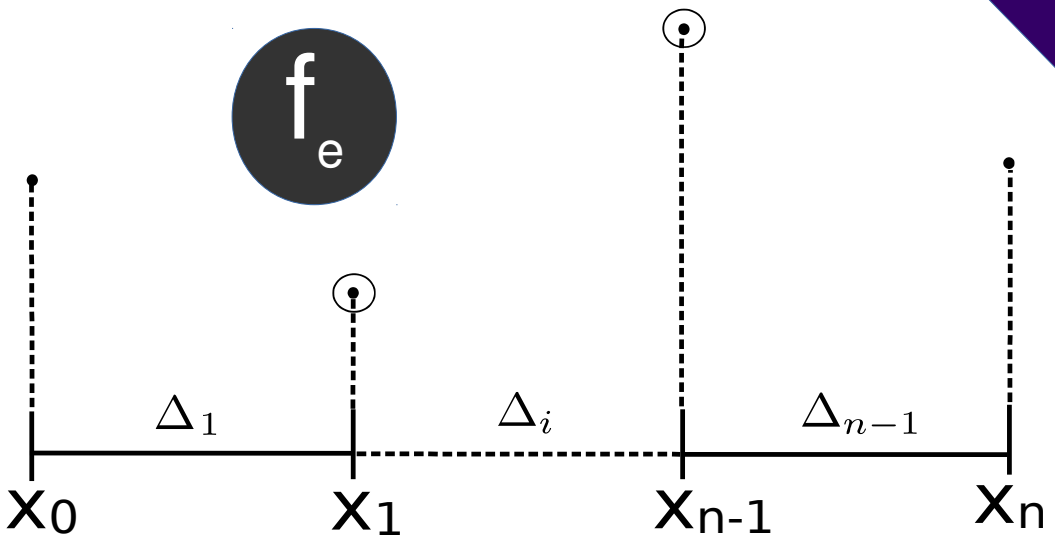
f_e

From ModSpec to table-based model

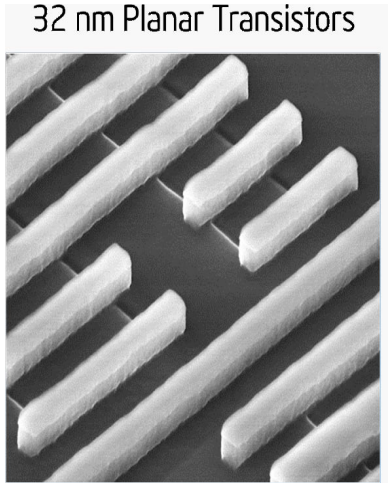


Machine Translation is openly available – VAPP

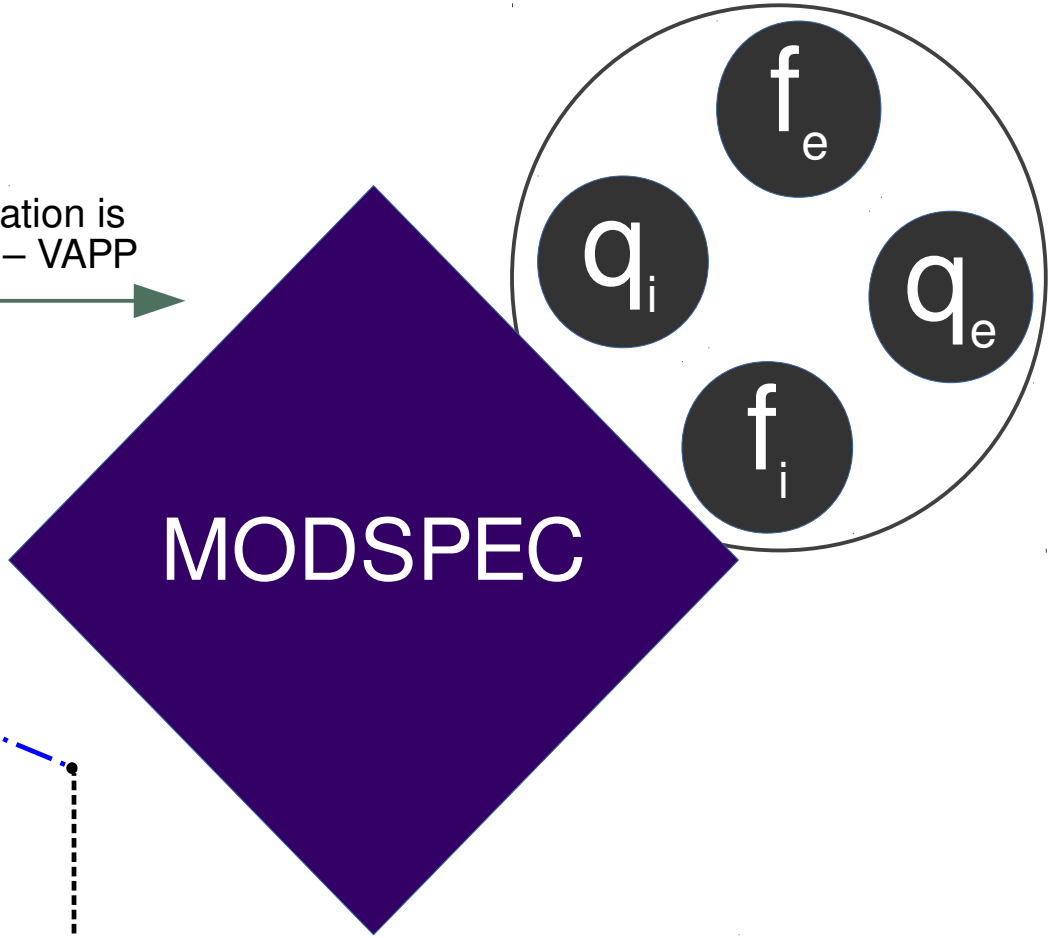


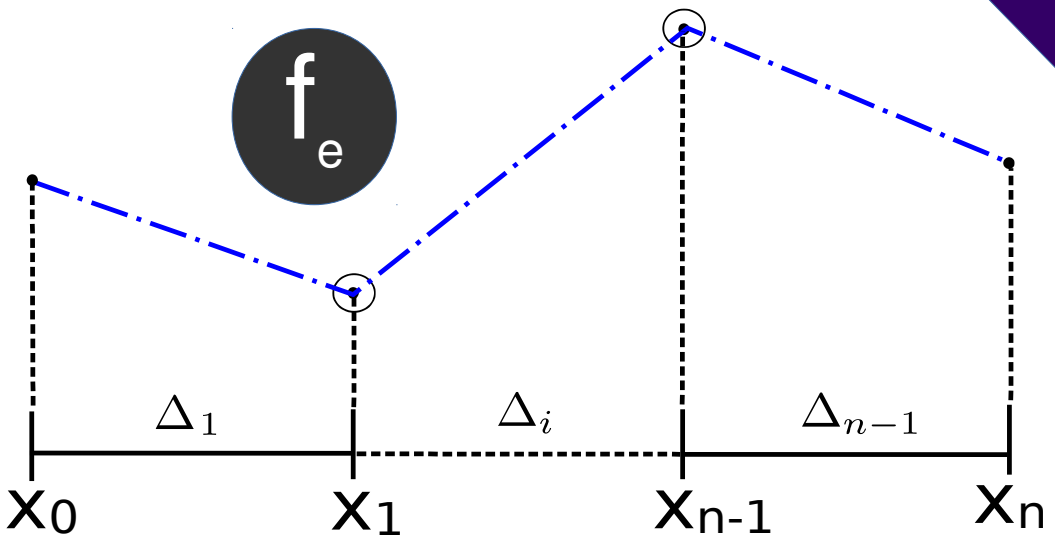


From ModSpec to table-based model

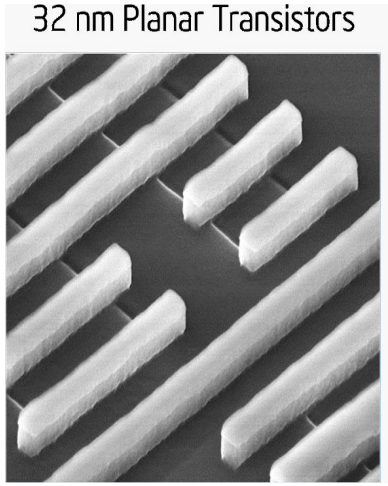


Machine Translation is openly available – VAPP

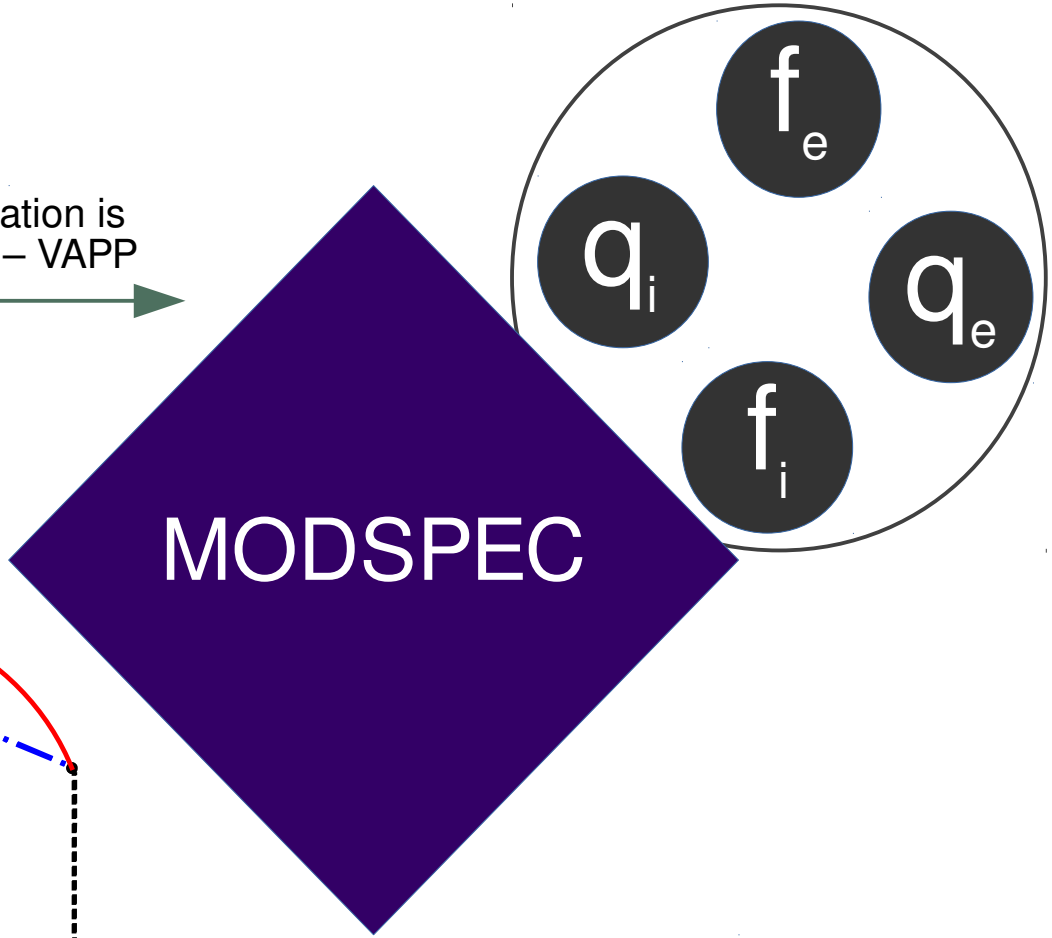


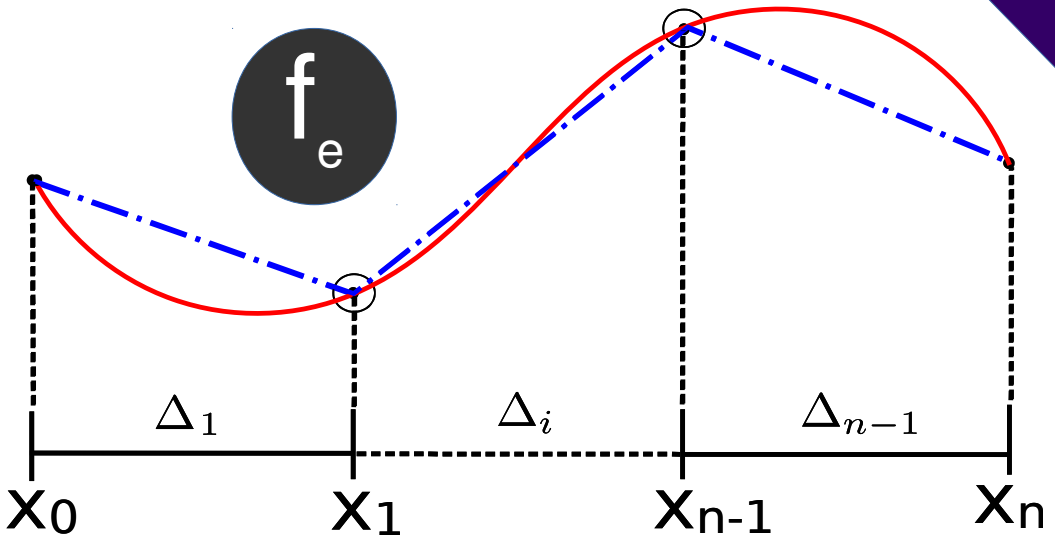


From ModSpec to table-based model

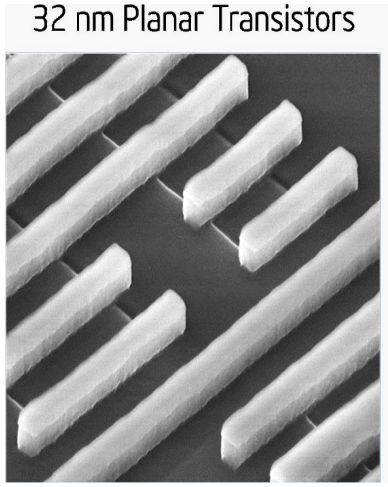


Machine Translation is openly available – VAPP

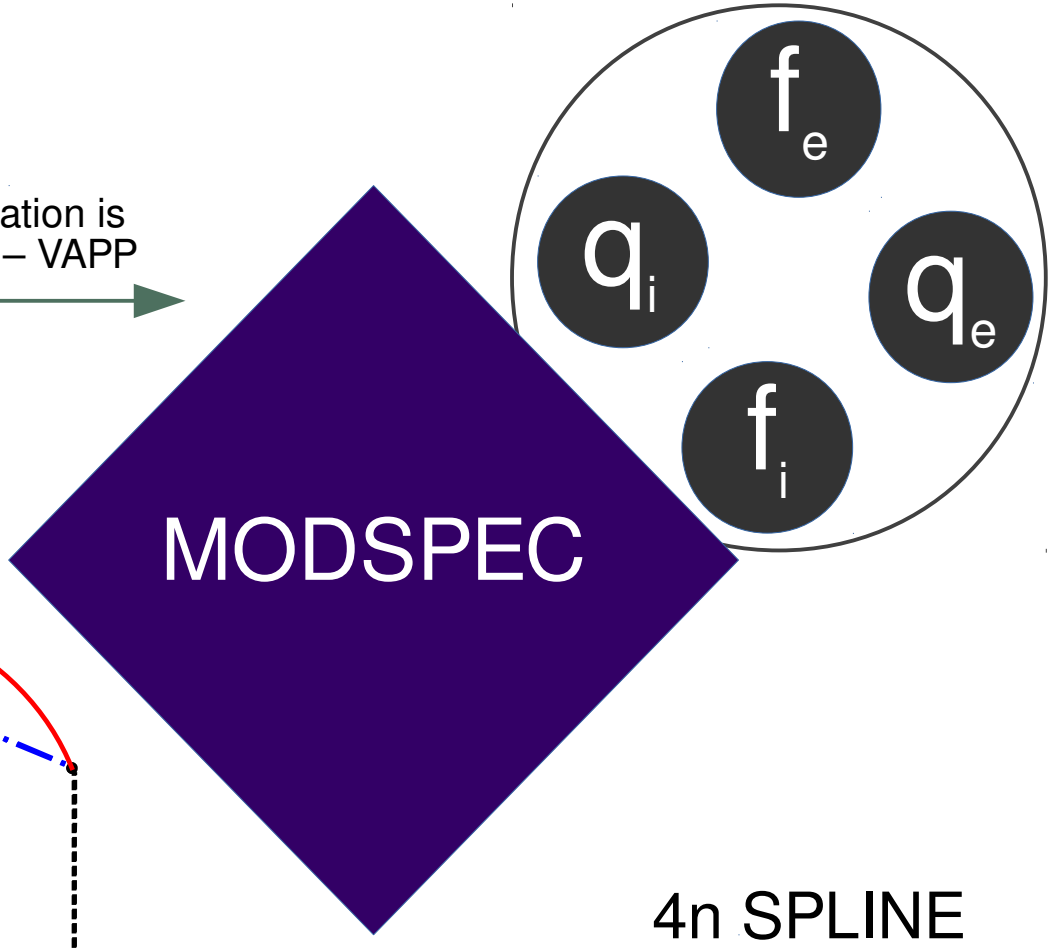


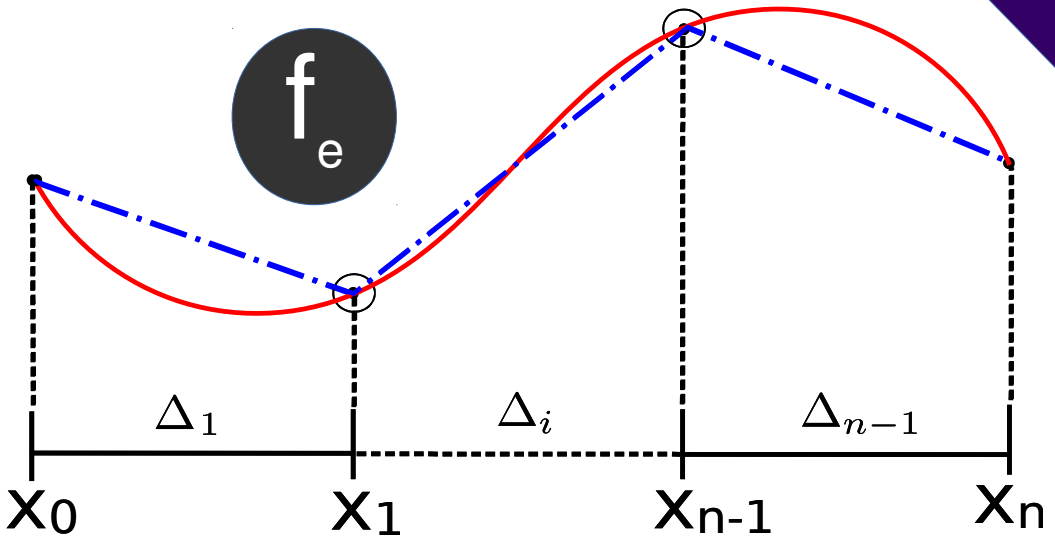


From ModSpec to table-based model



Machine Translation is openly available – VAPP





4n SPLINE COEFFICIENTS

Extrapolation!

Extrapolation!

- Solvers typically follow a method of prediction-correction.

Extrapolation!

- Solvers typically follow a method of prediction-correction.
