Translinear - based neural circuit primitives

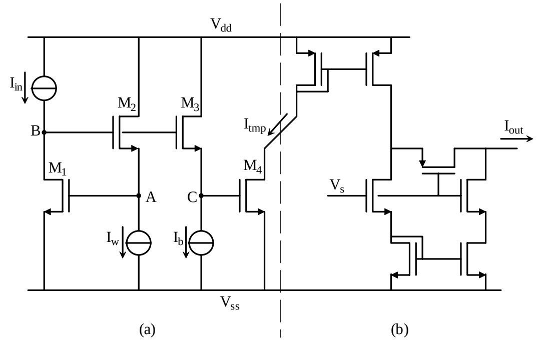
Maurizio Valle

Two quadrant synaptic multiplier (1)

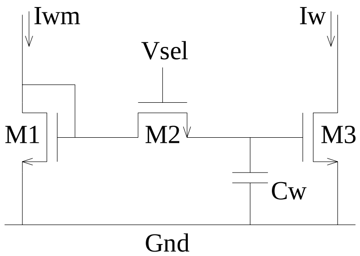
$$I_{in} I_w = I_b I_{tmp}$$

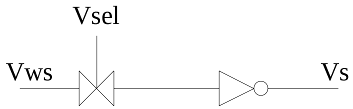
$$I_{tmp} = I_{in} \frac{I_w}{I_b}$$

$$\begin{aligned} &if && V_5 = V_{ss} \\ & && I_{out} = +I_{tmp} = +I_{in} \frac{I_w}{I_b} \\ &elseif && V_5 = V_{dd} \\ & && I_{out} = -I_{tmp} = -I_{in} \frac{I_w}{I_b} \\ &end && \end{aligned}$$

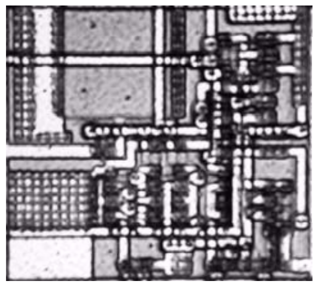


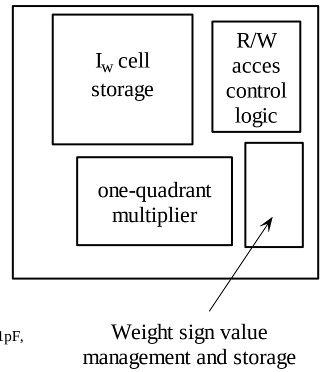
Two quadrant synaptic multiplier(2)





Two quadrant synaptic multiplier(3)

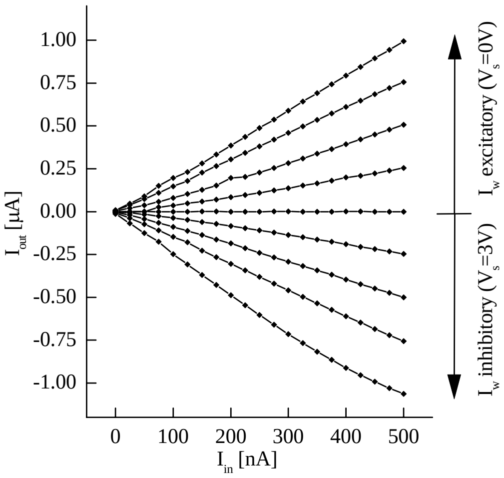




- the one-quadrant multiplier;
- a current copier cell for the dynamic storage of the weight current (with a 1pF, 30µm×30µm MOS capacitor);
- the circuit for the weight sign value management and storage (a simple DRAM cell);
- the control logic for the read/write (R/W) access operation.

