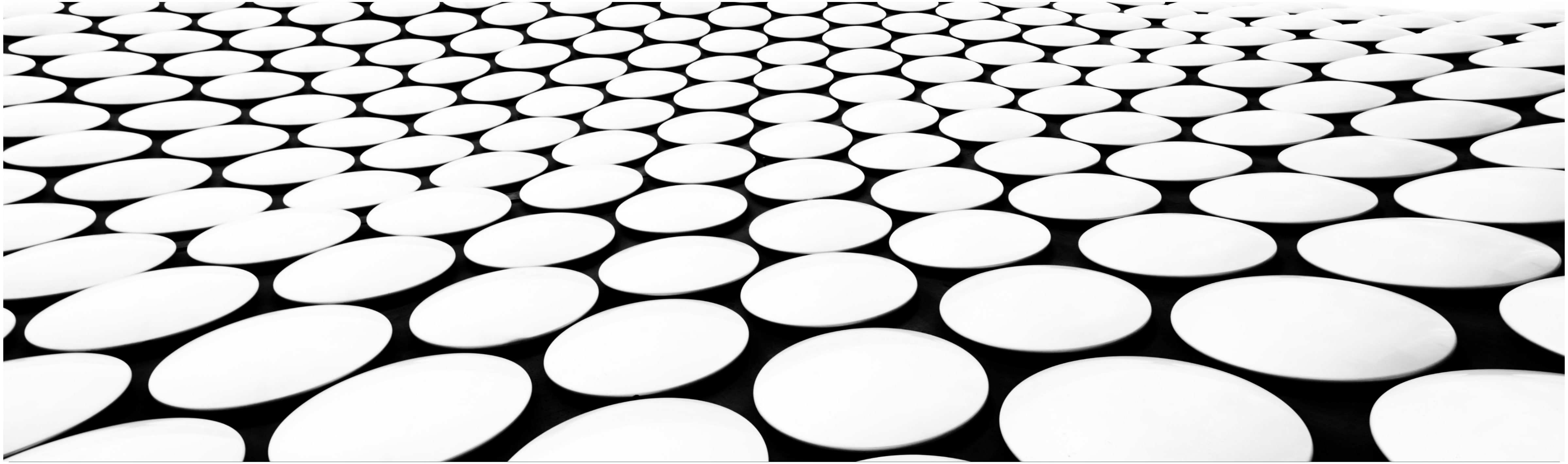
Statistical Analysis of MOSFET Extracted Parameters for n-MOS Mismatch Modeling

JUAN PABLO MARTINEZ BRITO
SERGIO BAMPI
CEITEC S.A. / UFRGS
BRAZIL

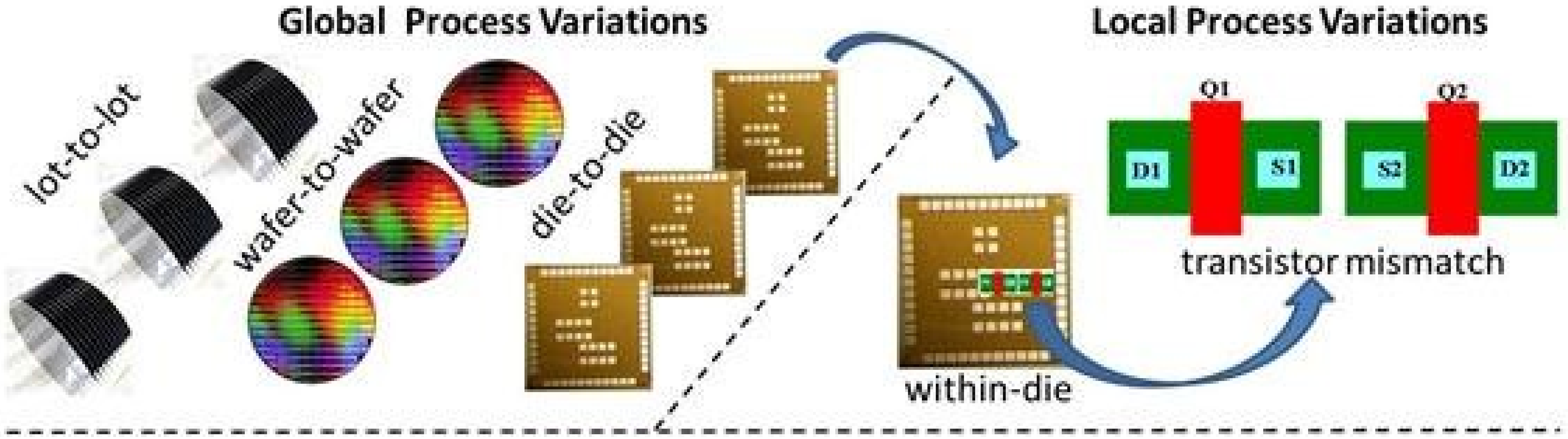
There are more things in heaven and earth, Horatio,
Than are dreamt of in your philosophy.
- Hamlet (1.5.167-8), Hamlet to Horatio



OUTLINE

1. MOSFET Mismatch Modeling Intro
2. Test Structure Approach
3. Variability Data Visualization
4. Statistical Transistor Parameter Distribution
5. Effective Mismatch Area
6. Conclusion

VARIABILITY CLASSIFICATION

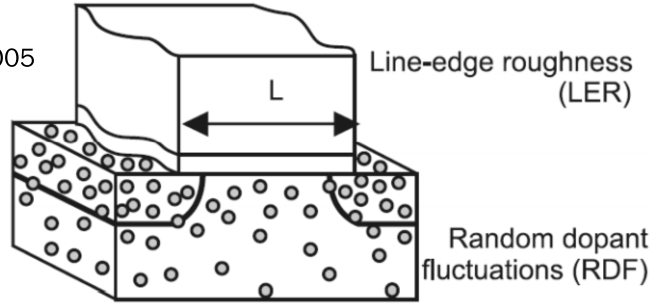


- ➢ Systematic behavior.
- ➢ Independent on device size.
- ➢ Somehow controllable.

- ➢ Random in nature.
- ➢ Dependent on device size.
- ➢ Somehow uncontrollable.

Primary sources of local variation

ROY; ASENOV, 2005



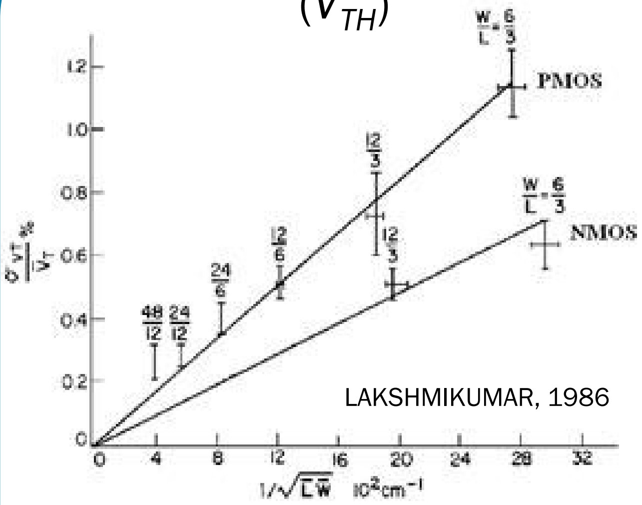
PELGROM MOSFET EMPIRICAL MISMATCH MODEL

Power content of fluctuations of parameter (Δ*P*) in the Fourier domain, which in turn can be interpreted as the variance of Δ*P*

$$\sigma^2(\Delta P) = \frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |H(\omega_x,\omega_y)|^2 S_P(\omega_x,\omega_y)d\omega_x d\omega_y$$

MOSFET Parameters that dominates MOS mismatches:

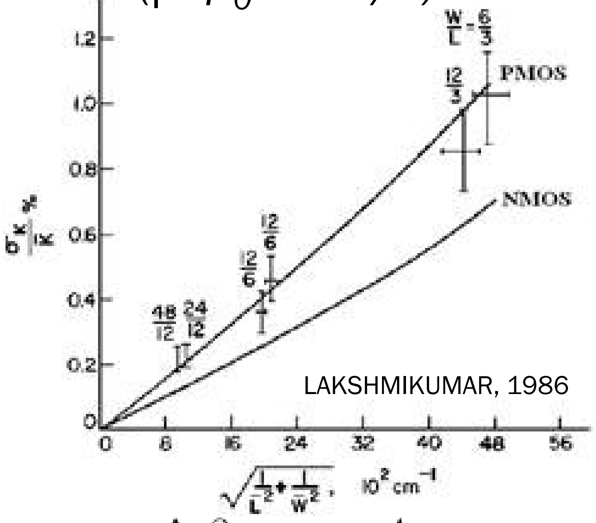
Threshold Voltage (V_{TH})



LAKSHMIKUMAR, 1986

$$\sigma(\Delta V_{TH}) = \frac{A_{V_{TH}}}{\sqrt{WL}}$$

Current Factor ($\beta = \mu_0 Cox.W/L$)



LAKSHMIKUMAR, 1986

$$\sigma(\frac{\Delta\beta}{\beta}) = \frac{A_\beta}{\sqrt{WL}}$$

PELGROM, 1989

$$\sigma^2(\Delta P) = \frac{{A_P}^2}{WL} + {S_P}^2 D^2$$

size dependent

distance dependent

KINGET; STEYAERT, 2005

$$\sigma^2\left(\frac{\Delta I_{DS}}{I_{DS}}\right) = \sigma^2\left(\frac{\Delta\beta}{\beta}\right) + \left(\frac{gm}{I_D}\right)^2 \cdot \sigma^2(\Delta V_{TH})$$

General MOSFET current mismatch model for all regions of operation

ACM (ADVANCED COMPACT MODEL) MISMATCH TRANSISTOR MODEL

ACM is an inversion charge-based MOS transistor model based on the inversion charge densities at the two ends (Source and Drain) of the channel.

ANA, CARLOS, MARCIO, 1998

Linear dependence between the inversion charge density (Q'_{IS}) and surface potential (ϕ_S) which is valid for all operating regions of the transistor (weak, moderate, strong).

$$I_D = I_F - I_R = \mu\eta C'_{ox}\frac{W}{L}\frac{\phi_t^2}{2}\left[\left(\frac{Q'_{IS(D)}}{\eta C'_{ox}\phi_t}\right)^2 - \frac{2Q'_{IS(D)}}{\eta C'_{ox}\phi_t}\right]$$

GALUP-MONTORO et al., 2005

$$N_{oi} = \int_0^{yd} N_a\left(1-\frac{y}{y_d}\right)^2 dy$$

- One of the first mismatch model to consider the influence of the vertical profile of dopants fluctuations in the depletion region by the *Noi* factor.

$$\sigma^2\left(\frac{\Delta I_{DS}}{I_{DS}}\right) = \frac{q^2 N_{oi}\mu}{L^2\eta C'_{ox} I_D}\ln\left(\frac{\eta C'_{ox}\phi_t - Q'_{IS}}{\eta C'_{ox}\phi_t - Q'_{ID}}\right)$$

- Integral made over the charge density of the inversion channel.

$$\sigma^2\left(\frac{\Delta I_{DS}}{I_{DS}}\right) = \frac{1}{WL}\left[\frac{N_{oi}}{N^{*2}}\frac{1}{(i_f - i_r)}\ln\left(\frac{1+i_f}{1+i_r}\right) + B_{ISQ}^2\right]$$

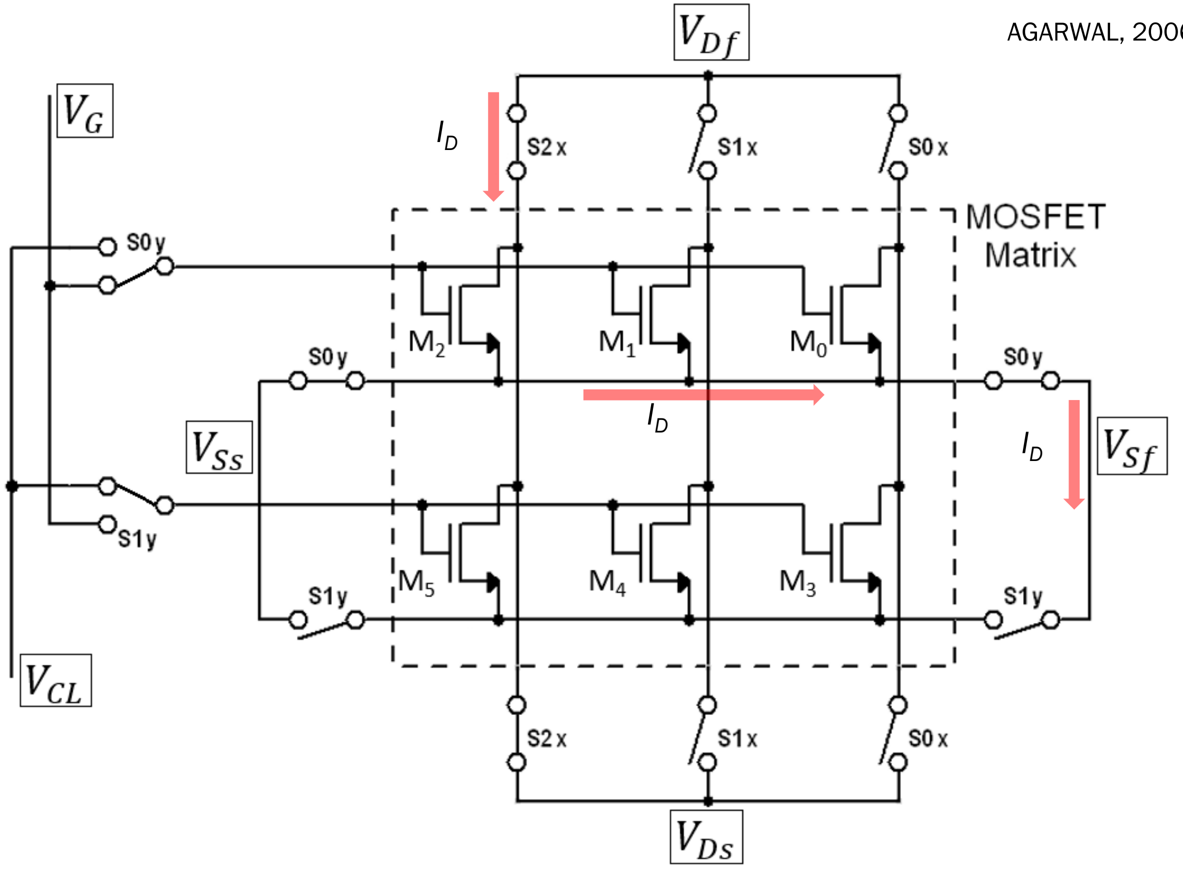
- Mismatch model expressed in terms of the normalized forward and reverse currents.