  - Predictions can be wrong and UNREASONABLE

Extrapolation!

- Solvers typically follow a method of prediction-correction.
  - Predictions can be wrong and UNREASONABLE
- Extrapolation plays a very important role in this context

Extrapolation!

- Solvers typically follow a method of prediction-correction.
  - Predictions can be wrong and UNREASONABLE
- Extrapolation plays a very important role in this context
  - Determines whether or not the an unreasonable prediction will return to a correct region of operation

Extrapolation!

- Solvers typically follow a method of prediction-correction.
  - Predictions can be wrong and UNREASONABLE
- Extrapolation plays a very important role in this context
  - Determines whether or not the an unreasonable prediction will return to a correct region of operation
  - How long will this “return” take

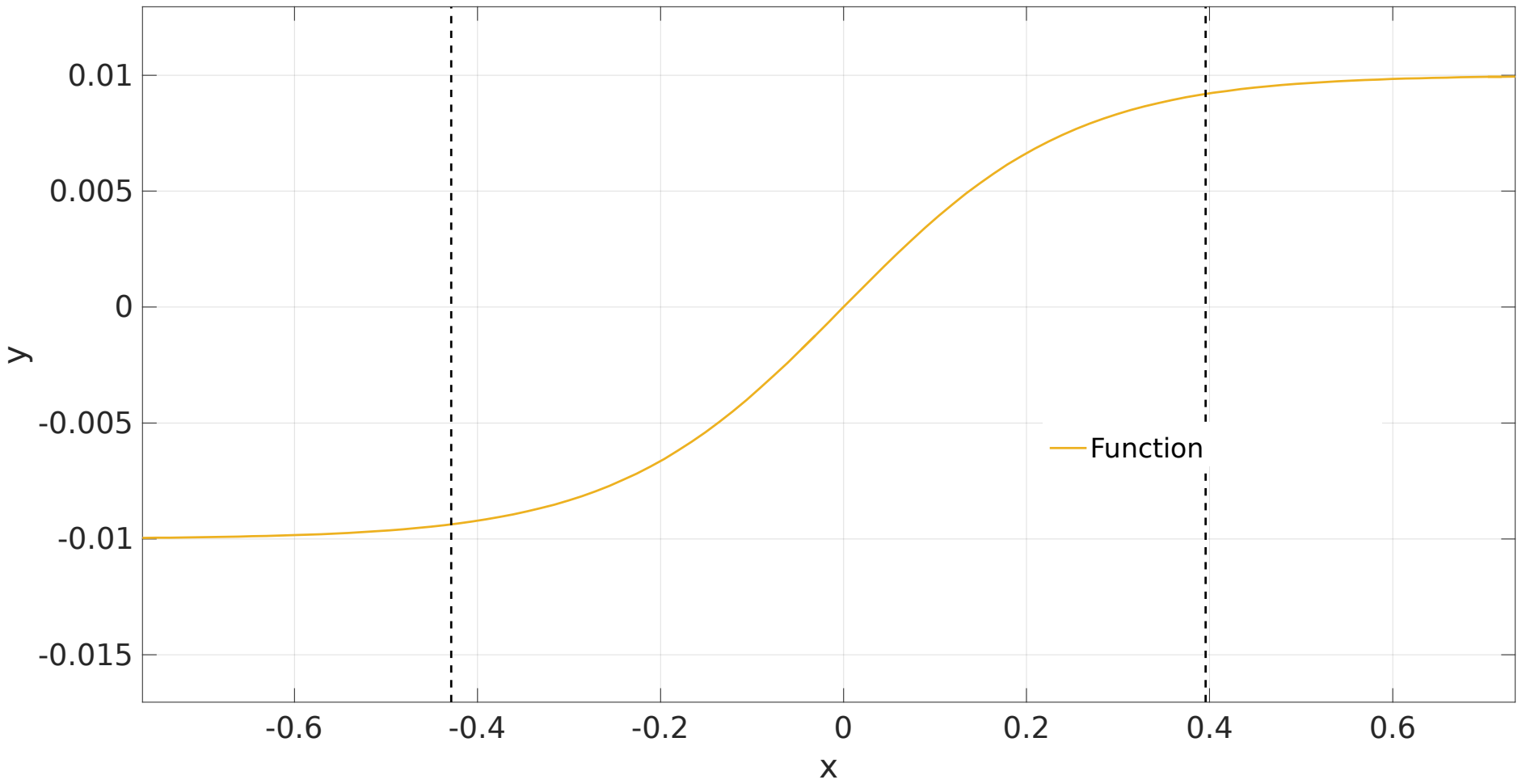
Extrapolation!