Two quadrant synaptic multiplier(4)

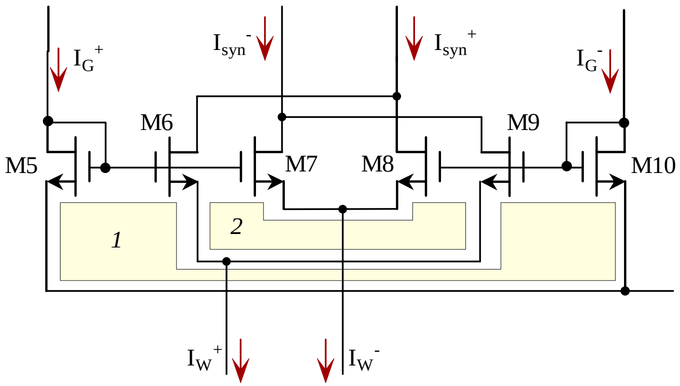
- ***Analog current mode multiplier***
- ***Power consumption***: < 30 μW
- ***Processing delay***: < 300ns
- ***Weight refresh time***: 2 ms
- ***Size***: 100 x 100 μm^2 (ECPD10)
- ***No. of transistors***: 31



Analog four quadrant multiplier

- differential coding of information ⇒ high noise and interference immunity;
- current mode of operation which ⇒ : i) high robustness with respect to spread of the technological parameters; ii) wide dynamic range of signals; iii) easy implementation of sums;
- weak inversion region of operation of transistors ⇒ low power / low voltage operation;
- translinear circuits ⇒ versatility and efficiency (i.e. small power consumption)

Analog four quadrant multiplier: basic circuit



Analog four quadrant multiplier: coding of information

$$I_{Syn} = x \cdot w \cdot I_B$$

$$I_G^+ = (1+x)\frac{I_B}{2},\ I_G^- = (1-x)\frac{I_B}{2} \qquad I_G = I_G^+ - I_G^- = xI_B$$

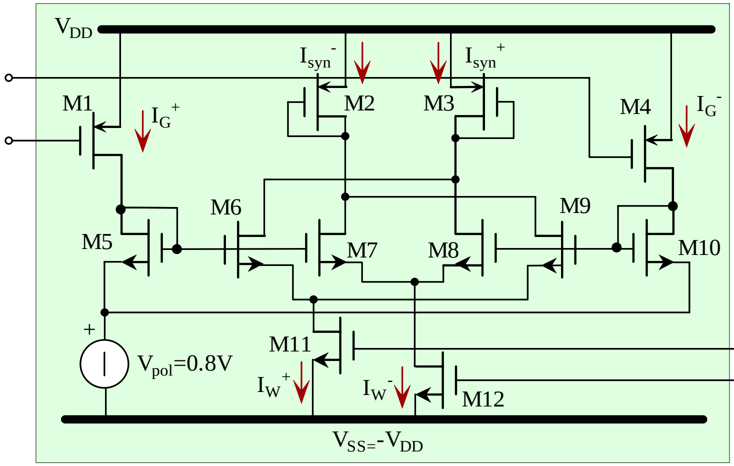
$$I_W^+ = (1+w)\frac{I_B}{2},\ I_W^- = (1-w)\frac{I_B}{2} \qquad I_W = I_W^+ - I_W^- = wI_B$$

$$I_{Syn}^+ = (1+x\cdot w)\frac{I_B}{2}$$

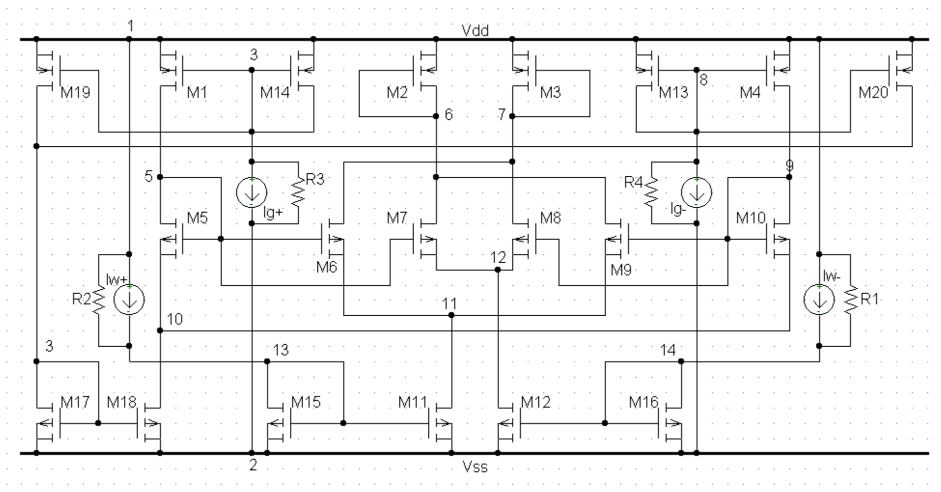
$$I_{Syn} = I_{Syn}^+ - I_{Syn}^- = x\cdot w\cdot I_B$$

$$I_{Syn}^- = (1-x\cdot w)\frac{I_B}{2}$$

Analog four quadrant multiplier: the overall circuit (1)