MOSFET 2D ADDRESSABLE MATRIX

FEATURES:

- X and Y Address Decoders.
- The (switches) transmission gates are thick oxide MOSFETs (IO MOSFETs) → low leakage.
- Force-Sense lines for Drain and Source terminals → (Keep V_{DS} constant)
- Negative voltage (NMOS case) applied to the gate → unselect the devices.

AGARWAL, 2006



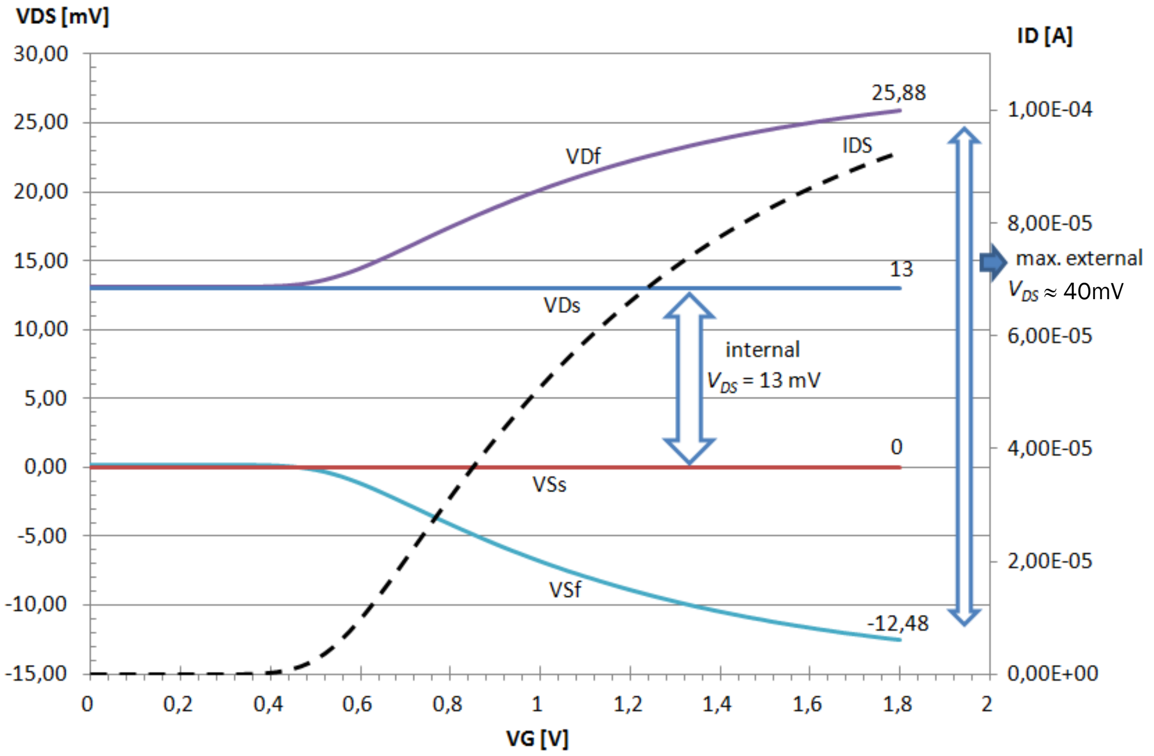
Example of a 3 column x 2 row MOSFET matrix.

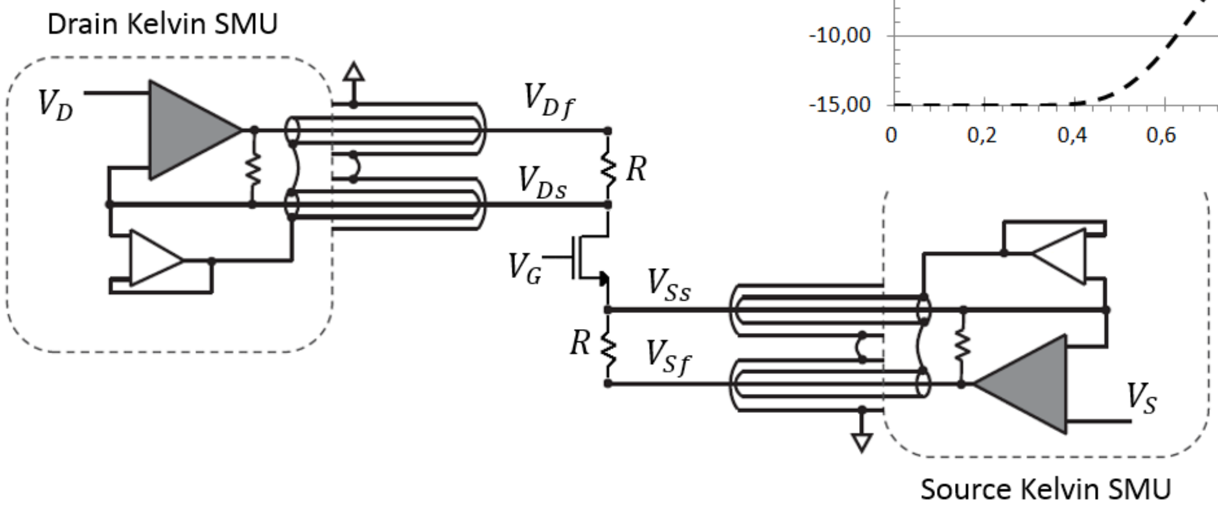
MOSFET MATRIX KELVIN APPROACH EFFICIENCY

To guarantee that the V_{DS} voltage inside the structure is exactly 13mV, the **Force-Sense** terminals must change accordingly.

→ V_{Df} goes more positive than V_{Ds}

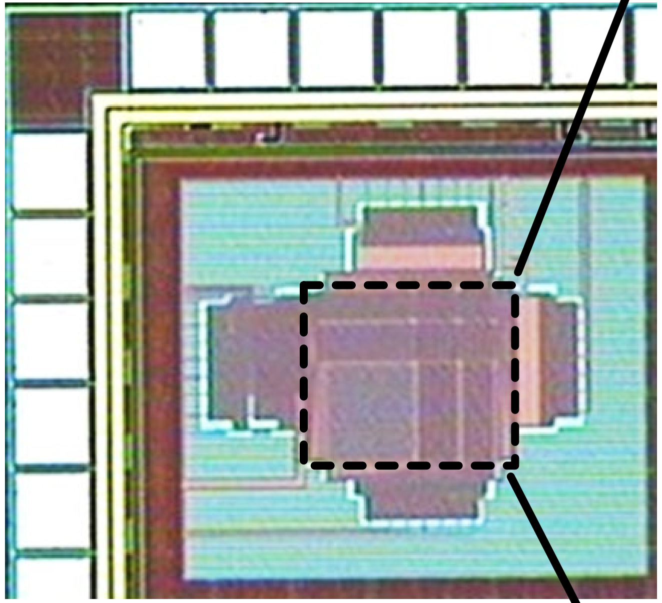
→ V_{Sf} goes more negative than V_{Ss}



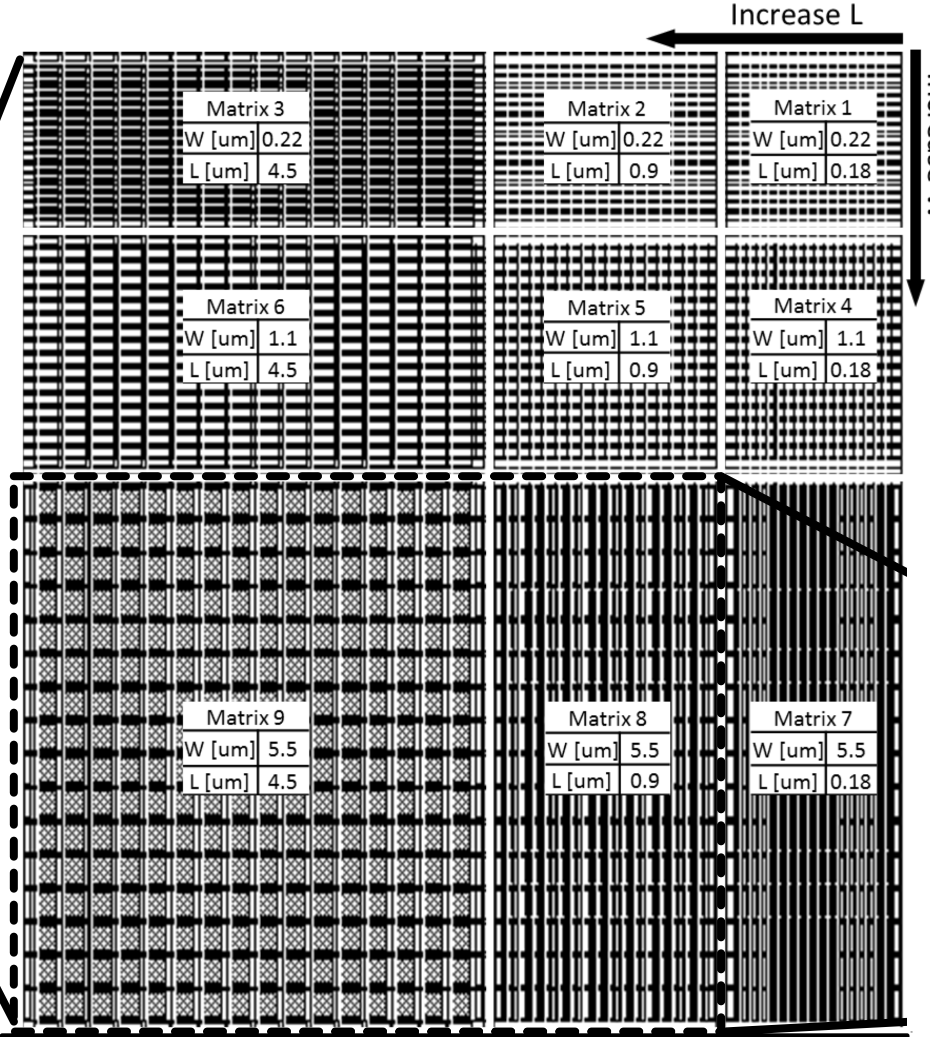


MOSFET 2D ADDRESSABLE MATRIX - IMPLEMETANTION

48 x 48 DUT MOSFET Matrix (2304 devices divided into 9 sub-groups of 256 same size transistors.



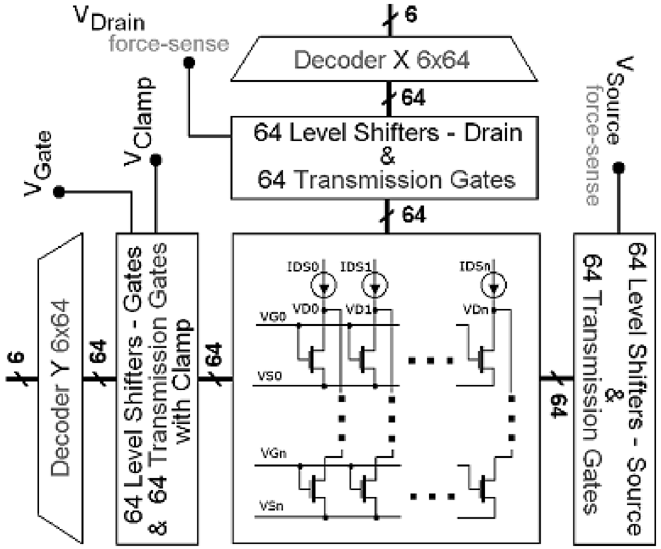
Uses 2 6-64 Address decoders
Includes 3 blocks of 64 Level shifters (from 1.8V to 3.3V) and 3 blocks of 64 Transmission Gates.