- Solvers typically follow a method of prediction-correction.
  - Predictions can be wrong and UNREASONABLE
- Extrapolation plays a very important role in this context
  - Determines whether or not the an unreasonable prediction will return to a correct region of operation
  - How long will this “return” take

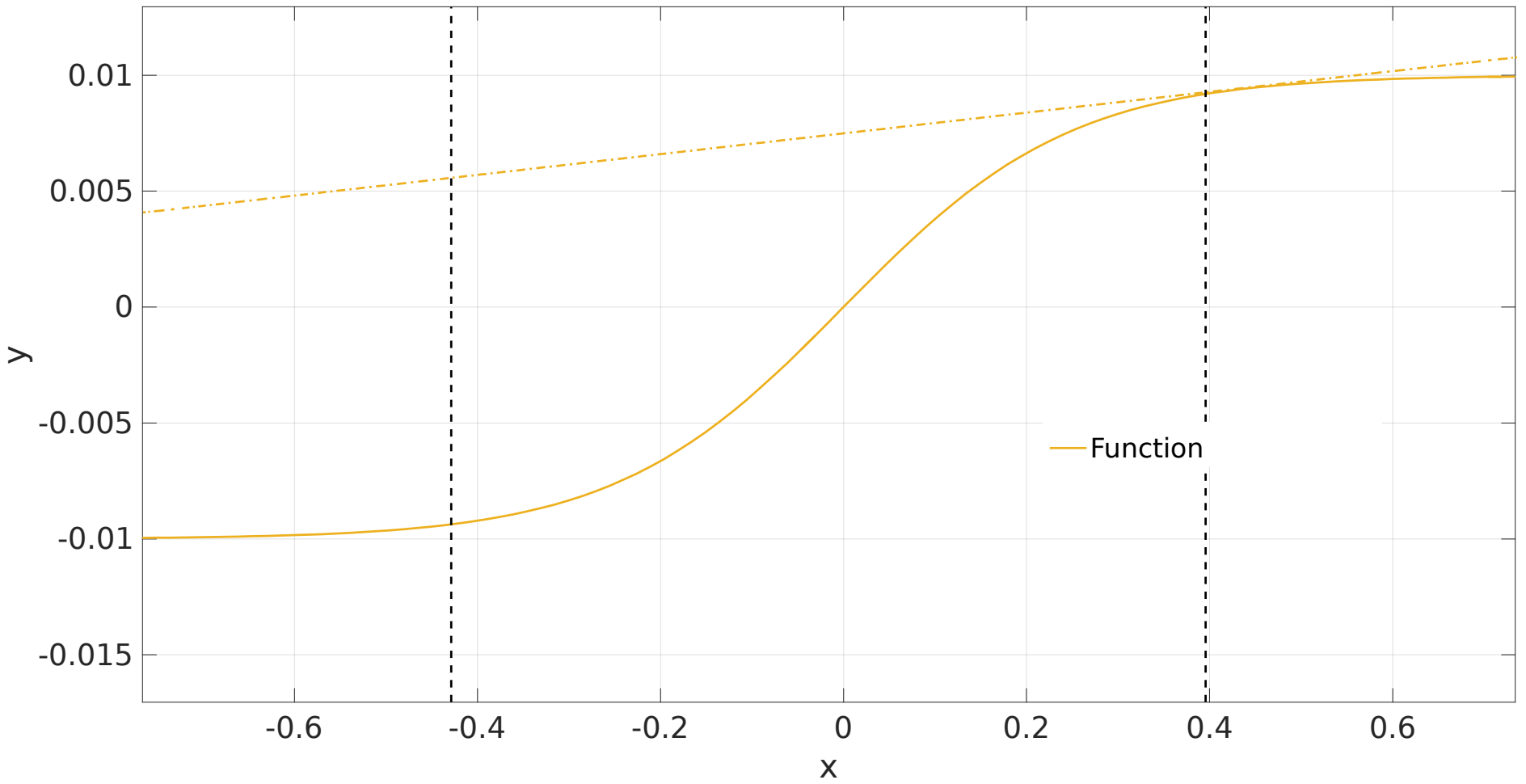
Linear extrapolation is an option

Extrapolation!

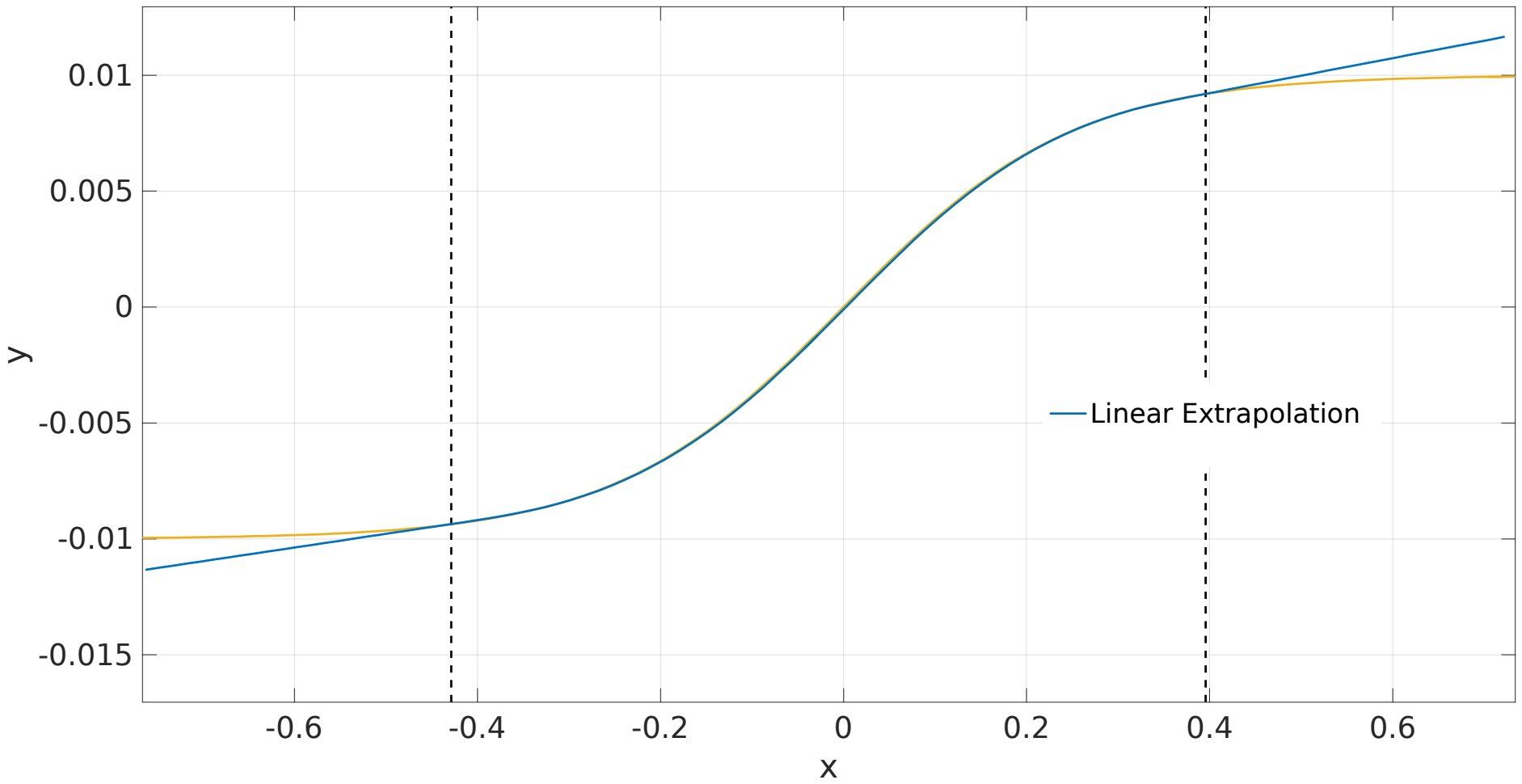
Extrapolation!