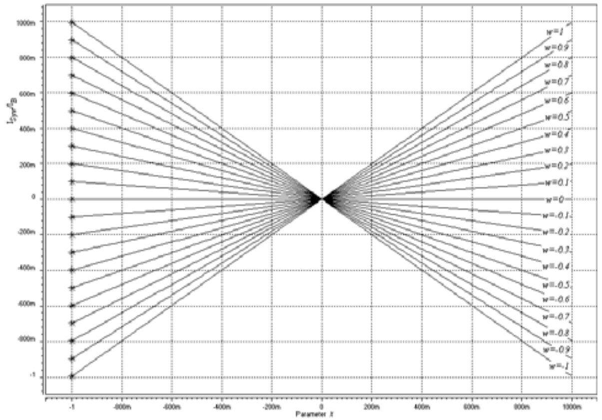


Analog four quadrant multiplier: the overall circuit (2)



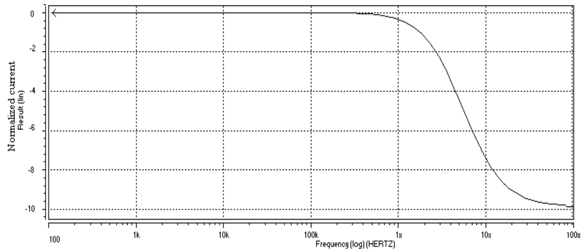
Results (simulations)

- Technology: AMS CMOS 0.8 μm
- Power supply voltage: ±1.25 V
- I_B = 250 nA



DC transfer characteristic: I_{Syn}/I_B=x×w vs x

Results (simulations)

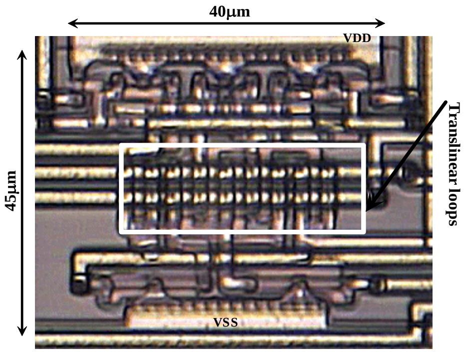


Simulated small signal frequency response of the multiplier: w=1, the ac input is applied to x. f_{-3dB} = 1 MHz

Total Harmonic Distortion (THD) is about 3%

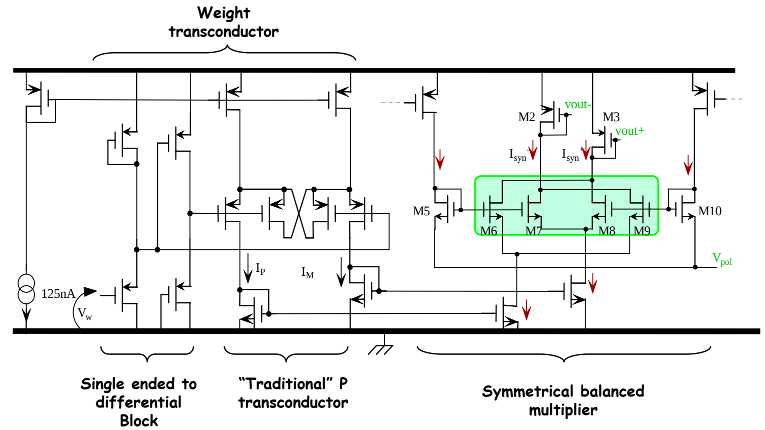
w was set to 1 (the bias current for I_W is I_B) and a sinusoid was applied to the x input: with $\omega = 2\pi\ 10^5$ rad/sec. This implies $I_G = I_{pp}\sin(\omega t)$ where I_{pp} = 100 nA and $\omega = 2\pi\ 10^5$ rad/sec.