SilTerra 0.18µm bulk CMOS process

Matrix #	W [μm]	L [μm]	$1/\sqrt{(WL)}$	W/L
M1	0.22	0.18	5.025	▲1.222
M2	0.22	0.90	✦ 2.247	✚0.244
M3	0.22	4.50	■ 1.005	0.049
M4	1.10	0.18	✦ 2.247	▼6.111
M5	1.10	0.90	■ 1.005	▲1.222
M6	1.10	4.50	● 0.449	✚0.244
M7	5.50	0.18	■ 1.005	30.556
M8	5.50	0.90	● 0.449	▼6.111
M9	5.50	4.50	0.201	▲1.222

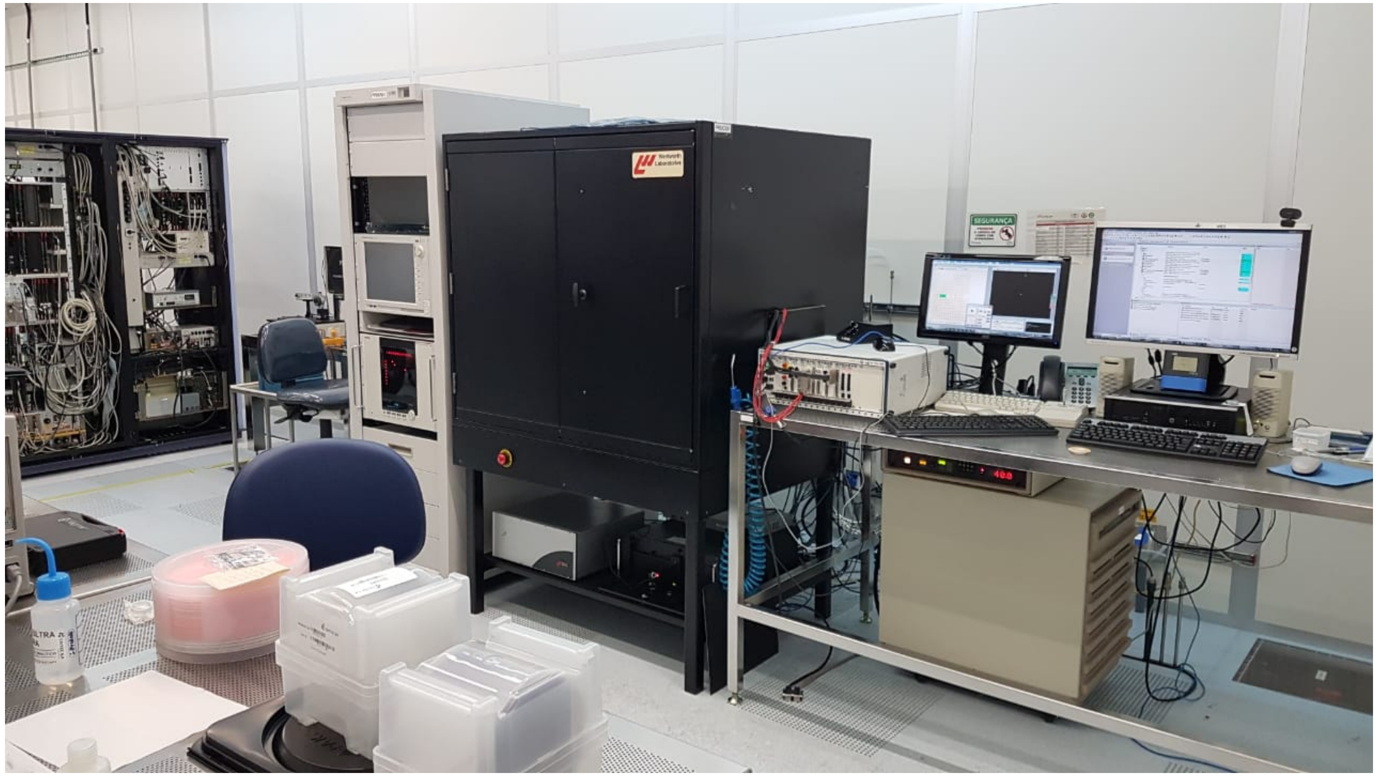
Final device sizes.



MEASUREMENT SETUP – INSIDE CEITEC'S CLEAN ROOM

- Wentworth Semi-auto probe station with thermal Chuck.
- B1500 Parameter Analyzer.
- Keysight ENA E5071C.
- Keithley Switch matrix
- Keithley S530 Parametric Test System.
- Agilent IC-CAP for parameter extraction

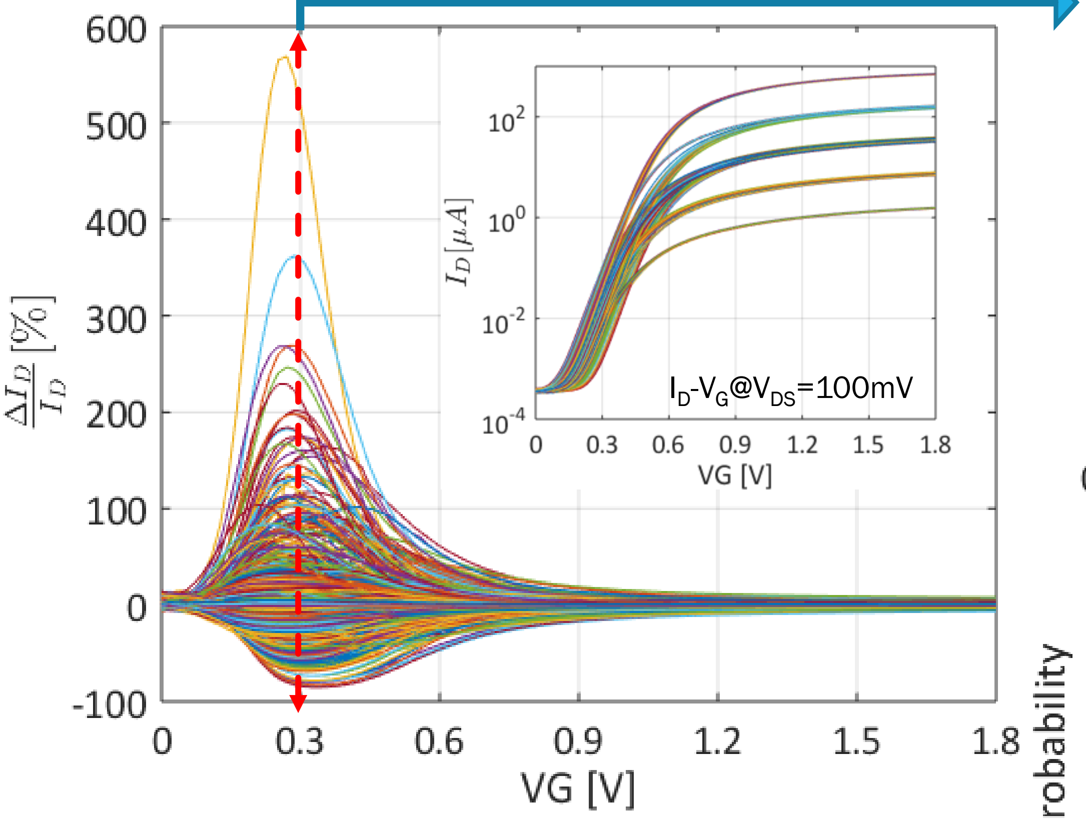
Inside class 10000 clean room – my work place

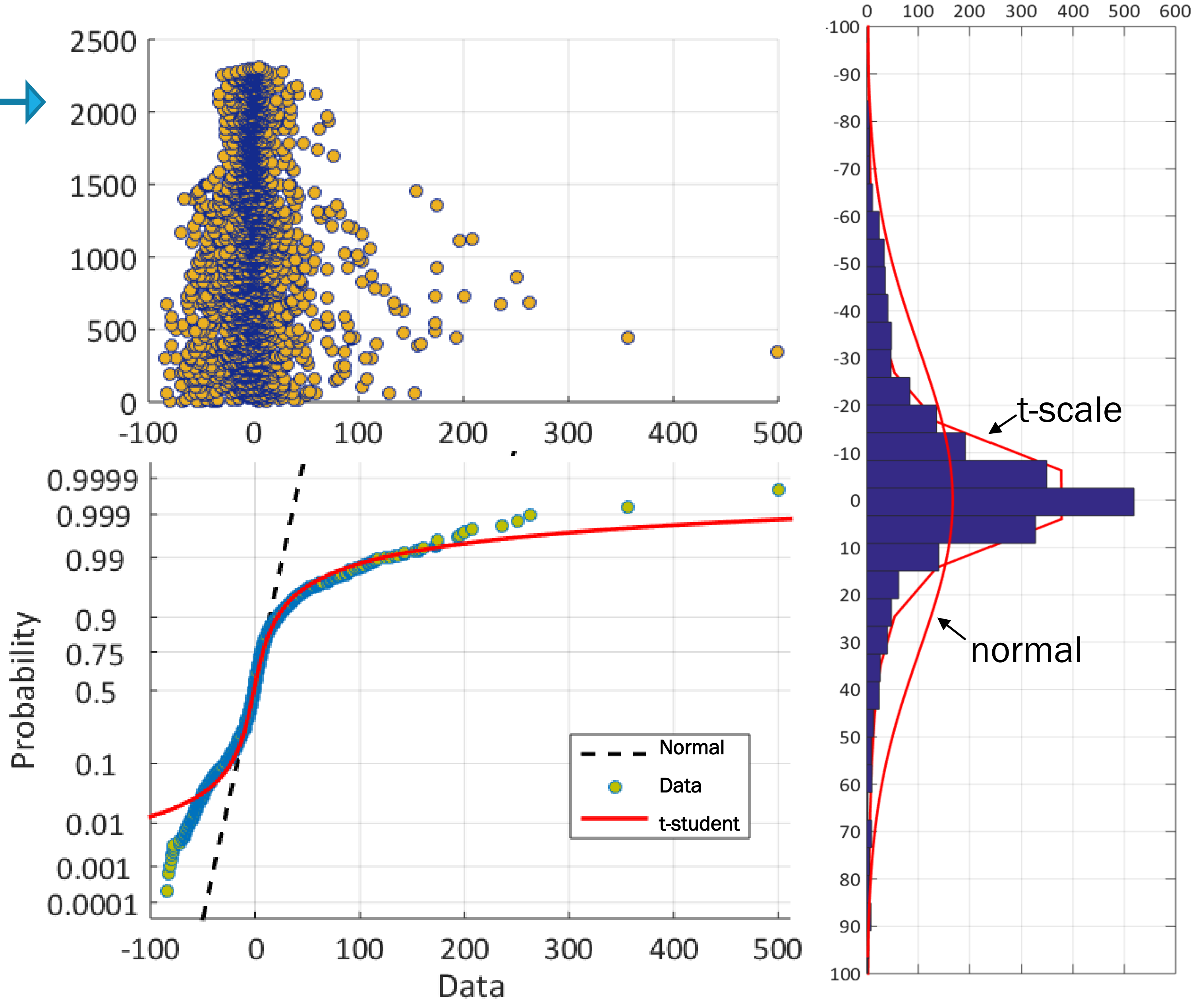


Brazilian state-owned semiconductor fab and design center. Class 100 and 10000 cleanrooms. Up and running: Wafer dicing, grinding, packaging, probe testing, and failure analysis.

ID-VG DATA DISTRIBUTION

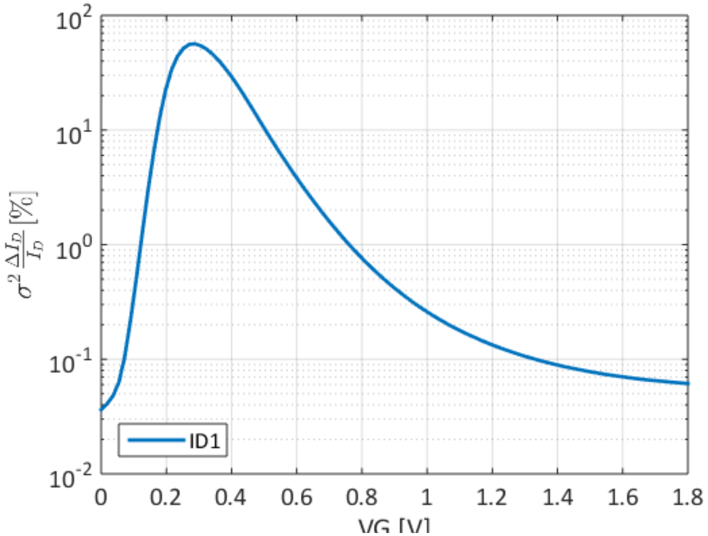
Data @ VG=300mV 'cross section'





In probability and statistics, **Student's *t*-distribution** (or simply the ***t*-distribution**) is any member of a family of continuous probability distributions that arise when estimating the mean of a normally-distributed population in situations where the sample size is small and the population's standard deviation is unknown.