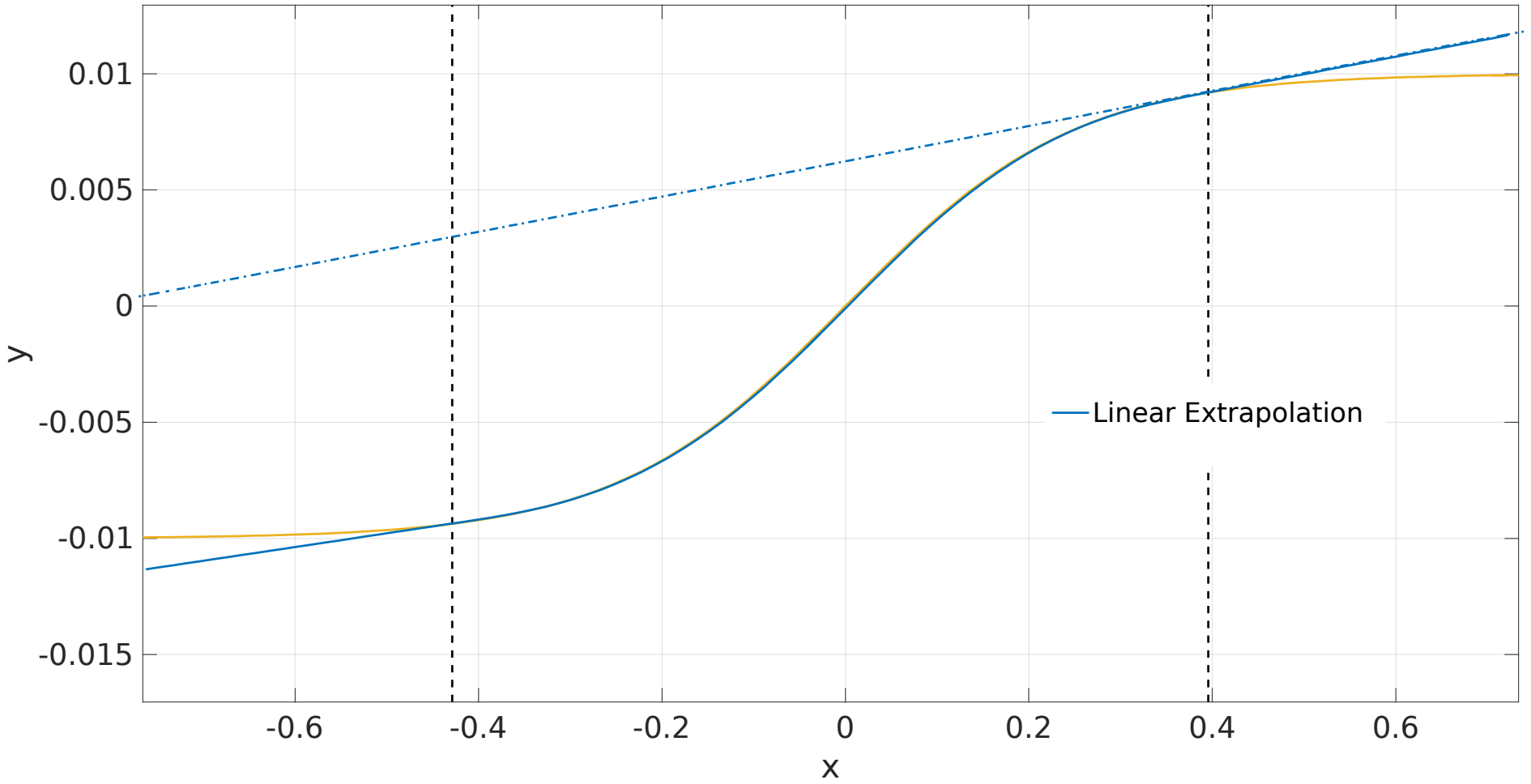
Extrapolation!



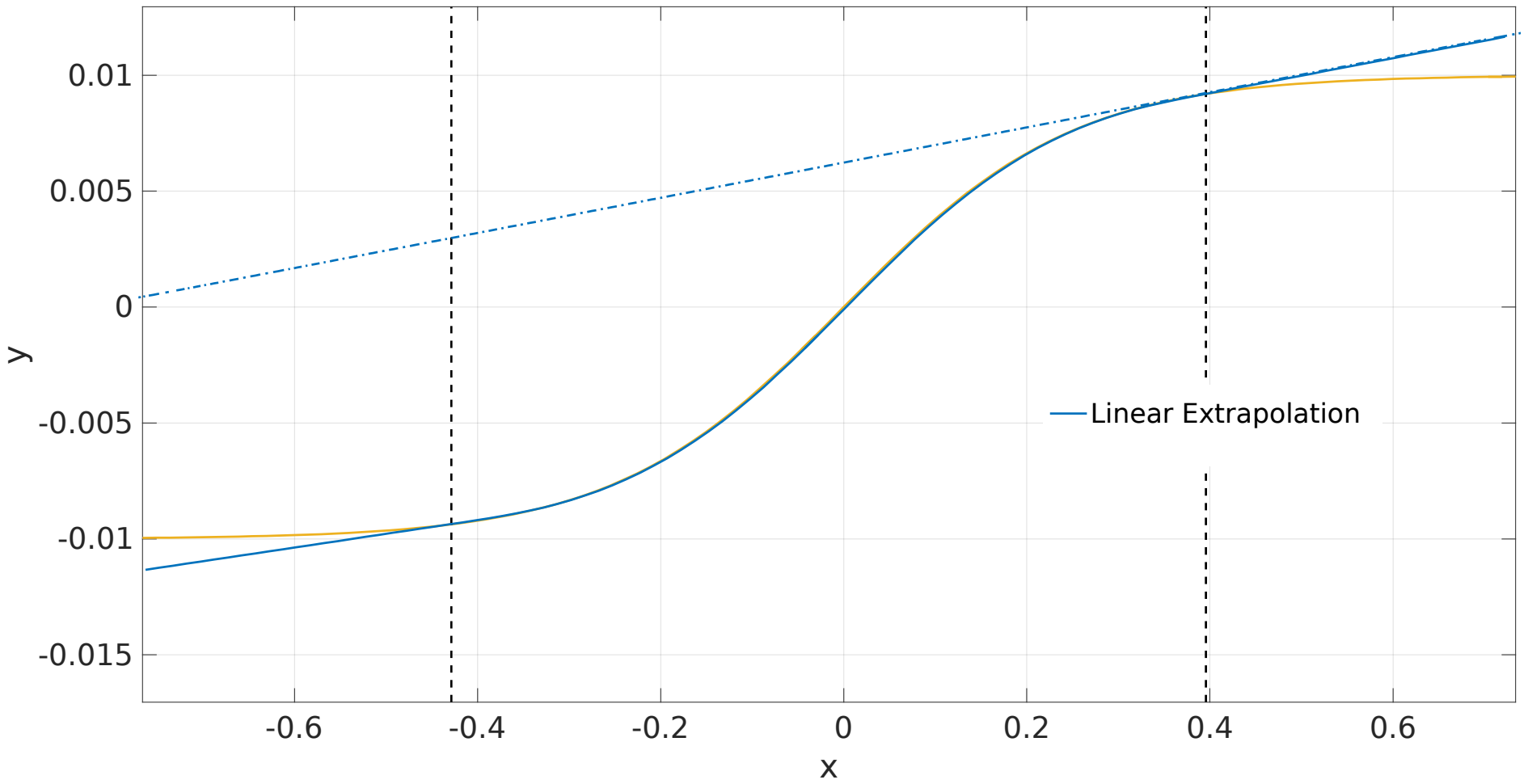
Extrapolation!



Extrapolation!

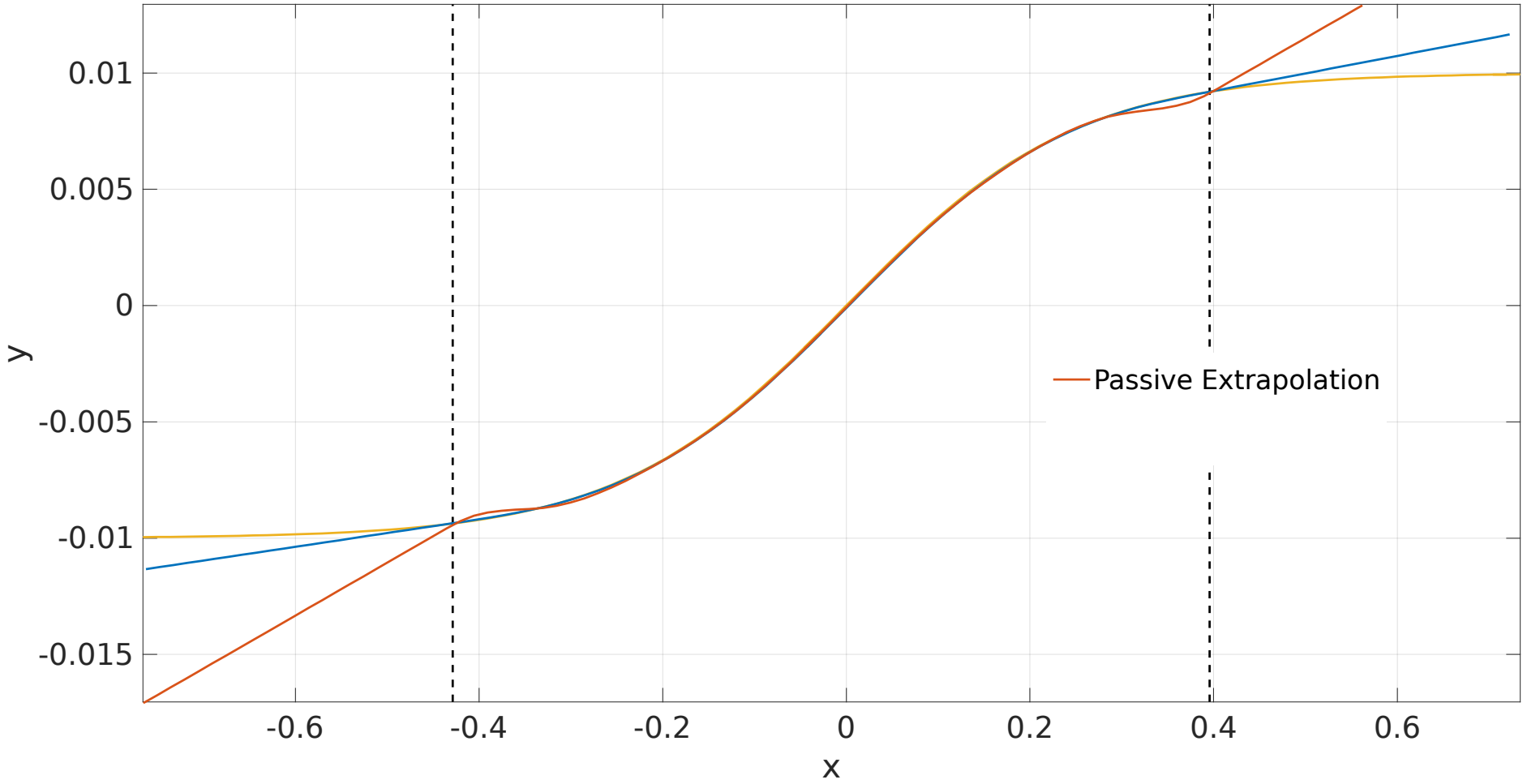


Extrapolation!