Results (measurements)



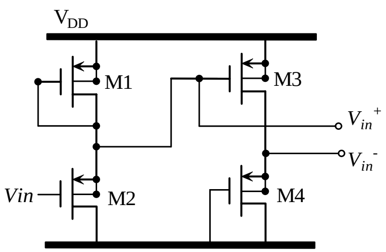
- Computational power: 2 MCPS/synapse
- Computational density: 133 MCPS/mm²
- Energy efficiency: 740 MCPS/mW

Results



Results (measurements)

Single ended to differential block

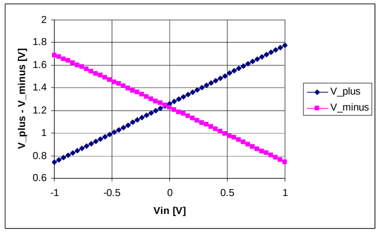


$$\frac{\mu C_{ox}}{2}\frac{W}{L}\left[\left(V_{DD}-V_{in}^{+}\right)-V_{TH}\right]^2 \equiv \frac{\mu C_{ox}}{2}\frac{W}{L}\left[\left(V_{in}^{+}-V_{in}\right)-V_{TH}\right]^2$$

$$V_{DD}-V_{in}^{+}\equiv V_{in}^{+}-V_{in} \qquad V_{in}^{+}\equiv \frac{V_{DD}}{2}+\frac{V_{in}}{2}$$

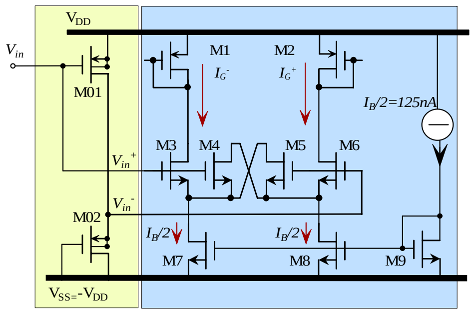
$$\frac{\mu C_{ox}}{2}\frac{W}{L}\left[\left(V_{DD}-V_{in}^{+}\right)-V_{TH}\right]^2 \equiv \frac{\mu C_{ox}}{2}\frac{W}{L}\left(V_{in}^{-}-V_{TH}\right)^2$$

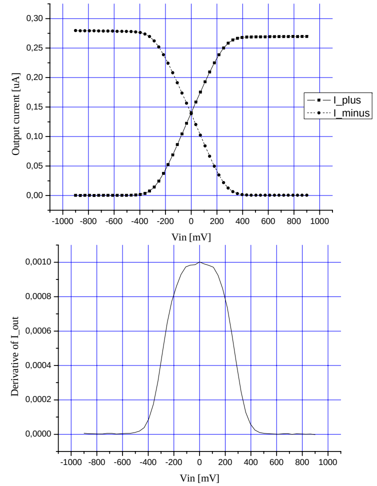
$$V_{in}^{-}\equiv \frac{V_{DD}}{2}-\frac{V_{in}}{2}$$



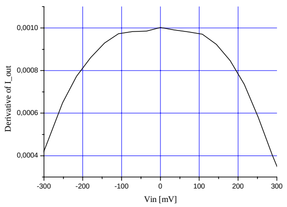
Results (measurements)

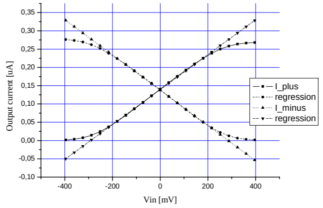
Input transconductor

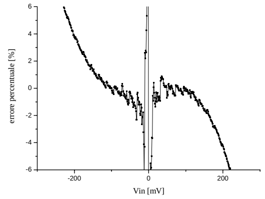




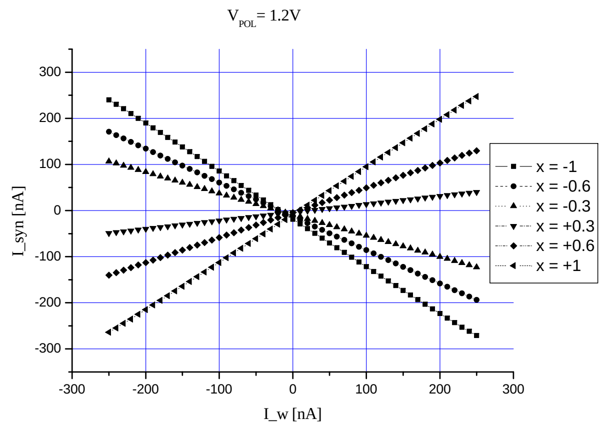
Results (measurements)