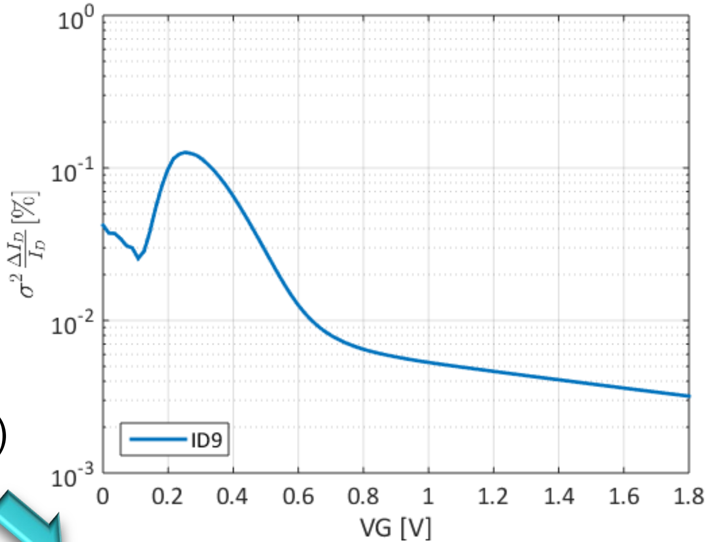
Wikipedia-Student's_t-distribution



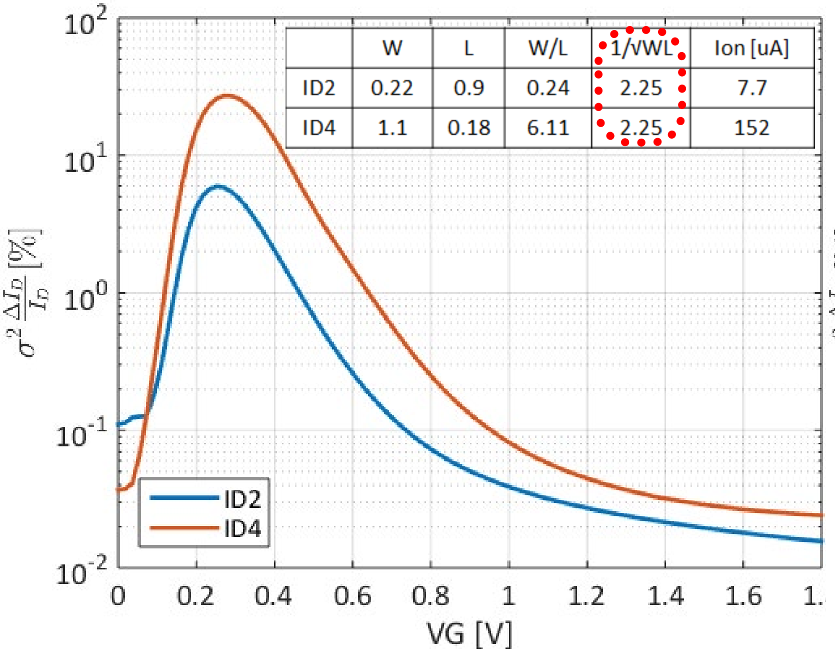
ΔID/ID VARIANCE

General model

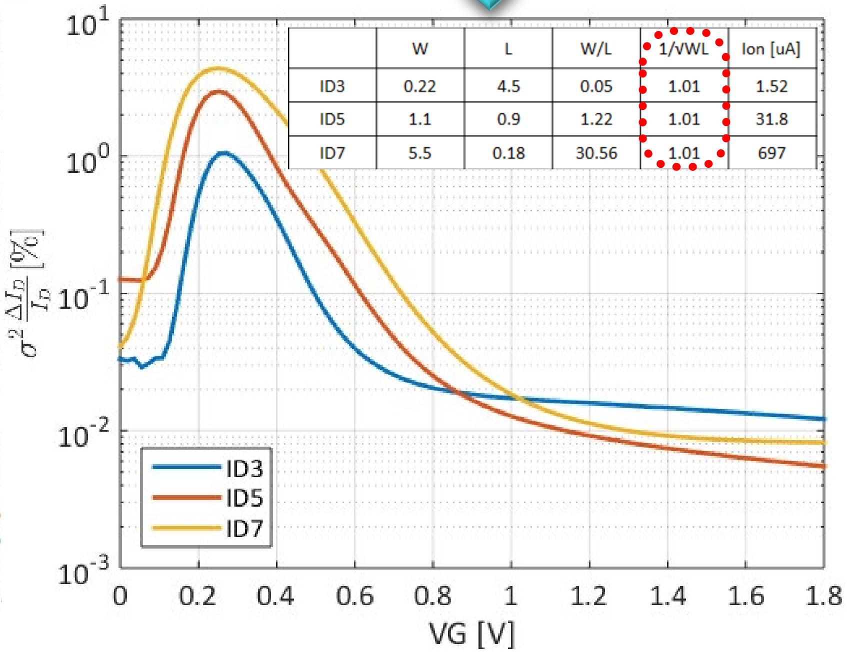
$$\sigma^2\left(\frac{\Delta I_D}{I_D}\right)=\frac{1}{WL}\left(A_\beta^2+\left(\frac{gm}{I_D}\right)^2 A_{V_{TH}}^2\right)$$



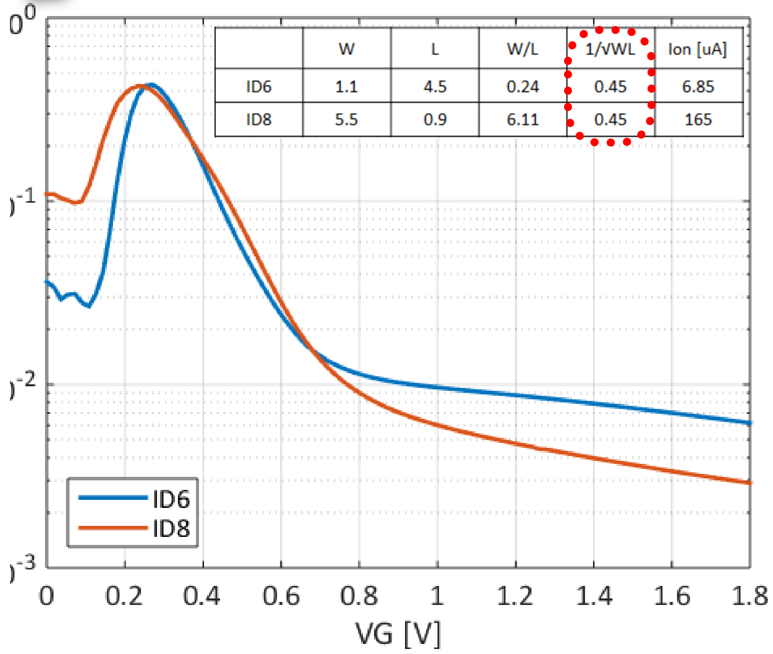
Curves of $\sigma^2(\Delta I_D/I_D)$ with the same $1/\sqrt{(WL)}$ term. Some unexplainable differences



	W	L	W/L	1/√WL	Ion [uA]
ID2	0.22	0.9	0.24	2.25	7.7
ID4	1.1	0.18	6.11	2.25	152



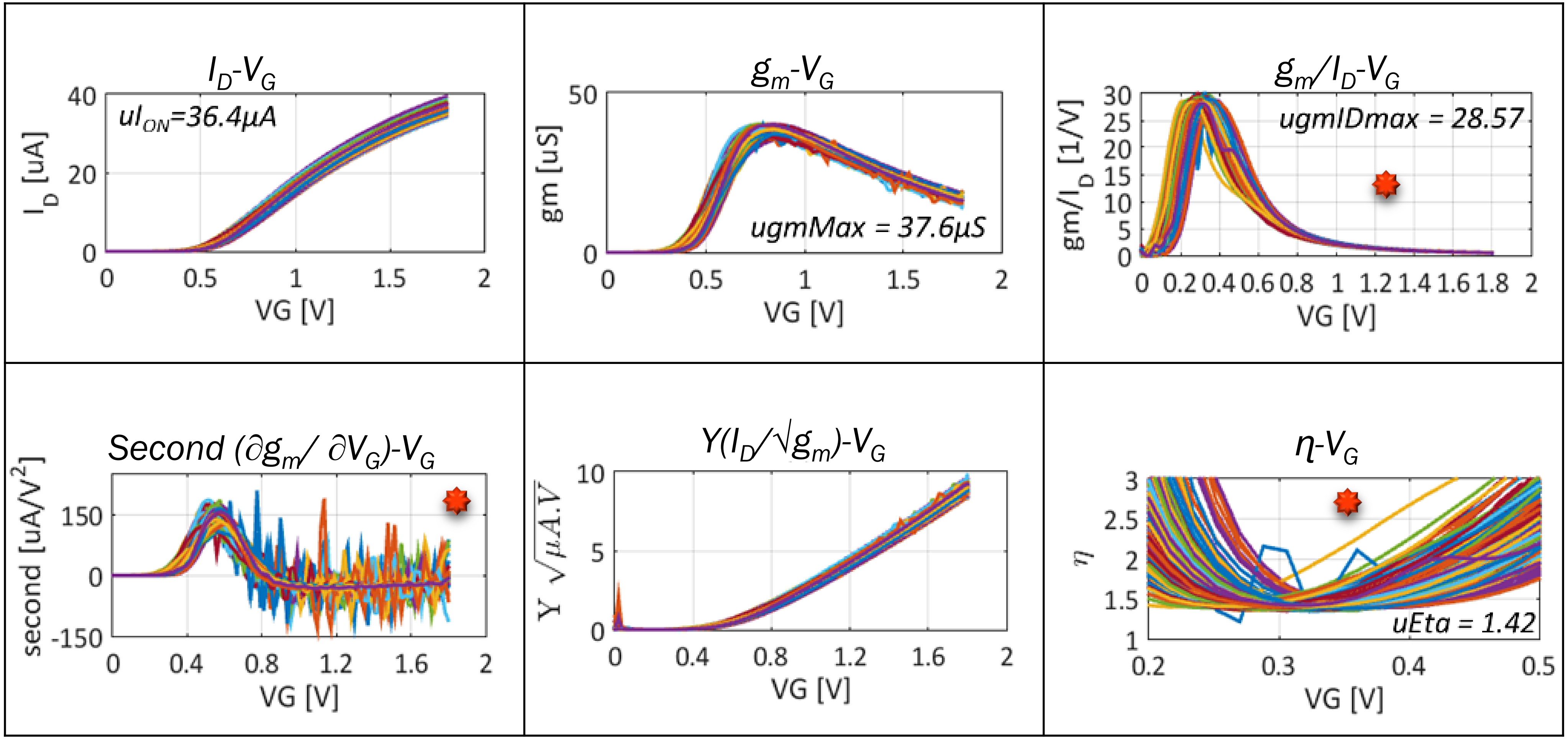
	W	L	W/L	1/√WL	Ion [uA]
ID3	0.22	4.5	0.05	1.01	1.52
ID5	1.1	0.9	1.22	1.01	31.8
ID7	5.5	0.18	30.56	1.01	697



	W	L	W/L	1/√WL	Ion [uA]
ID6	1.1	4.5	0.24	0.45	6.85
ID8	5.5	0.9	6.11	0.45	165

PARAMETER EXTRACTION - ALL NEEDED CURVES

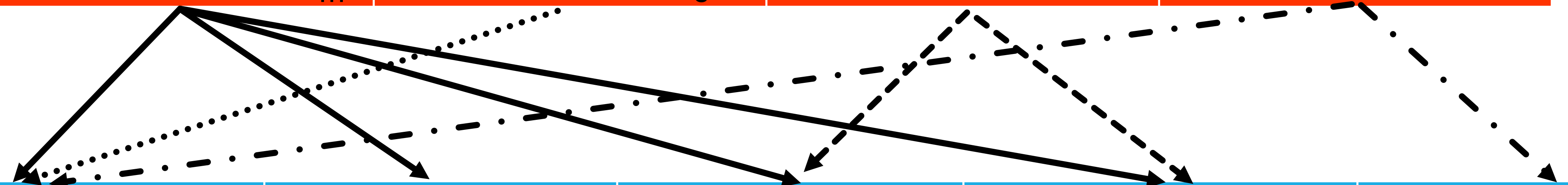
Very noisy values



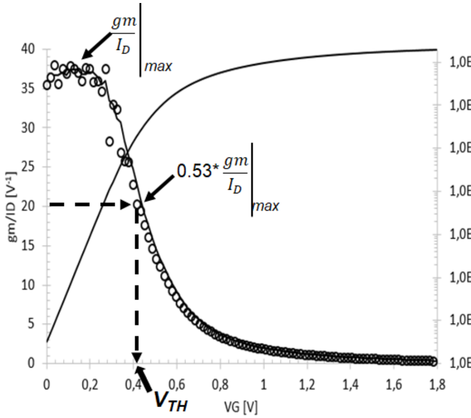
PARAMETER EXTRACTION

4 parameters vs 5 methods

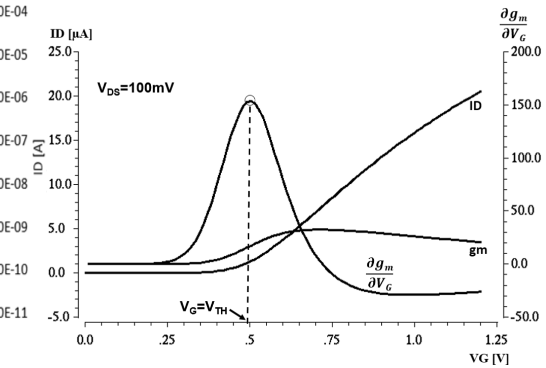
Threshold voltage (V_{TH})	Specific current (I_S)	Current gain (β)	Slope factor (η)



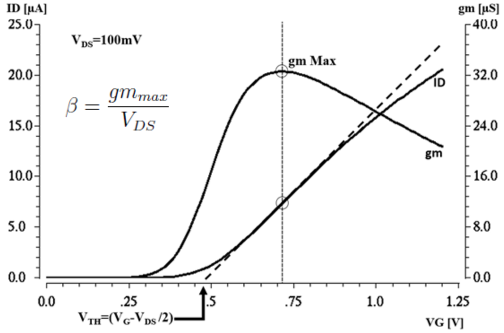
Transconductance-to-current ratio	Second derivative	Linear extrapolation	Y-function	Subthreshold slope