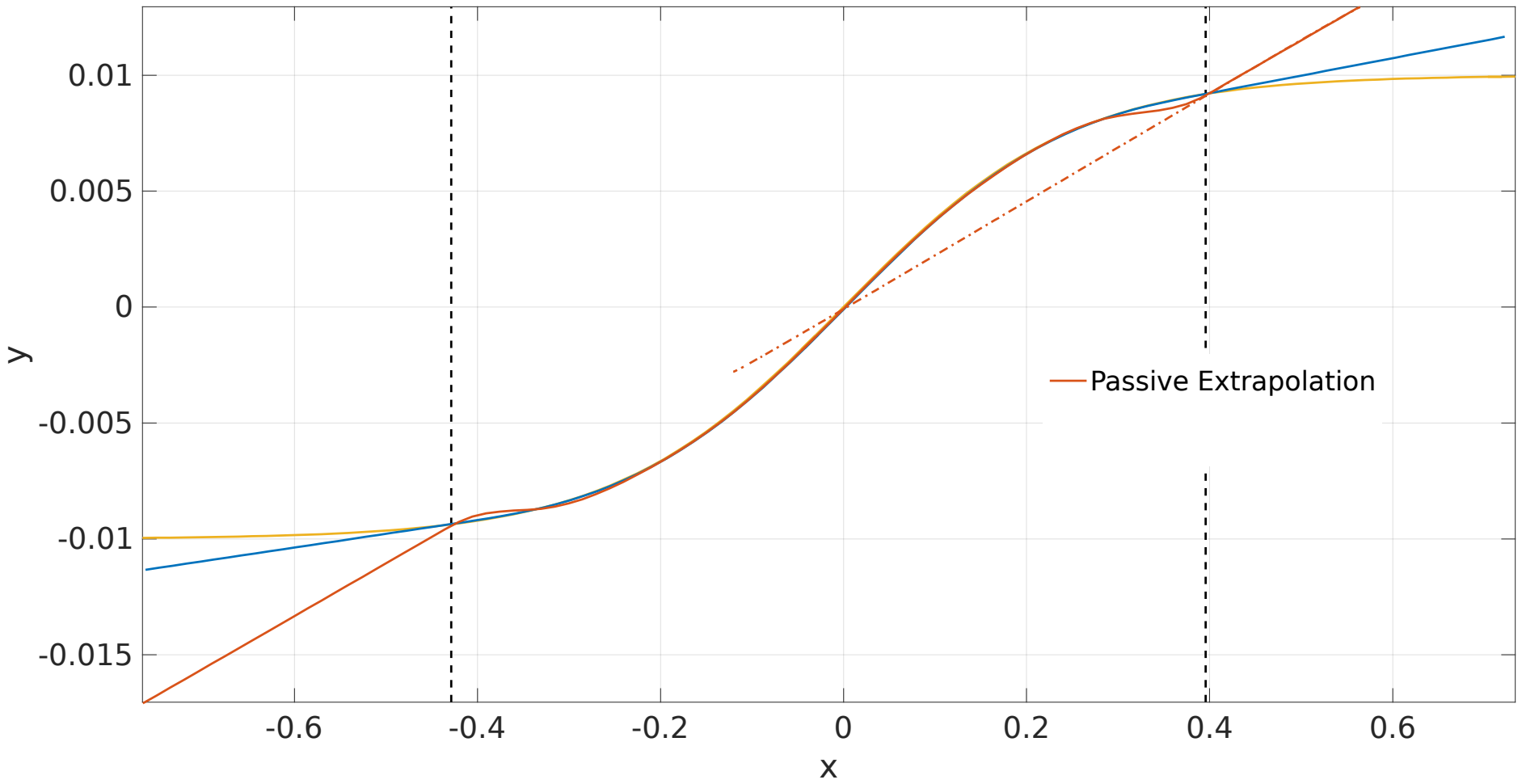
Linear extrapolation can be too restrictive!

Extrapolation!



Linear extrapolation can be too restrictive!

Extrapolation!



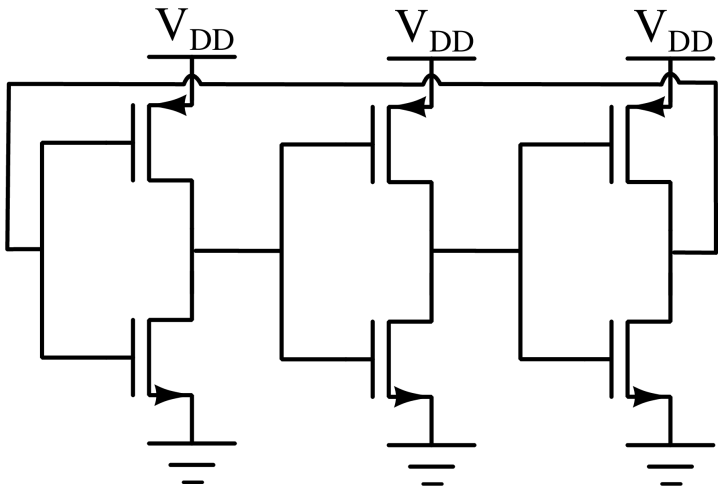
Linear extrapolation can be too restrictive!