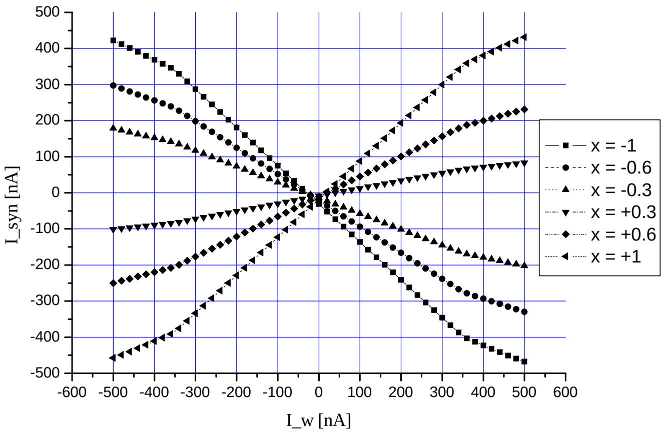




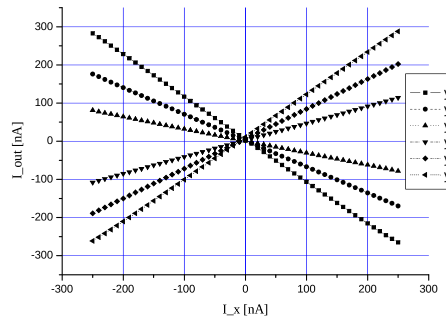
Results (measurements)

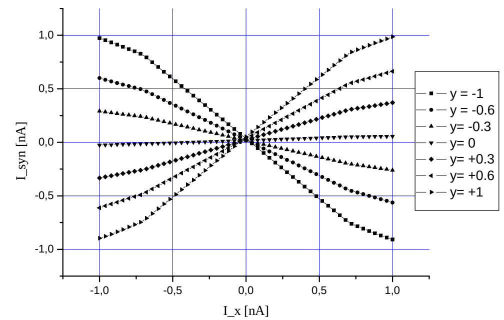


Results (measurements)

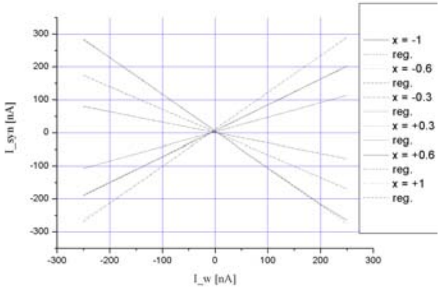


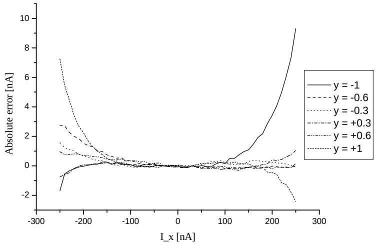
Results (measurements)





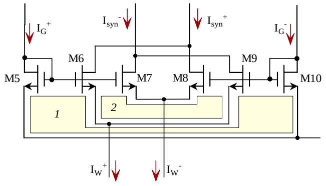
Results (measurements)





Mismatch modeling

$$I_{DS} = \frac{W}{L} I_{DC} \cdot e^{\frac{V_{GS}}{n\phi_t}} \left[1 - e^{-\frac{V_{DS}}{\phi_t}} \right] = \frac{W}{L} I_{DC} \cdot e^{\frac{V_{GS}}{n\phi_t}} \cdot \delta(V_{DS})$$



$$V_{GS\,i} = n\phi_t \ln\left(\frac{I_{DS_i}}{\left(W/L\right)_i I_{DC_i} \cdot \delta_i} \right)$$