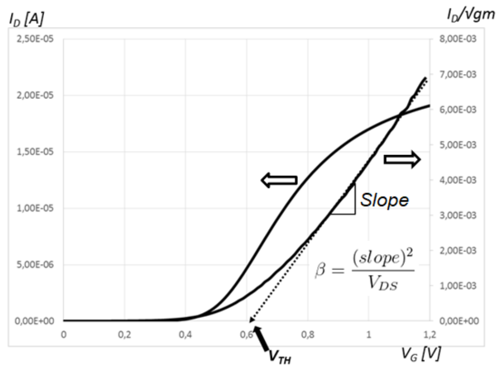
Schneider; Galup-montoro, 2010b



Wong et al., 1987

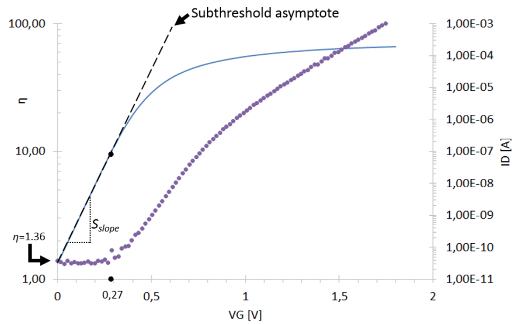


Schroder, 2006a



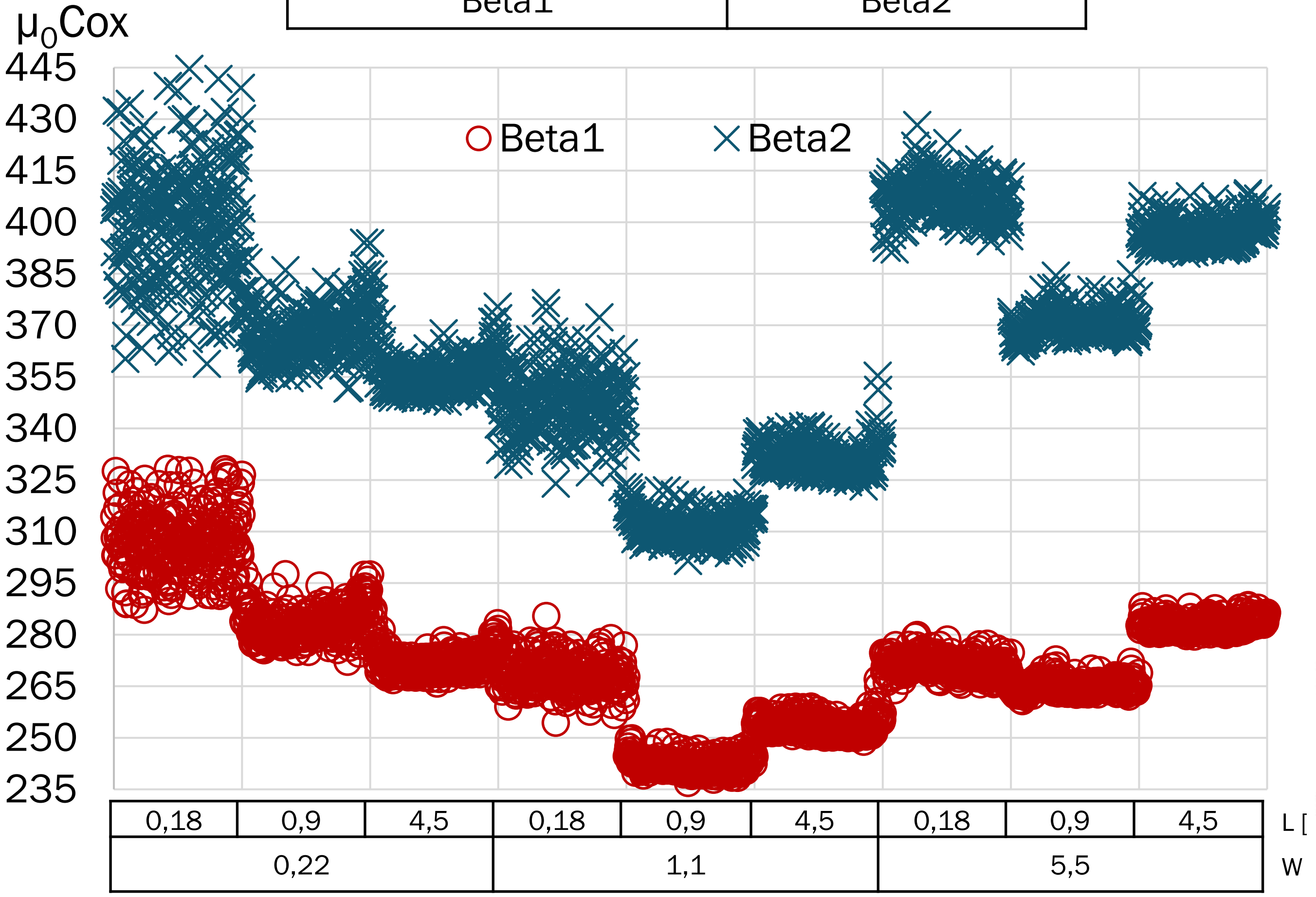
Ghibaudo, 1988

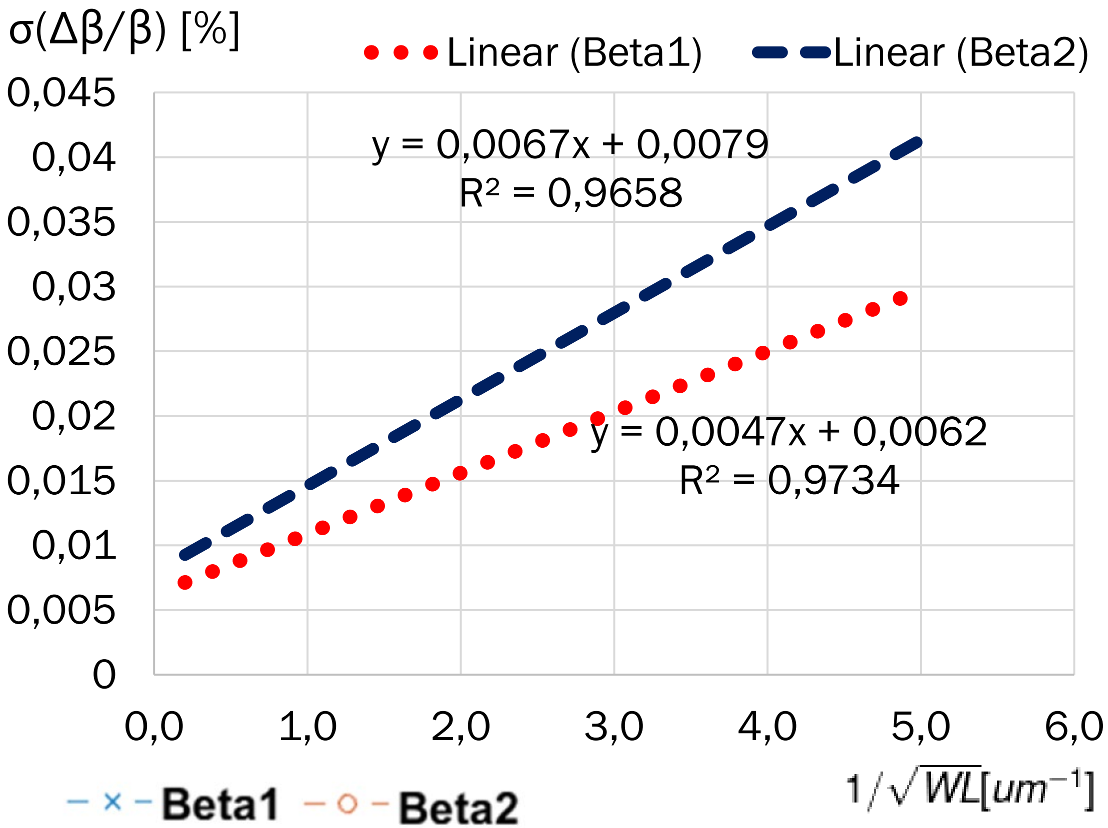
Sze, Ng; 2008

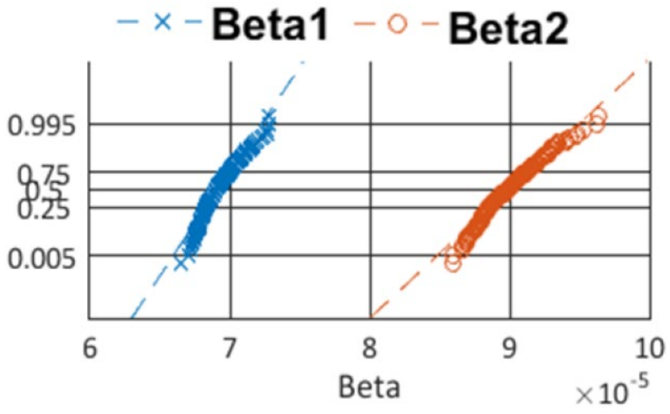


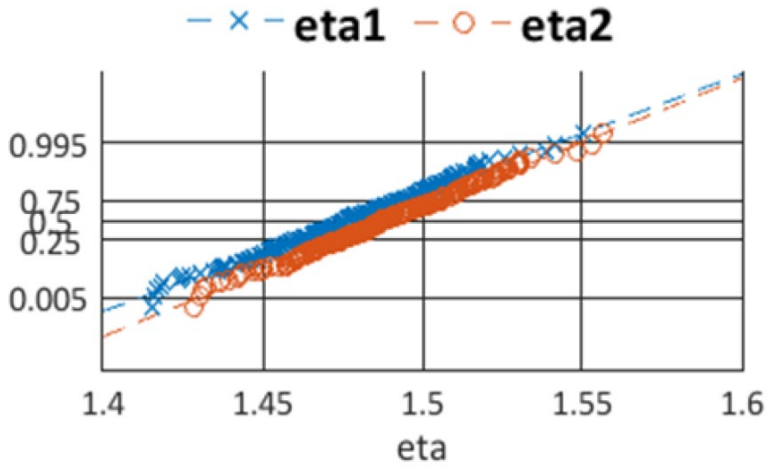
BETA NORMALIZED – L/W

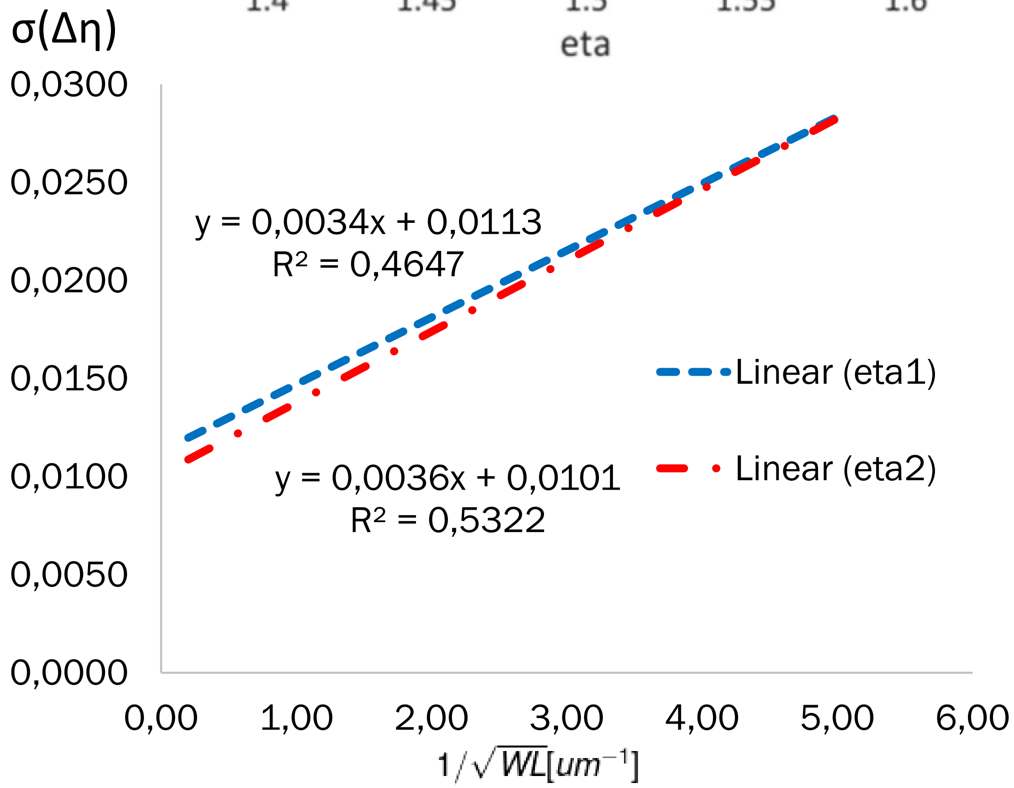
Linear extrapolation	Y-function
Beta1	Beta2