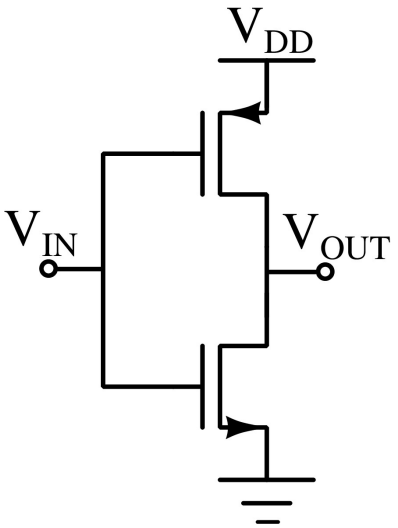
Speedup
Accuracy
Memory

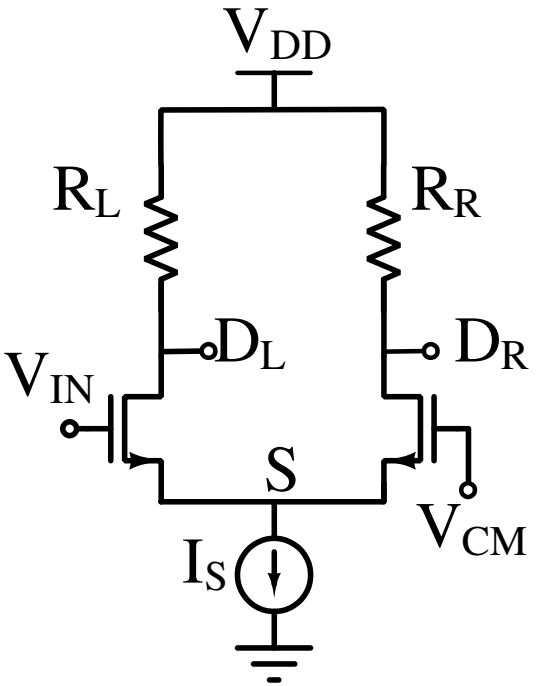
Simulation Results

Test Circuits and Waveforms

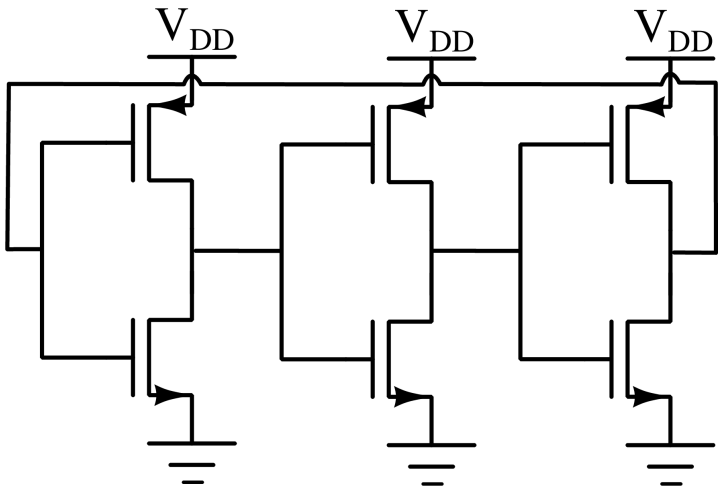
Test Circuits and Waveforms

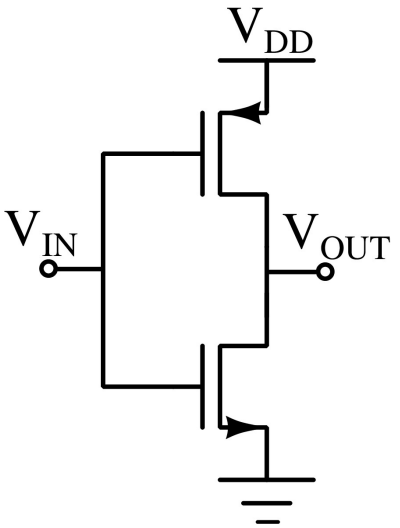


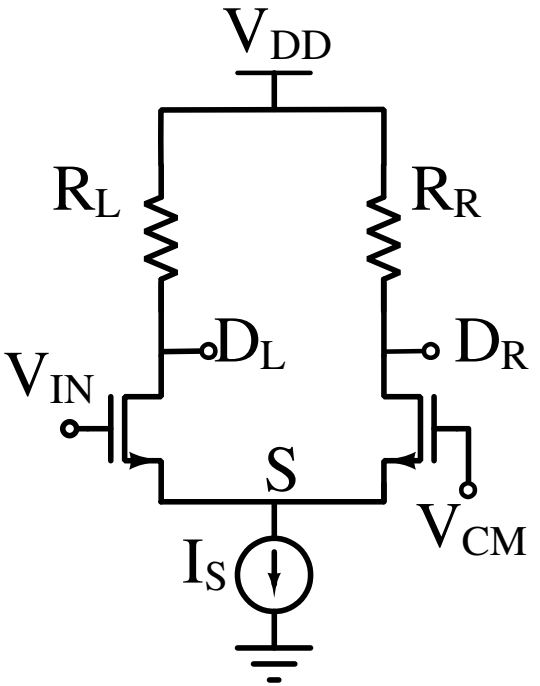


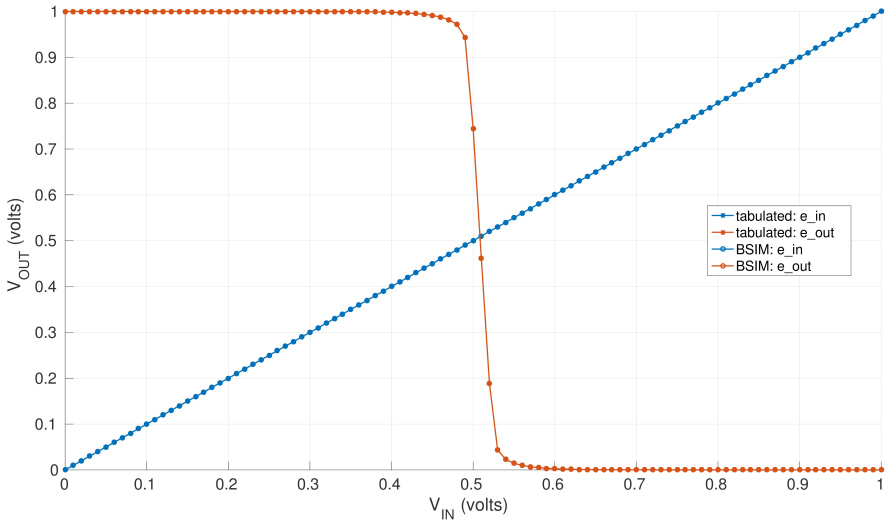


Test Circuits and Waveforms

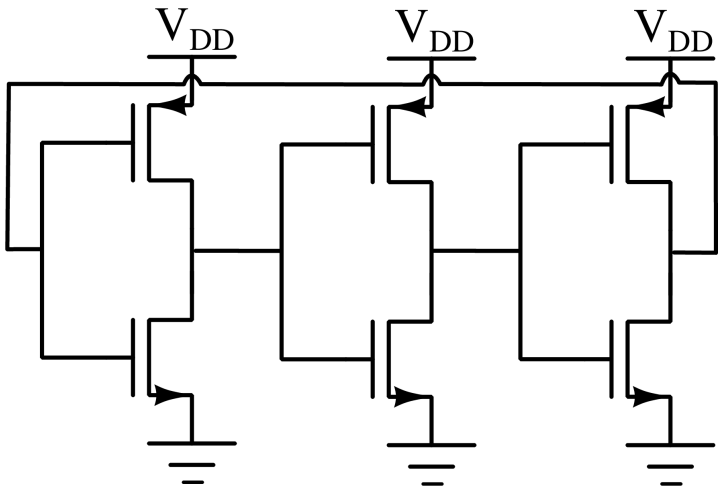


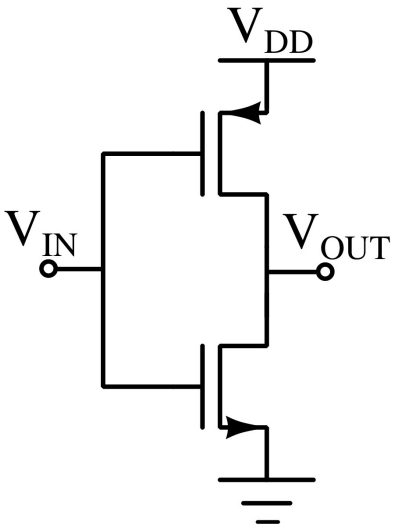


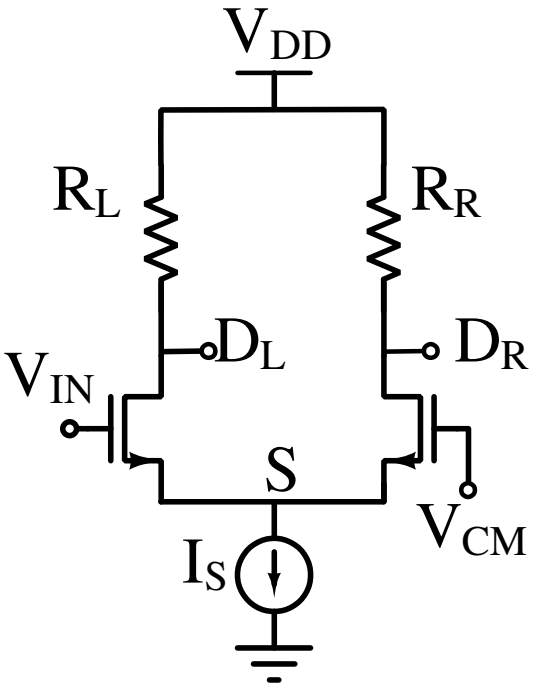


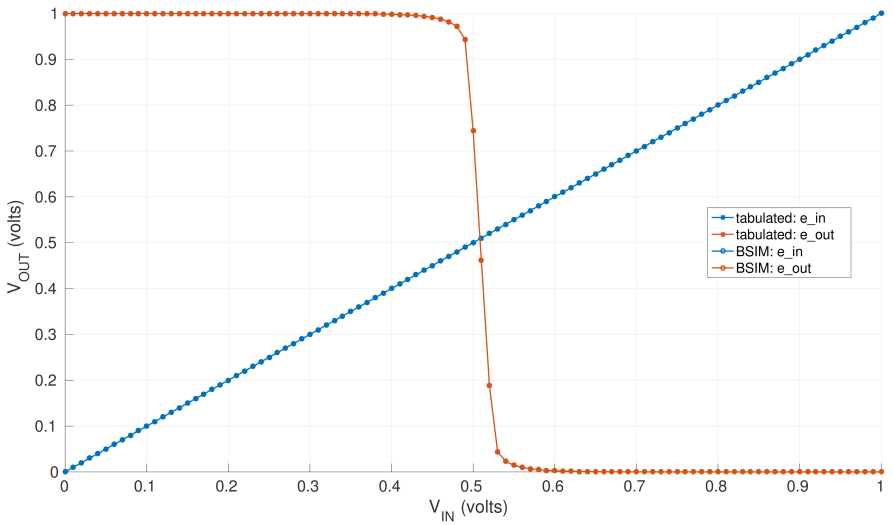


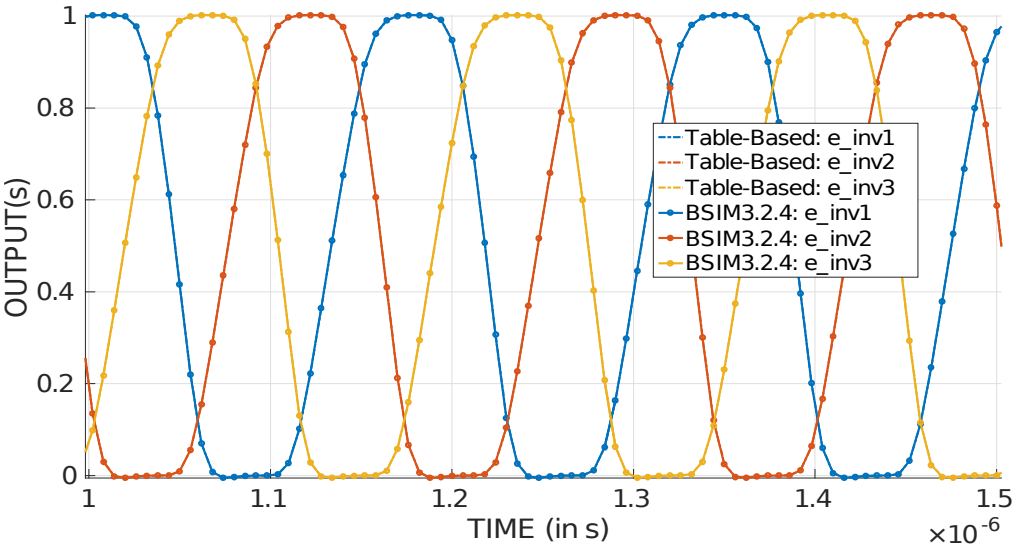
Test Circuits and Waveforms



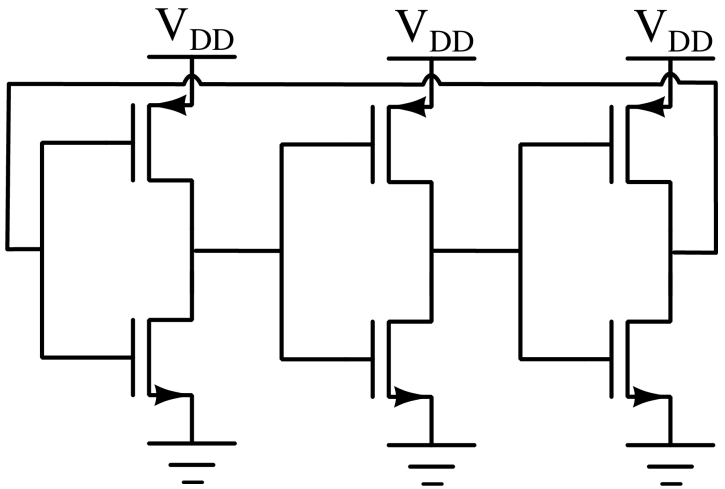


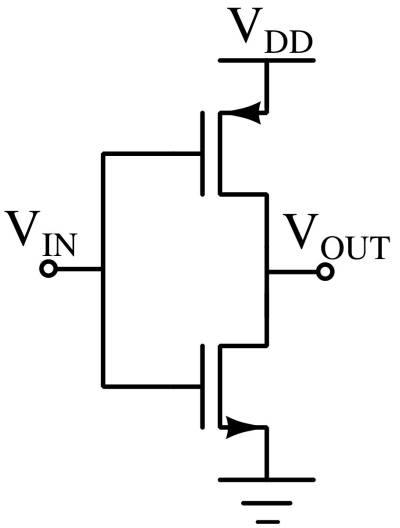