$$V_{GS6} - V_{GS7} + V_{GS8} - V_{GS9} = 0$$

$$\ln\left(\frac{I_{DS5}}{\left(W/L\right)_5 I_{DC_5} \delta_5} \right) - \ln\left(\frac{I_{DS7}}{\left(W/L\right)_7 I_{DC_7} \delta_7} \right) + \ln\left(\frac{I_{DS8}}{\left(W/L\right)_8 I_{DC_8} \delta_8} \right) - \ln\left(\frac{I_{DS10}}{\left(W/L\right)_{10} I_{DC_{10}} \delta_{10}} \right) = 0$$

$$\frac{I_{DS5}}{\left(W/L\right)_5 I_{DC_5} \delta_5} \cdot \frac{I_{DS8}}{\left(W/L\right)_8 I_{DC_8} \delta_8} = \frac{I_{DS7}}{\left(W/L\right)_7 I_{DC_7} \delta_7} \cdot \frac{I_{DS10}}{\left(W/L\right)_{10} I_{DC_{10}} \delta_{10}}$$

Mismatch modeling

$$\gamma_1 = \frac{\left(W/L\right)_5 \left(W/L\right)_8}{\left(W/L\right)_7 \left(W/L\right)_9} \cdot \frac{I_{DC_5} I_{DC_8}}{I_{DC_7} I_{DC_9}} \cdot \frac{\delta_5 \delta_8}{\delta_7 \delta_9} \qquad I_{DS_5} \cdot I_{DS_8} = \gamma_1 I_{DS_7} \cdot I_{DS_9}$$

$$I_G^+ I_{M9} = \gamma_1 I_{M6} I_G^- \qquad \text{loop 1}$$

$$I_{M6} I_{M8} = \gamma_2 I_{M7} I_{M9} \qquad \text{loop 2}$$

$$I_G{}^+ = I_{M5},\ I_G{}^- = I_{M10} \qquad I_{M6} + I_{M9} = I_W{}^+,\ I_{M7} + I_{M8} = I_W{}^-$$

$$I_{SYN}^+ - I_{SYN}^- = I_W^+ + I_W^- - \frac{2 I_G^- I_W^+}{1/\gamma_1 I_G^+ + I_G^-} - \frac{2 I_G^+ I_W^-}{I_G^+ + \gamma_1 \gamma_2 I_G^-}$$

Mismatch modeling

$$I_{OUT} = I_{SYN}^+ - I_{SYN}^- = I_W^+ + I_W^- - \frac{2 I_G^- I_W^+}{I_B + \left(1/\gamma_1 - 1\right) I_G^+} - \frac{2 I_G^+ I_W^-}{I_B + \left(\gamma_1 \gamma_2 - 1\right) I_G^-}$$

$$I_{OUT} = I_{SYN}^+ - I_{SYN}^- = I_B - \frac{(1-x)(1+w) I_B}{2 + \left(1/\gamma_1 - 1\right)(1+x)} - \frac{(1+x)(1-w) I_B}{2 + \left(\gamma_1 \gamma_2 - 1\right)(1-x)}$$

$$I_{OUT} = I_{SYN}^+ - I_{SYN}^- = \left[1 - n(x) - n'(x) w \right] I_B$$

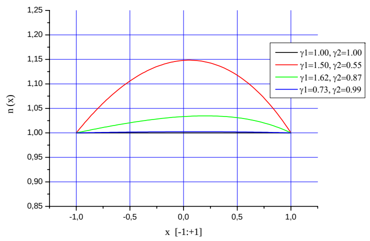
$$n_{\gamma_1,\gamma_2}(x) = \frac{(1-x)}{2 + \left(1/\gamma_1 - 1\right)(1+x)} + \frac{(1+x)}{2 + \left(\gamma_1 \gamma_2 - 1\right)(1-x)} \qquad n'_{\gamma_1,\gamma_2}(x) = \frac{(1-x)}{2 + \left(1/\gamma_1 - 1\right)(1+x)} - \frac{(1+x)}{2 + \left(\gamma_1 \gamma_2 - 1\right)(1-x)}$$