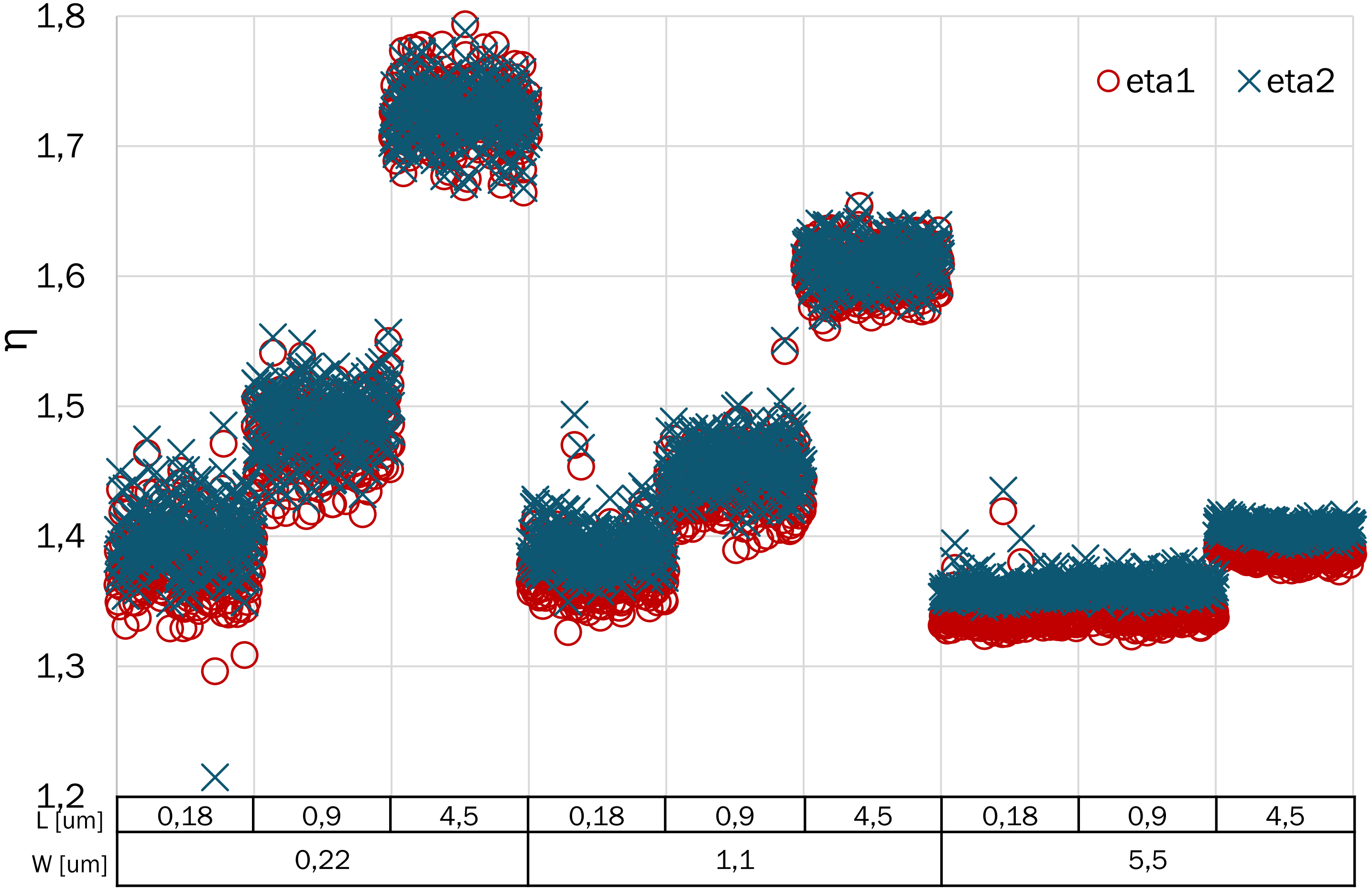






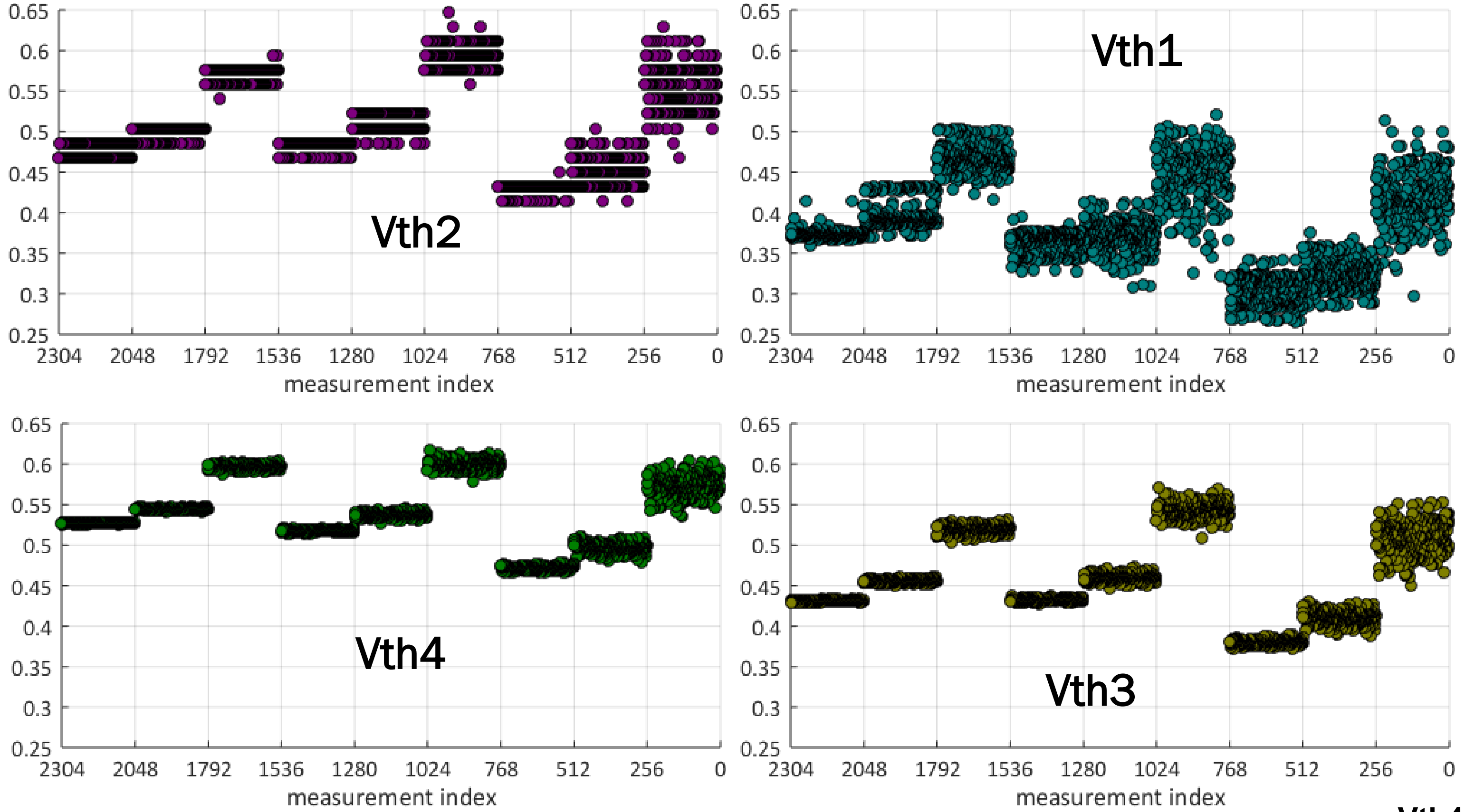
SLOPE FACTOR

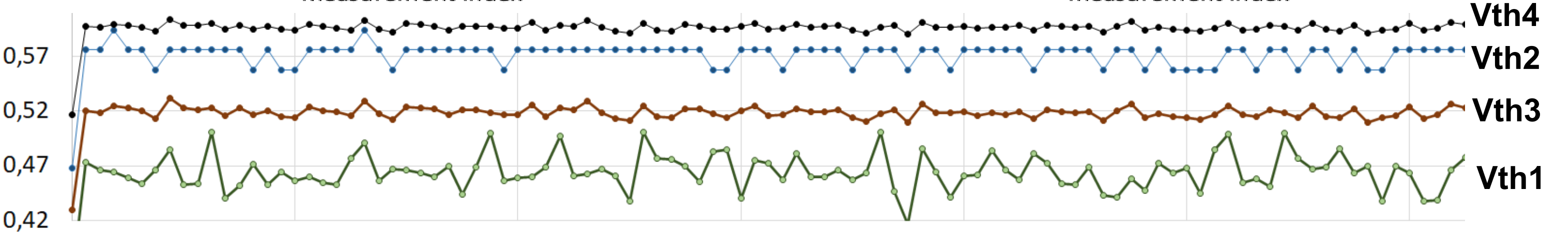
gm/I_D	Subthreshold slope
eta1	eta2

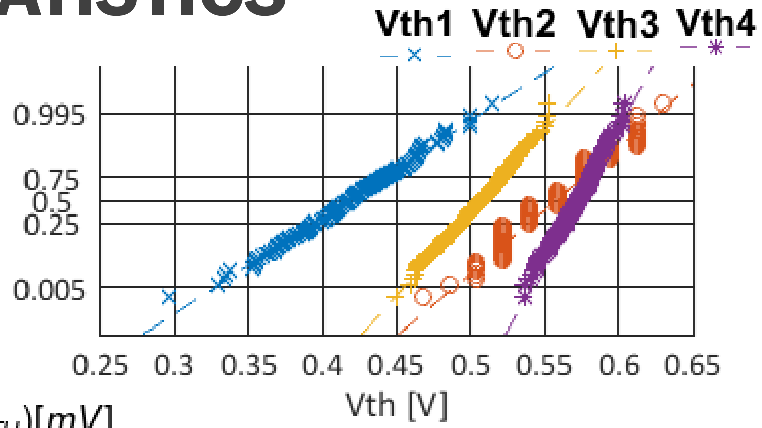


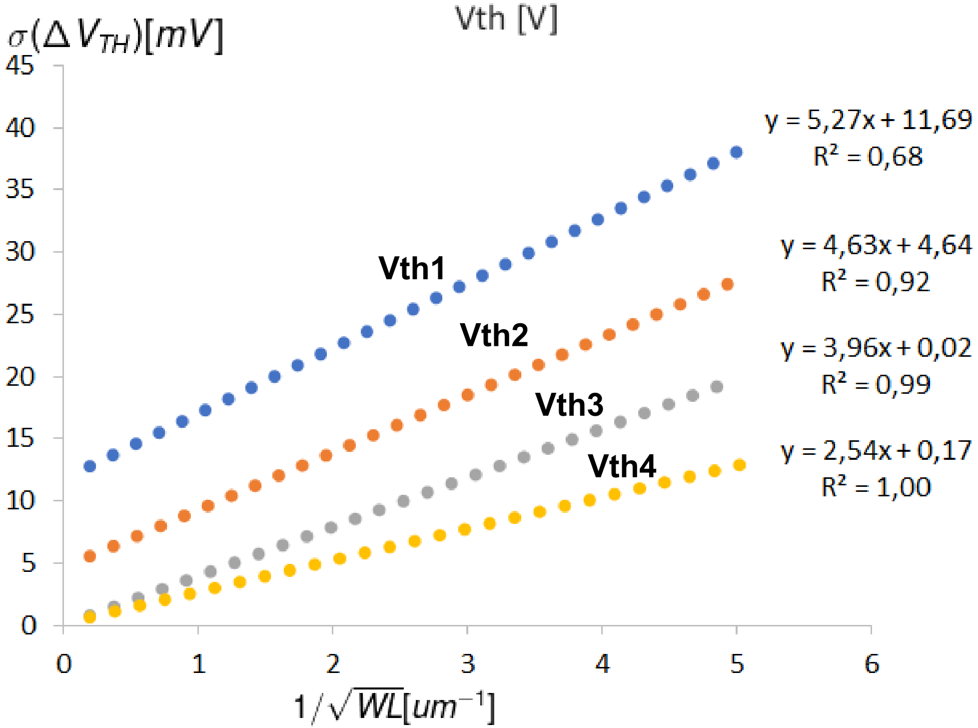
THRESHOLD VOLTAGE ‘VISUAL’ STATISTICS

gm/ID	Second derivative	Linear extrapolation	Y-function
Vth1	Vth2	Vth3	Vth4

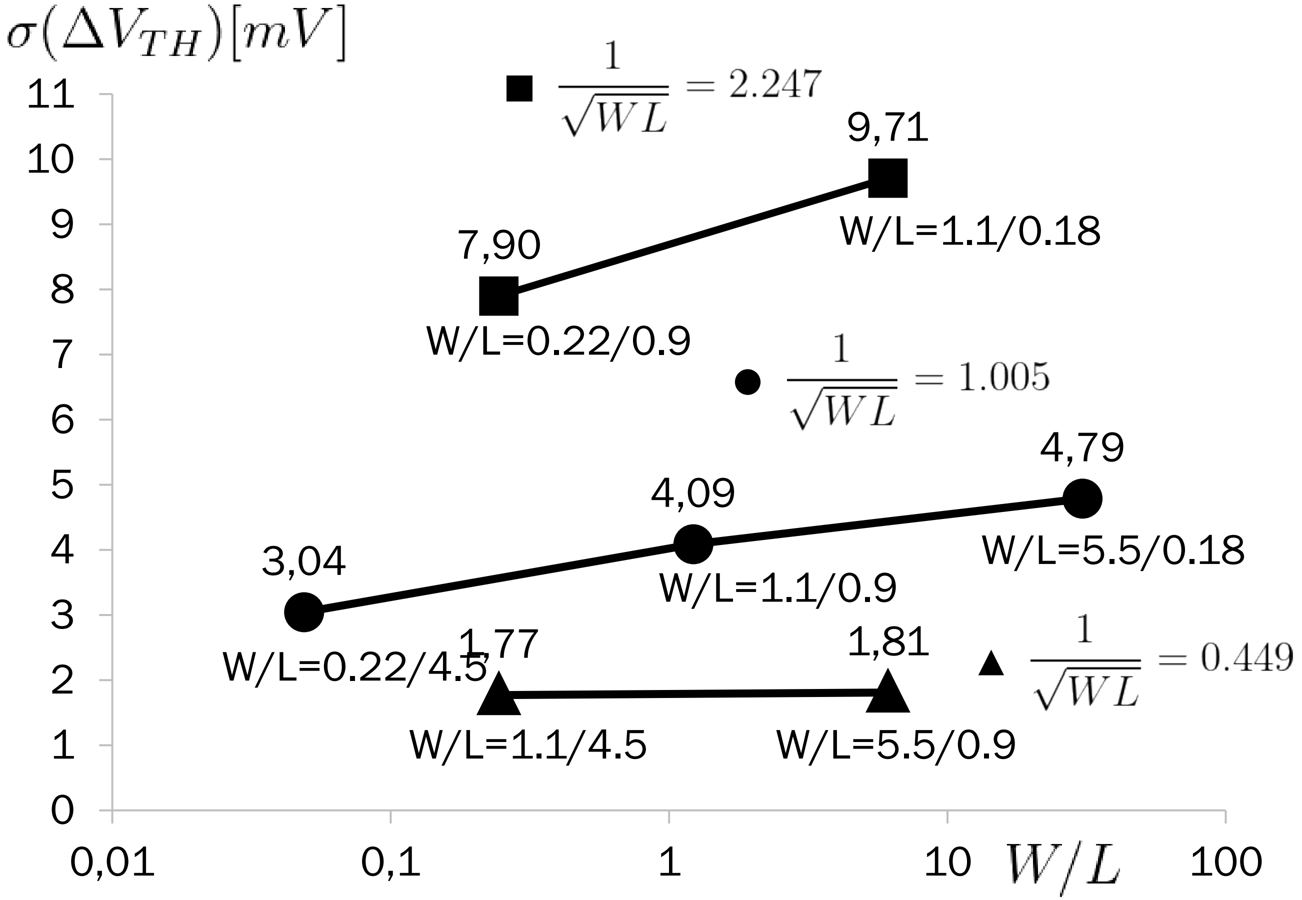


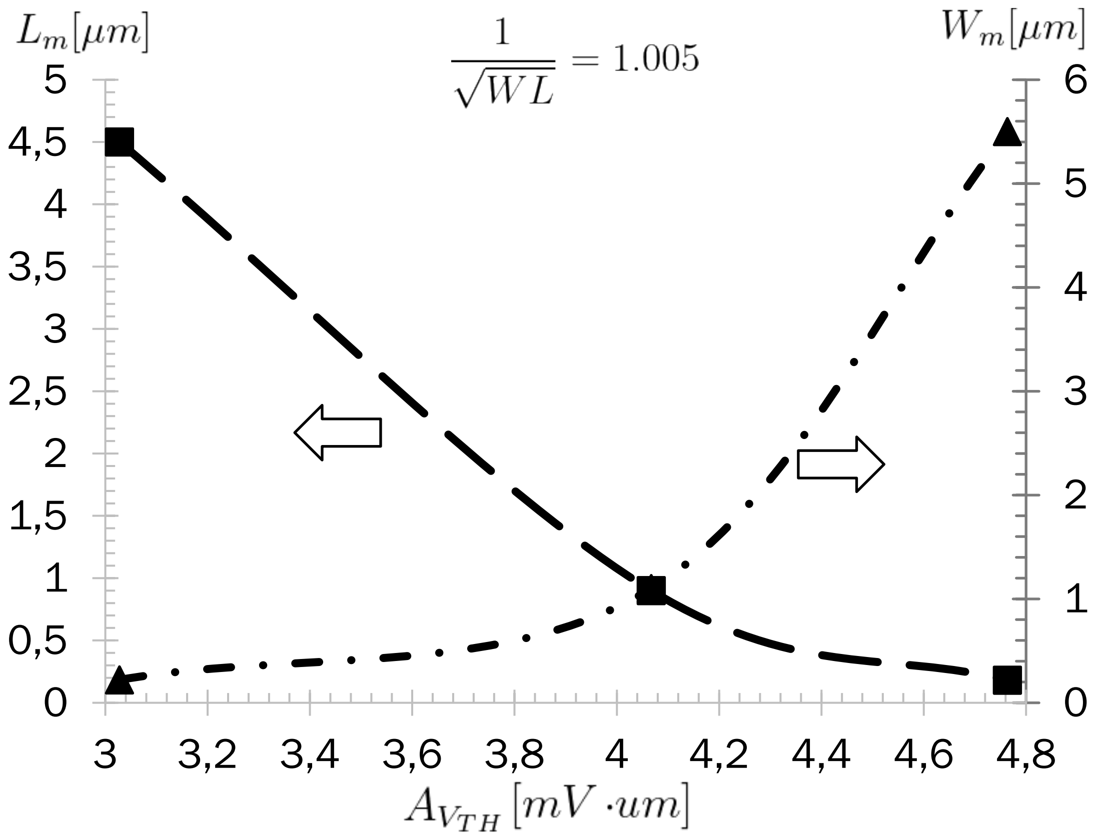






SAME ACTIVE AREA DEVICES ≠ STANDARD DEVIATION





Open questions:

1) Are these differences due to Mask size versus Effective channel sizes of the devices?
2) Are these differences due to the influence of short channel effects on sigma delta Vth?
3) Are this differences due to a combination of item 1 and 2?

EXTRACTION OF L_{EFF}: SHIFT-AND-RATIO METHOD DEFINITION

$$\frac{V_{\mathrm{DS}}}{I_{\mathrm{DS}}} = R_{\mathrm{tot}} = R_{\mathrm{ch}} + R_{\mathrm{sd}} = L_{\mathrm{eff}}\, f\,(V_{\mathrm{GS}} - V_{\mathrm{TH}}) + R_{\mathrm{sd}}$$

$$S = \frac{\partial R_{\mathrm{tot}}}{\partial V_{\mathrm{G}}} = L_{\mathrm{eff}} \frac{\partial f\,(V_{\mathrm{GS}} - V_{\mathrm{TH}})}{\partial V_{\mathrm{G}}} \quad ; \quad \frac{\partial R_{sd}}{\partial V_G} \approx 0$$