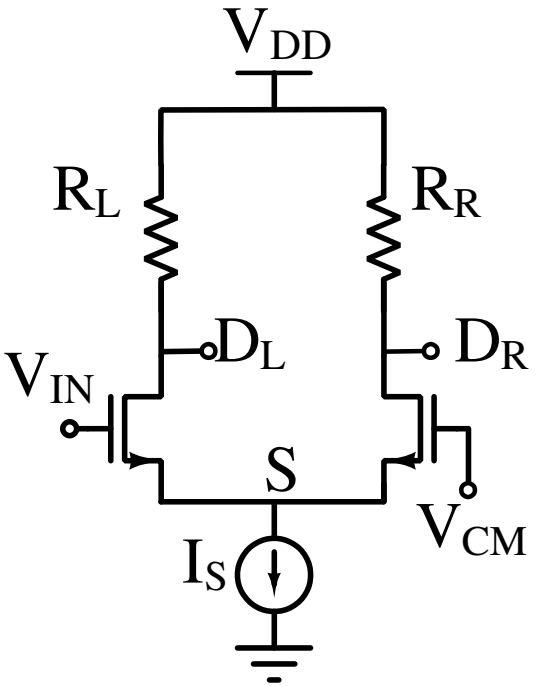


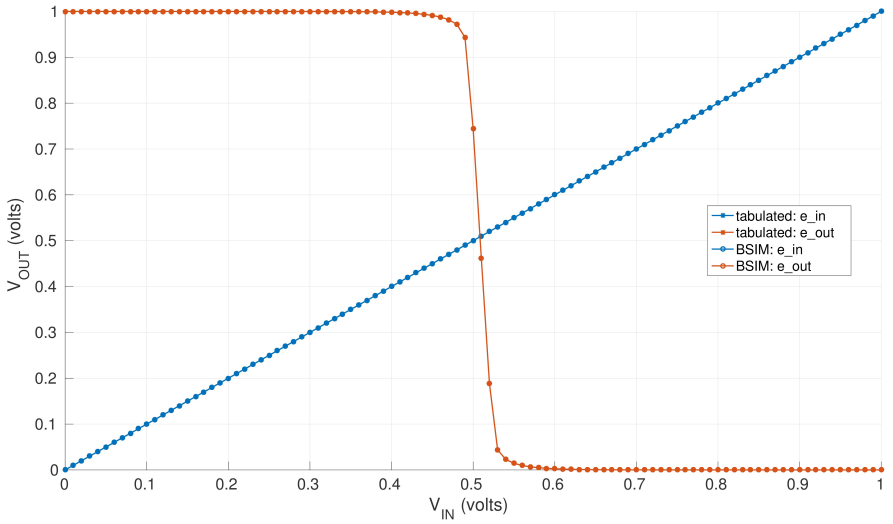


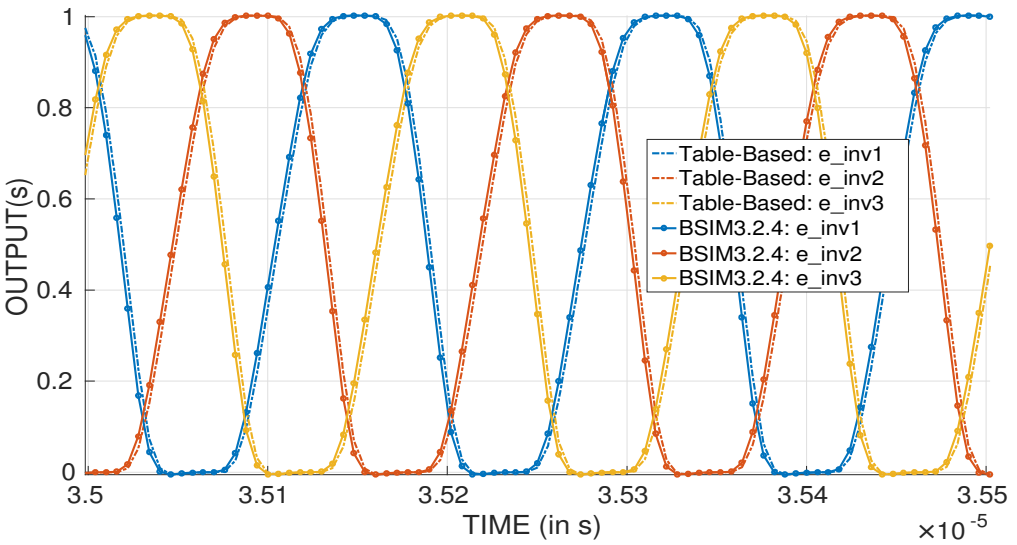
Test Circuits and Waveforms





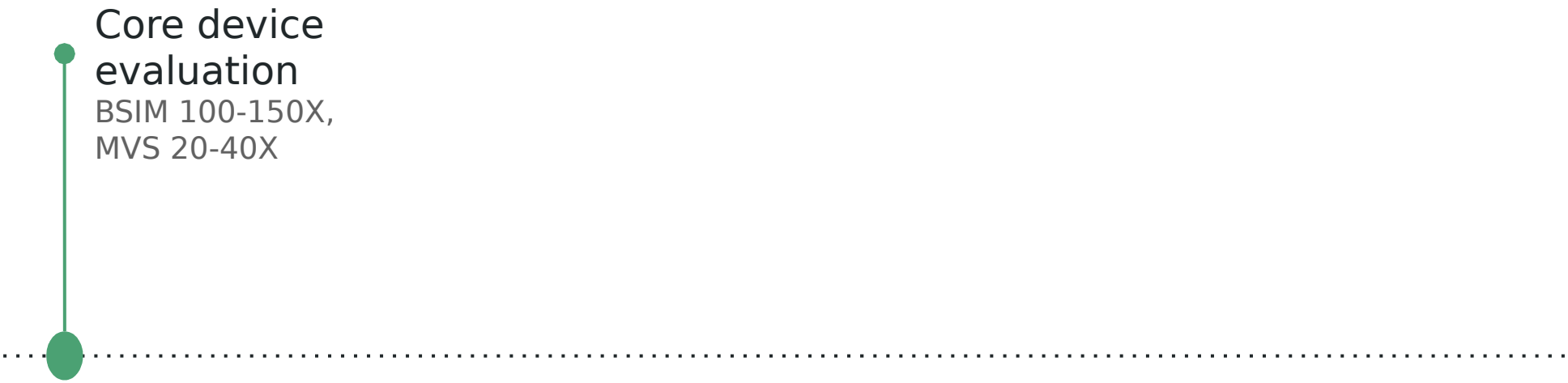






Speedup Evaluation

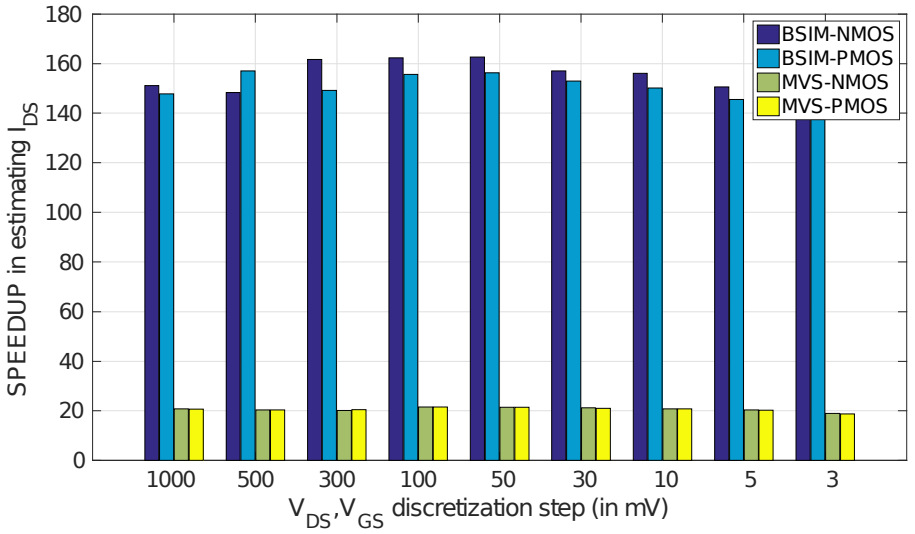
Speedup Evaluation



Speedup Evaluation

Core device evaluation

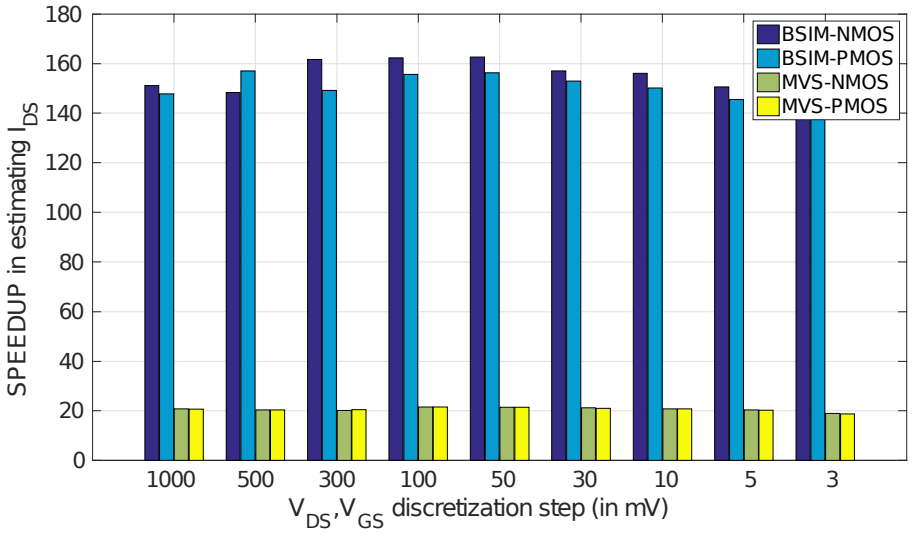
BSIM 100-150X, MVS 20-40X



Speedup Evaluation

Core device evaluation

BSIM 100-150X,
MVS 20-40X



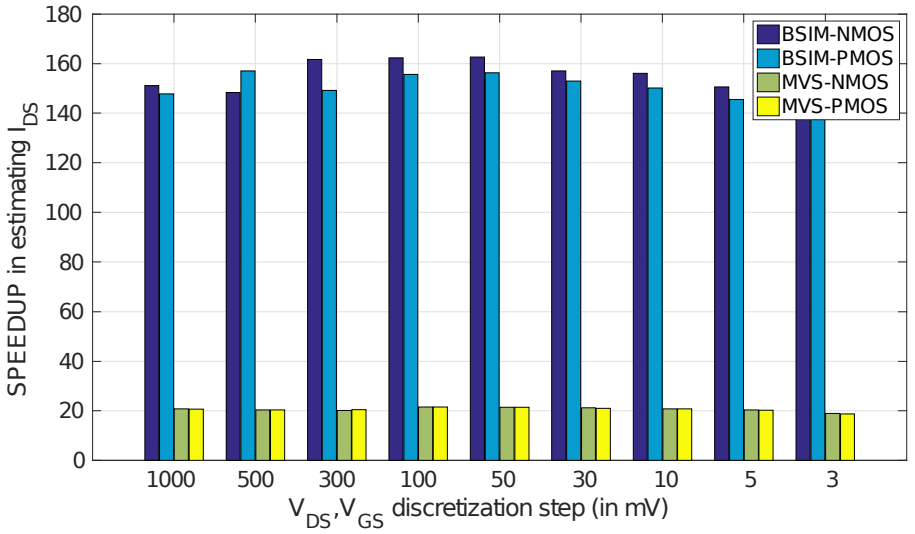
Analysis (BSIM)

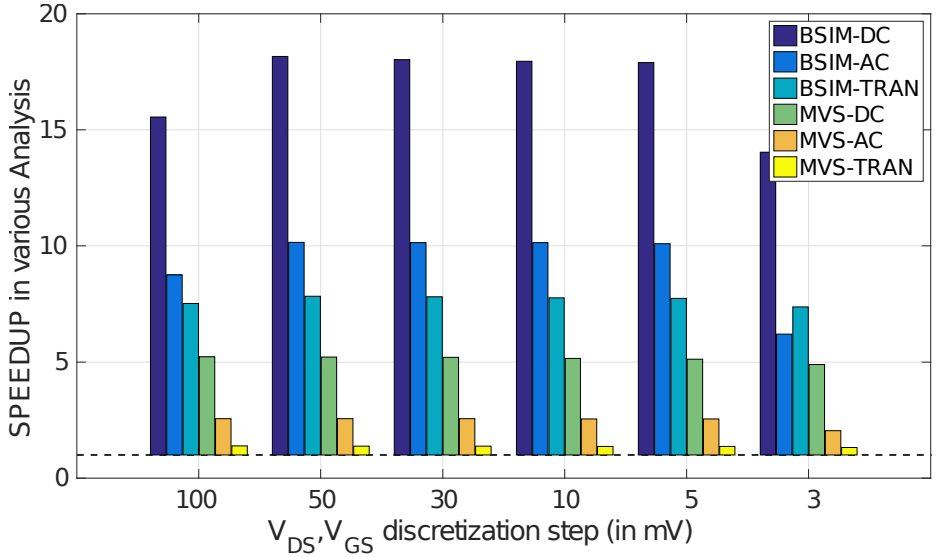
DC (QSS) 14-20X,
AC 7-10X,
TRANSIENT 6-8X

Speedup Evaluation

Core device evaluation

BSIM 100-150X,
MVS 20-40X





Analysis (BSIM)