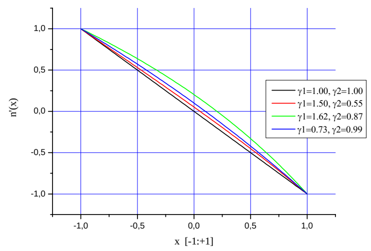
Mismatch modeling

	γ_1	γ_2
Chip n°1	1.5	0.55
Chip n°2	1.62	0.87
Chip n°3	1.23	0.99

$$\varepsilon = I'_{OUT} - I_{OUT} = [1 - n(x) - n'(x)w - xw]I_B$$

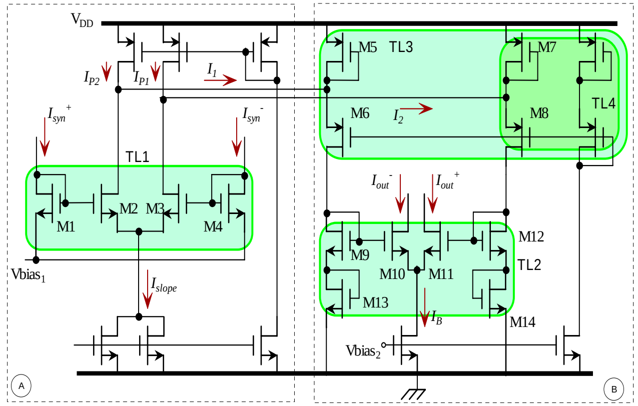
$$\varepsilon\% = \frac{I'_{OUT} - I_{OUT}}{I_{OUT}} = \frac{1 - n(x) - n'(x)w - xw}{xw}$$





The Neuron Circuit (block N) (1)

- The neuron transfer function slope control block (A)
- The neuron transfer function block (B)



The Neuron Circuit (block N) (1)

From the TL1 translinear loop equations

$$I_{M2} \cdot I_{Syn}^{-} = I_{M3} \cdot I_{Syn}^{+}$$
$$I_{M2} + I_{M3} = I_{slope} \equiv 2KI_B$$

We can obtain I_{M2} and I_{M3}, so:

$$I_1 = I_{P2} - I_{M2} = KI_B - I_{M2} = \frac{x \cdot w \cdot K \cdot I_B}{w-2}$$

$$I_2 = I_{P1} - I_{M3} = KI_B - I_{M3} = -\frac{x \cdot w \cdot K \cdot I_B}{w-2}$$

I_1 and I_2 are the single output current components in output from the neuron transfer function slope control block

The Neuron Circuit (block N) (3)

From the TL3 and TL4 translinear loops equations:

$$(I_{M6} - I_1) \cdot I_{M6} = I_B^2 \qquad (TL\ 3)$$
$$(I_{M8} - I_2) \cdot I_{M8} = I_B^2 \qquad (TL\ 4)$$

and From the TL2 translinear loop equations:

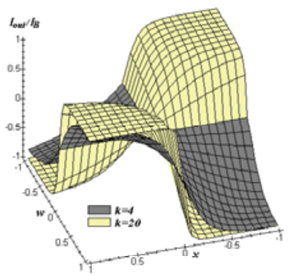
$$I_{M6}^2 \cdot I_{out^+} = I_{M8}^2 \cdot I_{out^-} \qquad (TL\ 2)$$
$$I_{out^+} + I_{out^-} = I_B \qquad (KCL)$$

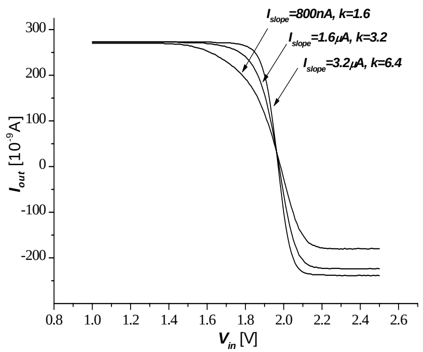
We can express I_{out}^{+} and I_{out}^{-} as function of the parameters *x* and *w* (and I_B too)

The Neuron Circuit (block N) (4)

$$I_{out} = I_{out}^{+} - I_{out}^{-} = -I_B \frac{\frac{kxw}{w-2}}{2 + \left(\frac{kxw}{w-2}\right)^2 \sqrt{\left(\frac{kxw}{w-2}\right)^2 + 4}}$$

The previous function exhibits an hyperbolic tangent shape as shown in the following figures:





Results (measurements)