$$R(V_{\mathrm{G}}) = \frac{S^L}{S^S(\Delta\delta \equiv \Delta V_{\mathrm{TH}})} \Rightarrow \mathrm{constant}(V_{\mathrm{G}})$$

$$R = \frac{L^L_{\mathrm{eff}}}{L^S_{\mathrm{eff}}} = \frac{(L^L_{\mathrm{m}} - \Delta L)}{(L^S_{\mathrm{m}} - \Delta L)} \Rightarrow \Delta L = \frac{R L^S_{\mathrm{m}} - L^L_{\mathrm{m}}}{R - 1}$$

Y. Taur 1992

Compensates the threshold voltage difference due to L*f*(Vth)

Shifted ratio of the derivative to output resistance of two transistors, a Long and Short.

Allows us to analyzes the VG range and the dependence of ΔVth with ΔL

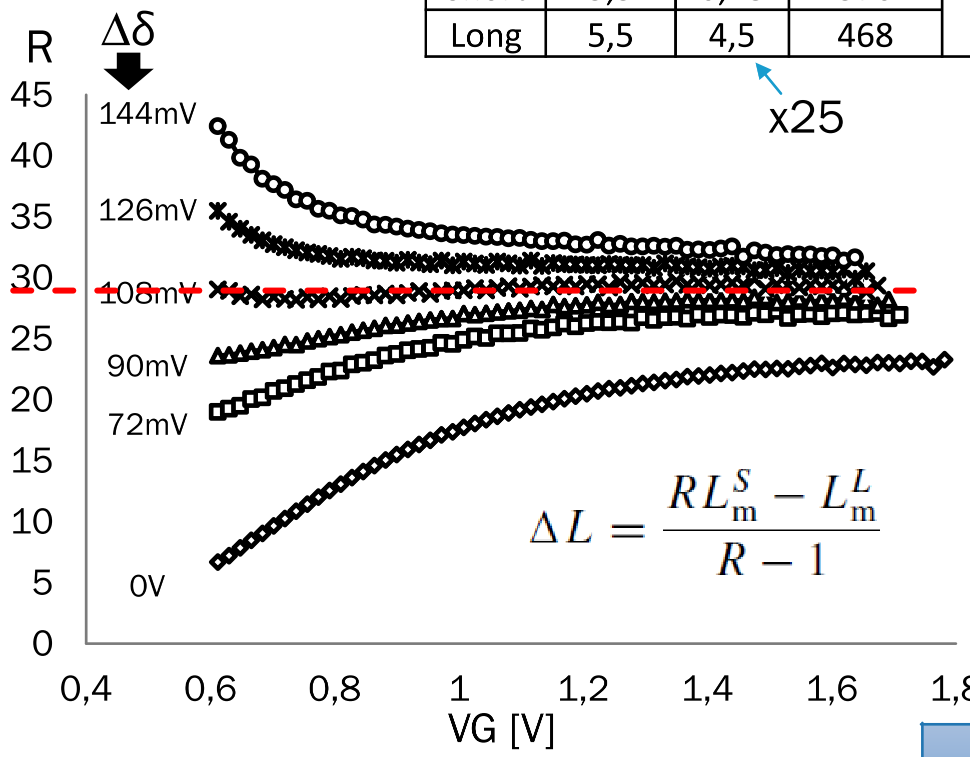
J. P. M. Brito and S. Bampi, "Local Variability Evaluation on Effective Channel Length Extracted With Shift-and-Ratio Method," in *IEEE Transactions on Electron Devices*, vol. 67, no. 11, pp. 4662-4666, Nov. 2020.

Thank you for the reviewers!

EXTRACTION OF L_{EFF}: SHIFT-AND-RATIO METHOD GRAPHICALLY

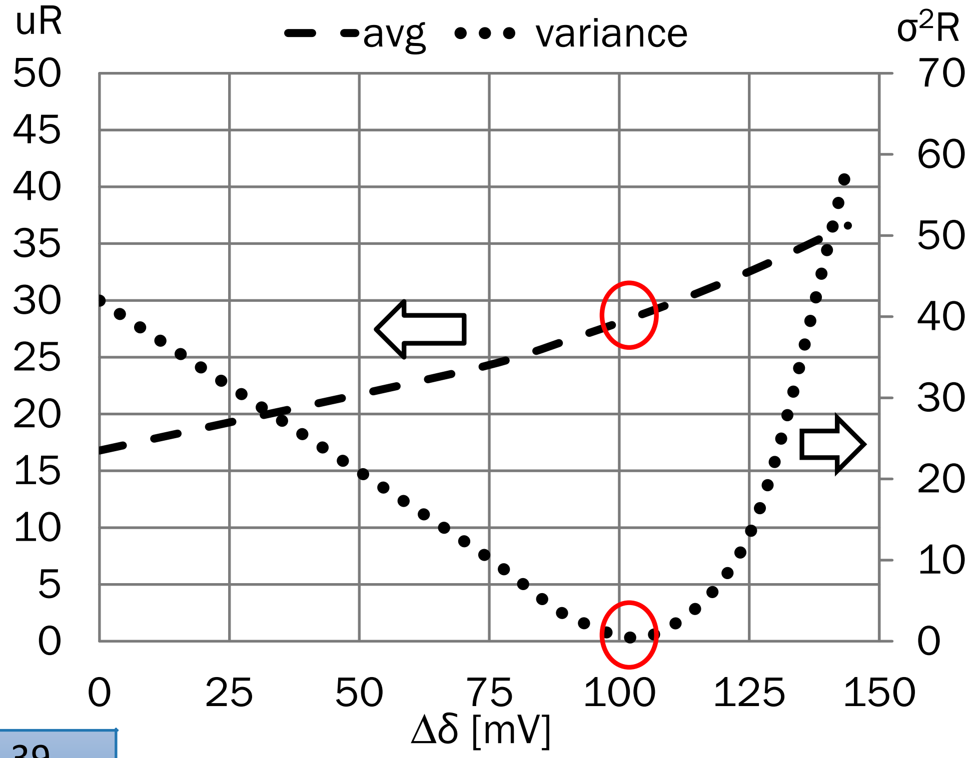
Device	W [µm]	L [µm]	Vth [mV]	ΔVth [mV]
Short	5,5	0,18	576	108
Long	5,5	4,5	468	

x25



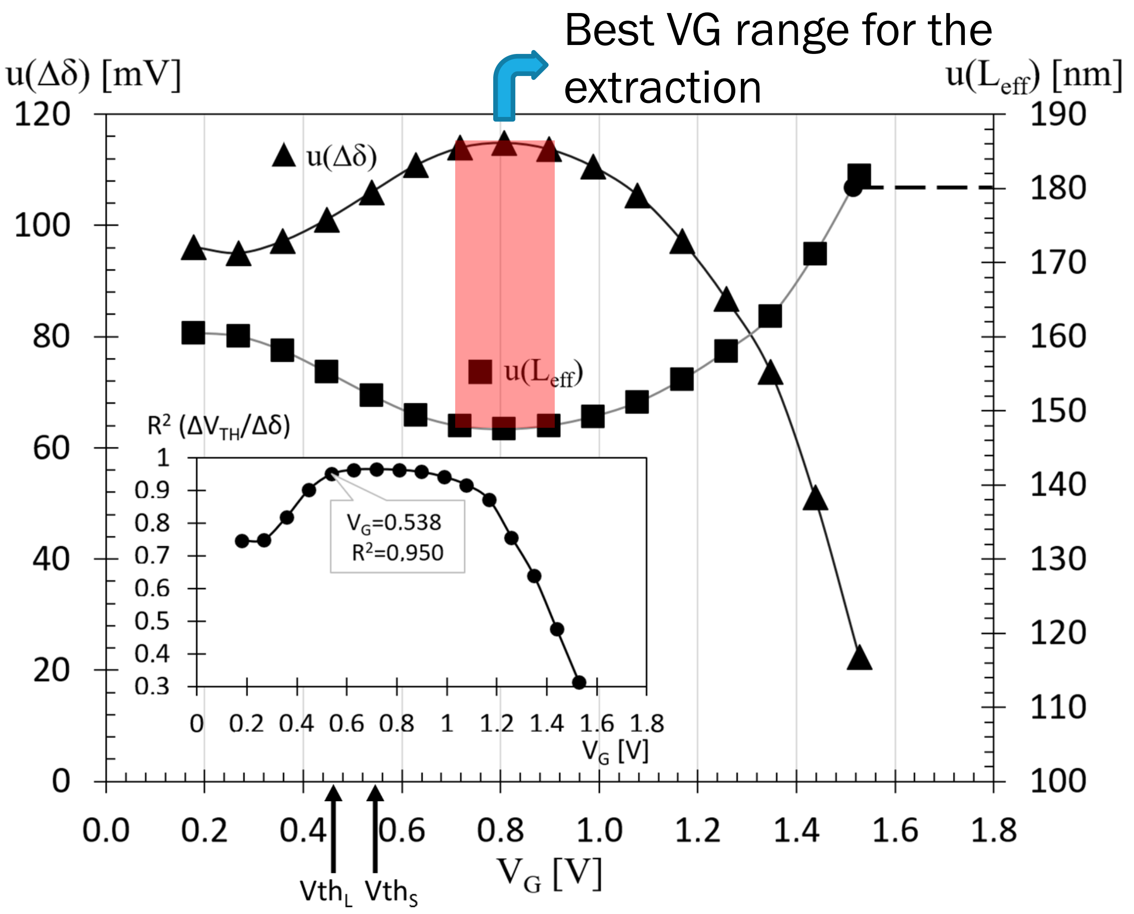
$$\Delta L = \frac{R L_m^S - L_m^L}{R - 1}$$