DC (QSS) 14-20X,
AC 7-10X,
TRANSIENT 6-8X

Error Evaluation

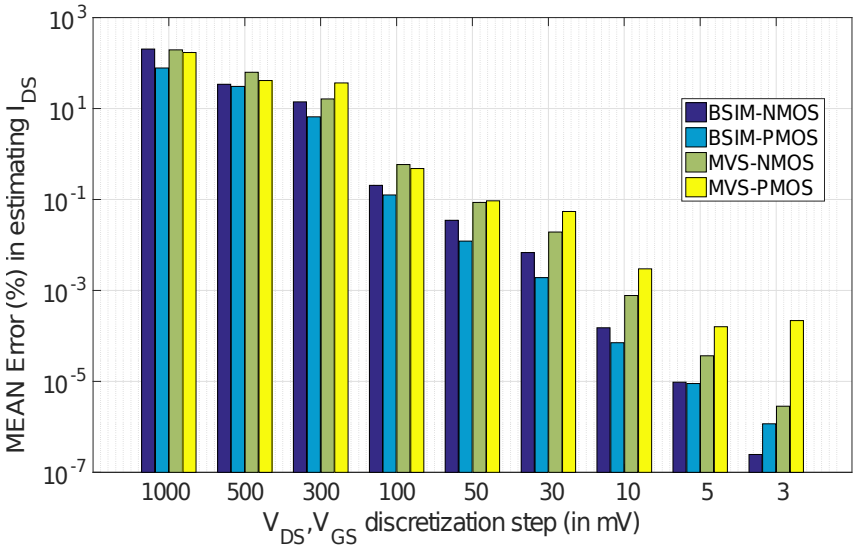
Error Evaluation

Error Evaluation

Core device evaluation

BSIM < 0.00001%

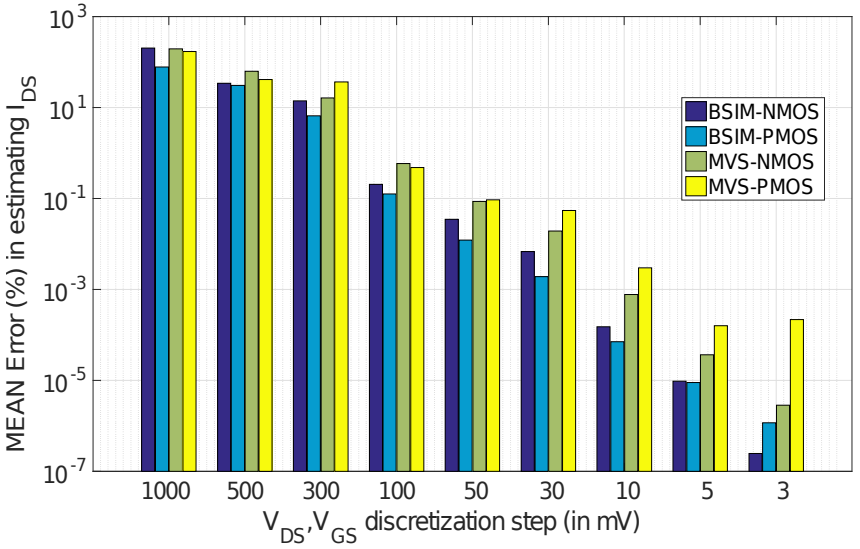
MVS < 0.001%



Error Evaluation

Core device evaluation

BSIM < 0.00001%
MVS < 0.001%



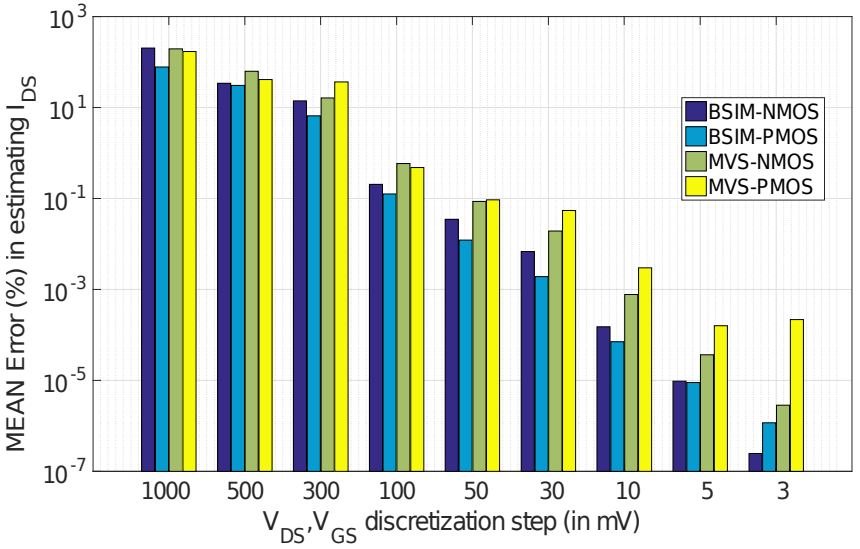
Analysis

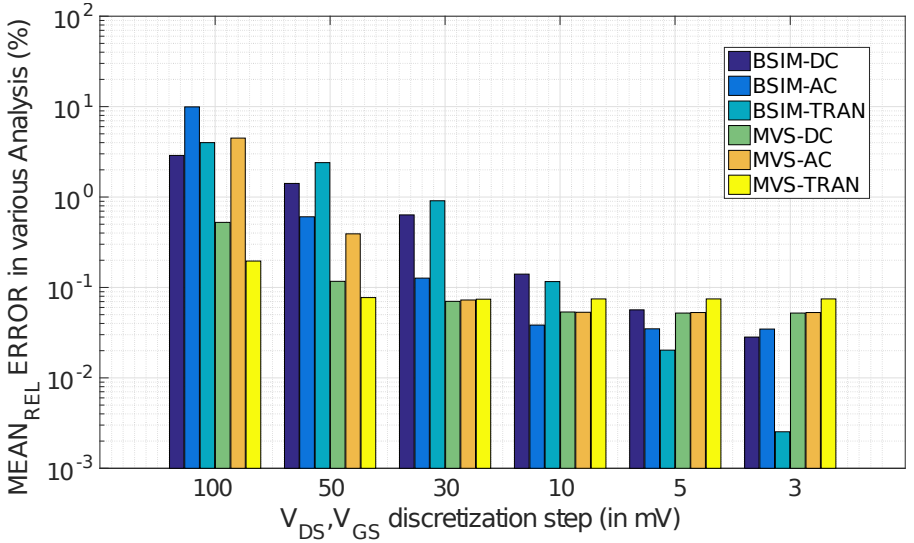
DC < 0.1%,
AC < 0.1%,
TRANSIENT < 0.01%

Error Evaluation

Core device evaluation

BSIM < 0.00001%
MVS < 0.001%



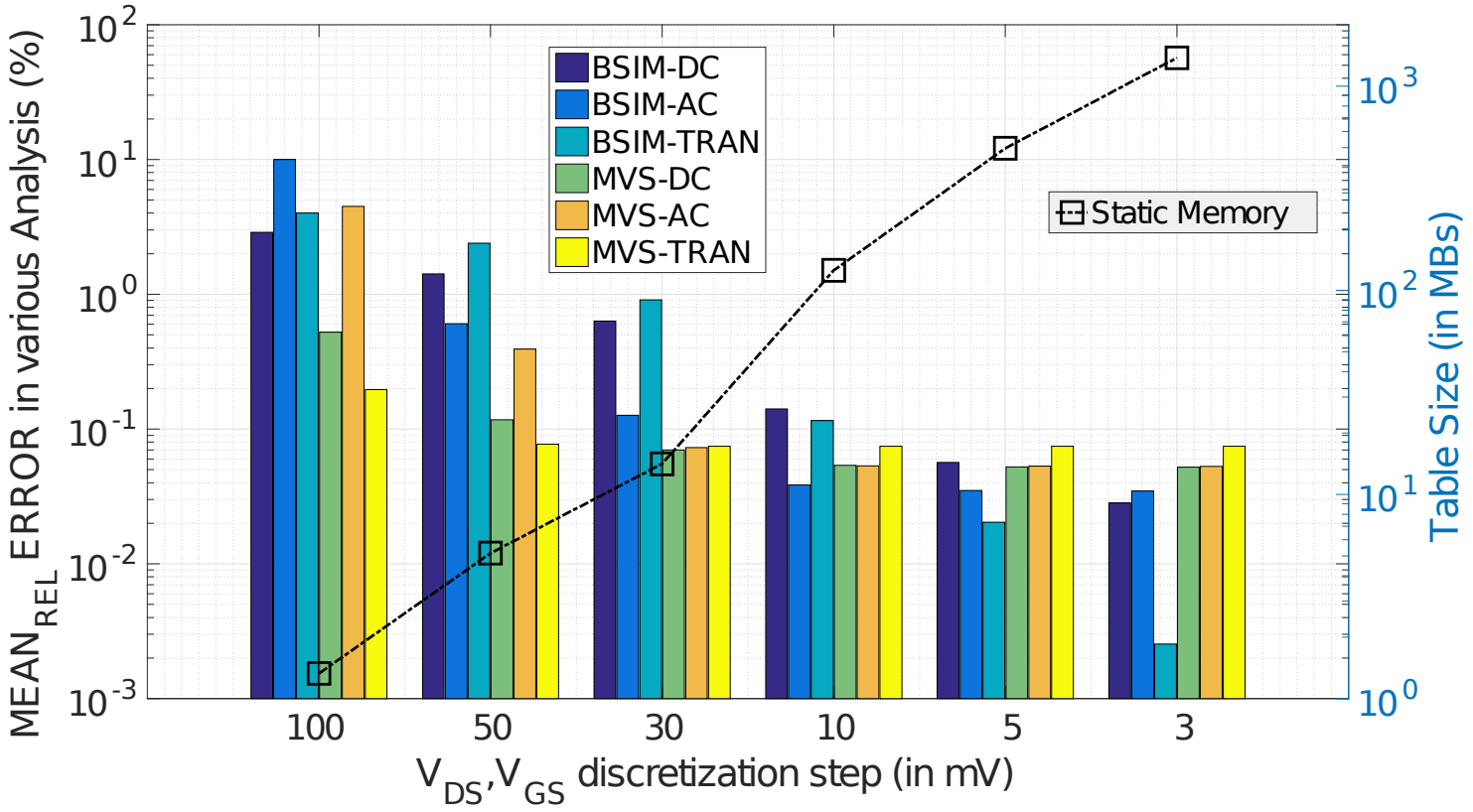


Analysis

DC < 0.1%,
AC < 0.1%,
TRANSIENT < 0.01%

The cost is Memory and table evaluation time

The cost is Memory and table evaluation time



Spline evaluation

Table sizes and computation times

Future Work and problems

Future Work and problems

- Memory requirements *‘can’* be lowered:

Future Work and problems

- Memory requirements *‘can’* be lowered:
  - The plots for reconstruction error seem to be saturating

Future Work and problems

- Memory requirements *‘can’* be lowered:
  - The plots for reconstruction error seem to be saturating
  - Splines are probably not the best interpolants, but work as a proof-of-concept

Future Work and problems

- Memory requirements *‘can’* be lowered:
  - The plots for reconstruction error seem to be saturating
  - Splines are probably not the best interpolants, but work as a proof-of-concept
- Exploiting structure in the polynomial coefficients obtained from table-based modelling

Future Work and problems

- Memory requirements *‘can’* be lowered:
  - The plots for reconstruction error seem to be saturating
  - Splines are probably not the best interpolants, but work as a proof-of-concept
- Exploiting structure in the polynomial coefficients obtained from table-based modelling
- Simulating circuits with measurement data!

QUESTIONS