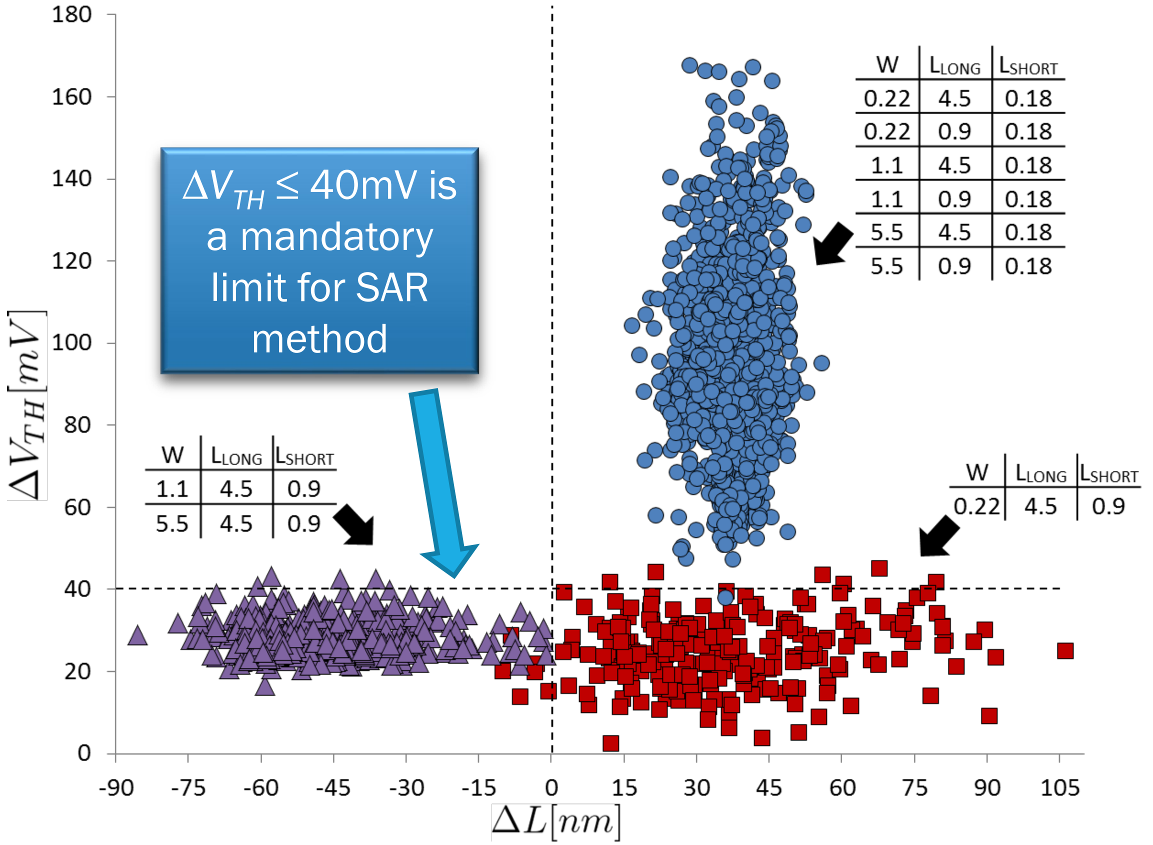
Y. Taur 1992



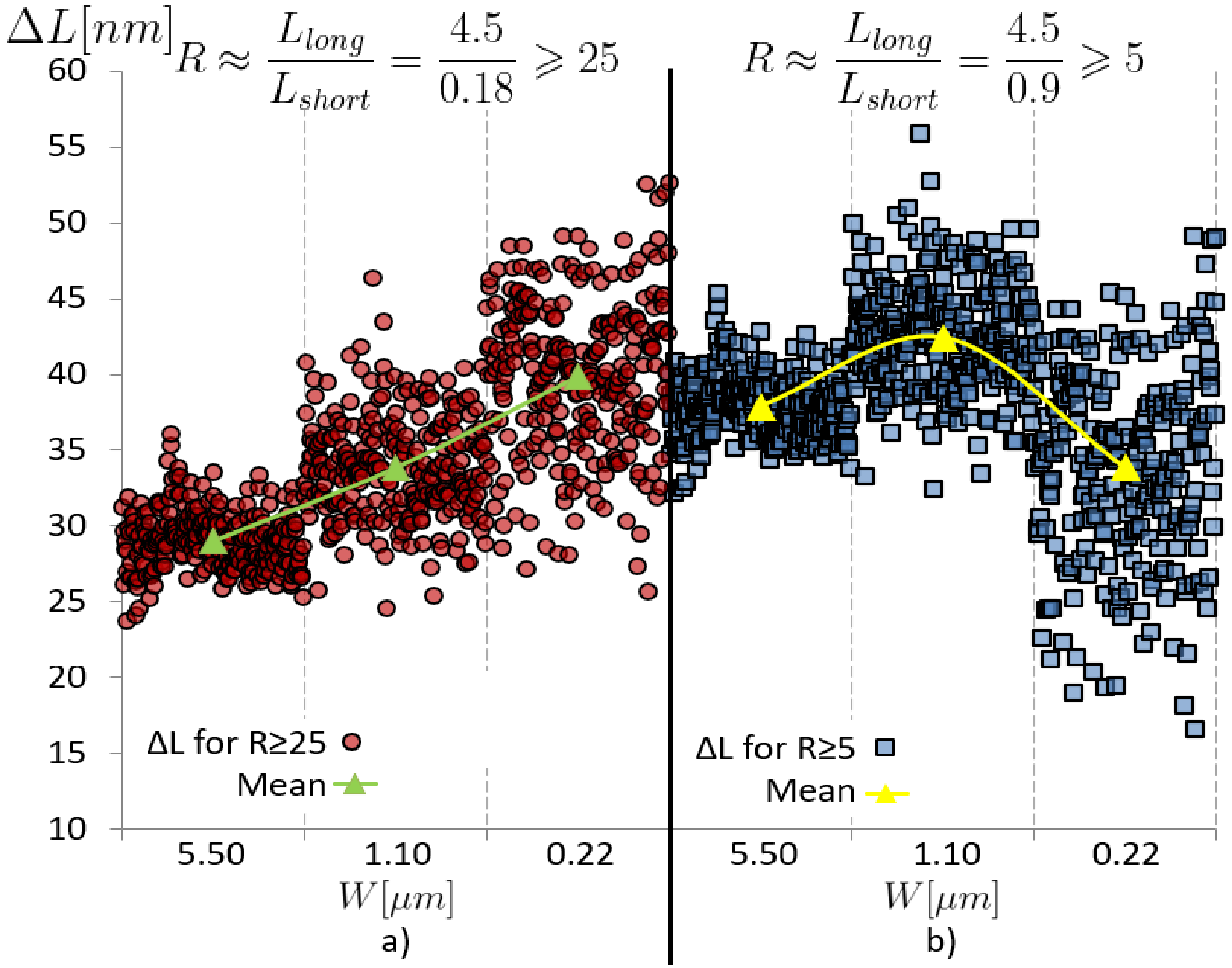
R	29.39
ΔL	27,84nm
Leff	153,11nm

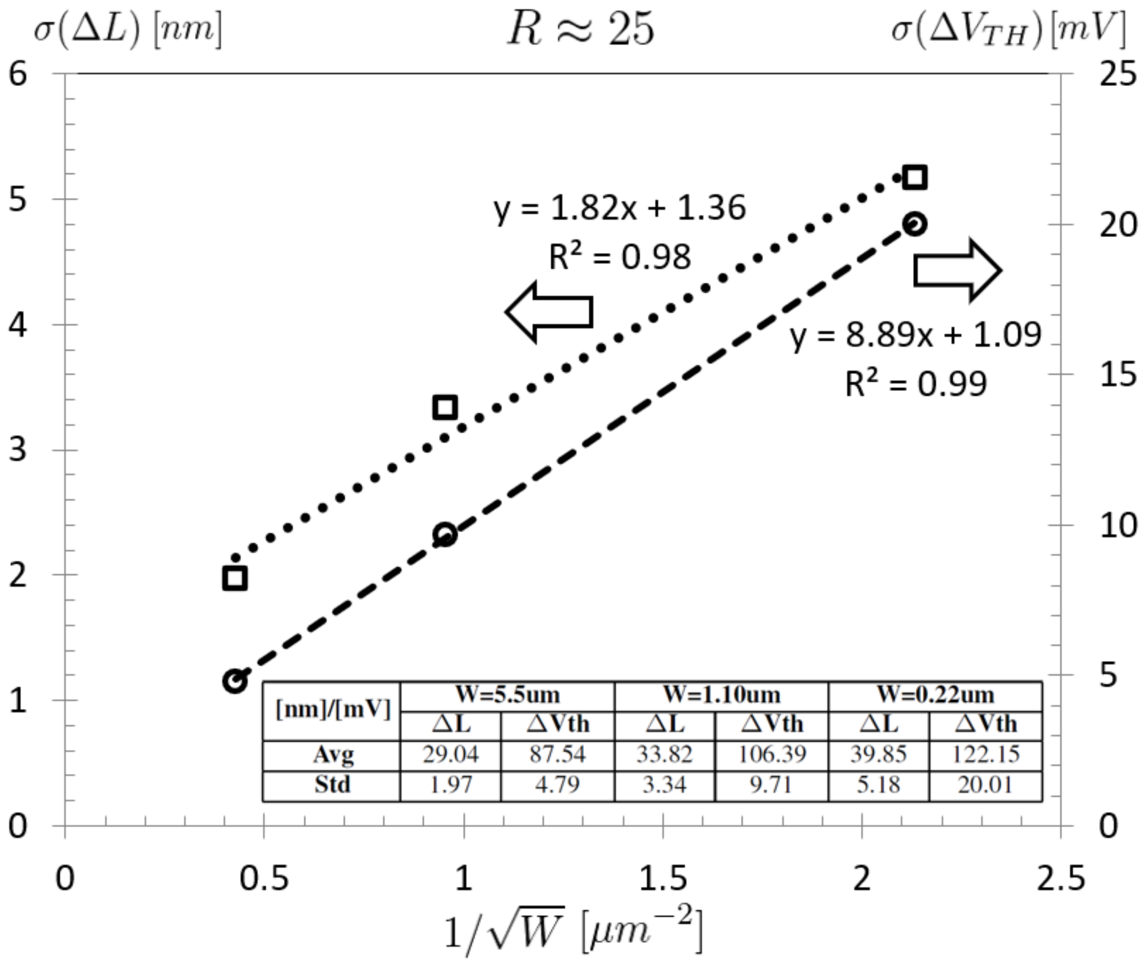
VG RANGE - ΔV_{TH} VERSUS ΔL CORRELATION





ΔL DISTRIBUTION AND VARIATION





[nm]/[mV]	W=5.5um		W=1.10um		W=0.22um	
	ΔL	ΔVth	ΔL	ΔVth	ΔL	ΔVth
Avg	29.04	87.54	33.82	106.39	39.85	122.15
Std	1.97	4.79	3.34	9.71	5.18	20.01

CONCLUSIONS

- Conclusions from the curves:
  - We suggest that the distribution data can be modeled by the t-student distribution.
  - Secondary device size effects need to be taken into account for a mismatch modeling equation.
- Conclusions from the parameters:
  - The current factor parameter (beta) is better extracted with the y-function for a statistical significance.
  - The threshold voltage parameter is better extracted with the linear extrapolation method in order to also have a statistical significance.
  - Transistors with the same 1/sqrt(area) term does not gives the Avt parameter.
  - Effective channel length it is not straightforward to analyze. In meantime, shift and ratio method gives good results for R around 25 and delta vt above 40mV.

- Thank you
- Obrigado

Any questions/suggestions ?

Sorry for mixing commas (,) with dots(.) 😐