Current Copier based Mixed-Signal Multiplier for Machine Learning Applications

Guide: Prof. Rajesh Zele

Mihir Kavishwar (17D070004)
Advanced Integrated Circuits and Systems Lab
Electrical Engineering, IIT Bombay

Outline

Introduction

- Most Machine Learning algorithms can be characterized as a sequence of Matrix Vector Multiplications (MVMs) interposed with non-linear transforms

Introduction

- Most Machine Learning algorithms can be characterized as a sequence of Matrix Vector Multiplications (MVMs) interposed with non-linear transforms
- **Multiply and Accumulate** (MAC) is the fundamental operation which gets repeated in MVMs: $y_i = \sum_{j=1}^{N} w_{ij}.x_j$

Introduction

- Most Machine Learning algorithms can be characterized as a sequence of Matrix Vector Multiplications (MVMs) interposed with non-linear transforms
- **Multiply and Accumulate** (MAC) is the fundamental operation which gets repeated in MVMs: $y_i = \sum_{j=1}^{N} w_{ij}.x_j$
- For **Edge Device** applications where we need to operate at very **low power** but still get **high performance**, Analog/Mixed-Signal implementations of MAC are proving to be much better alternative to the Digital counterpart

Introduction

- Most Machine Learning algorithms can be characterized as a sequence of Matrix Vector Multiplications (MVMs) interposed with non-linear transforms
- **Multiply and Accumulate** (MAC) is the fundamental operation which gets repeated in MVMs: $y_i = \sum_{j=1}^{N} w_{ij}.x_j$
- For **Edge Device** applications where we need to operate at very **low power** but still get **high performance**, Analog/Mixed-Signal implementations of MAC are proving to be much better alternative to the Digital counterpart
- In this work, I have explored a **Current-Copier based Mixed-Signal Multiplier** circuit in an attempt to come up with a novel implementation of MAC

System Architecture

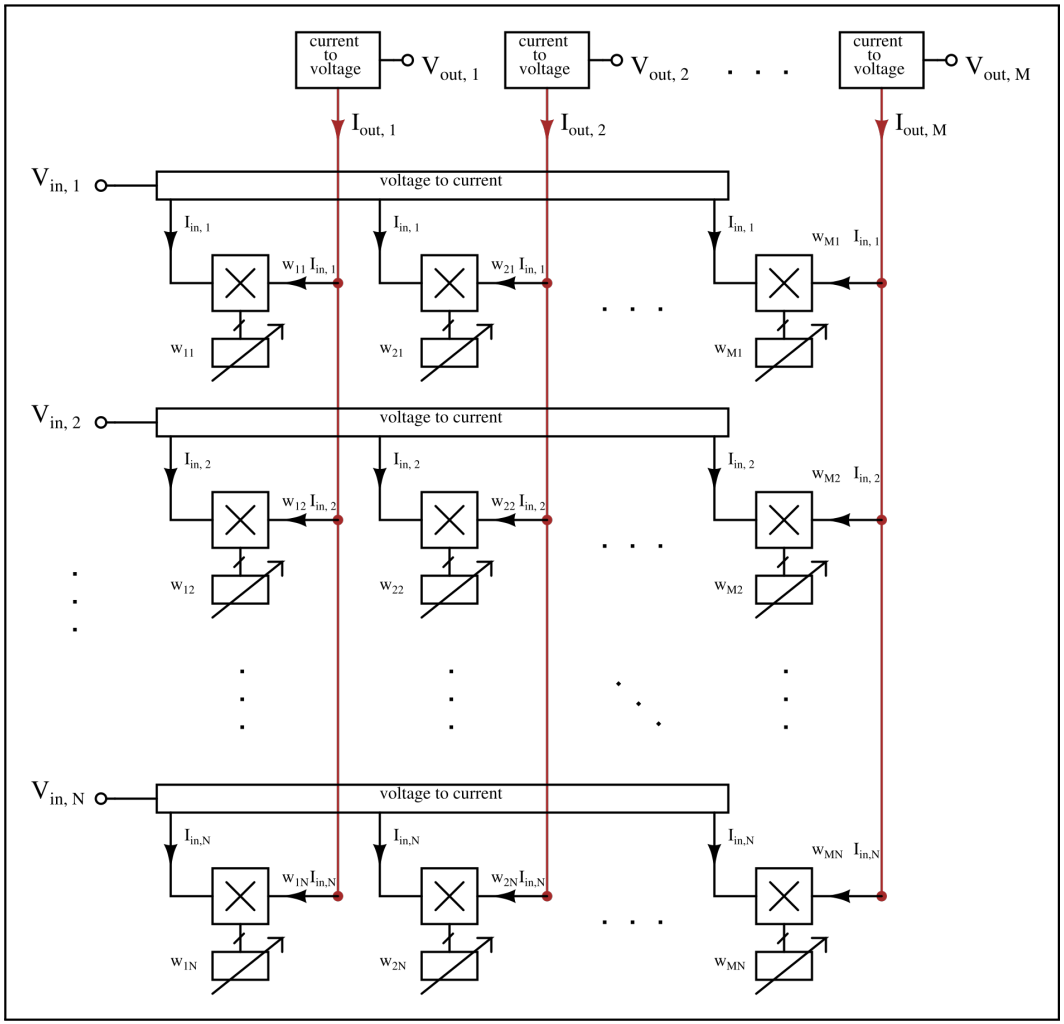


Figure: Mixed-Signal Matrix Vector Multiplier

$$\begin{bmatrix} I_{out},1 \\ I_{out},2 \\ \vdots \\ I_{out},M \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & \cdots & w_{MN} \end{bmatrix} \begin{bmatrix} I_{in},1 \\ I_{in},2 \\ \vdots \\ I_{in},N \end{bmatrix}$$

- Weights are stored digitally in local registers while the input and output currents are Analog
- Same multiplier cell is replicated several times
- Connecting multiplier output currents in parallel gives us the addition operation

Current Copier Principle

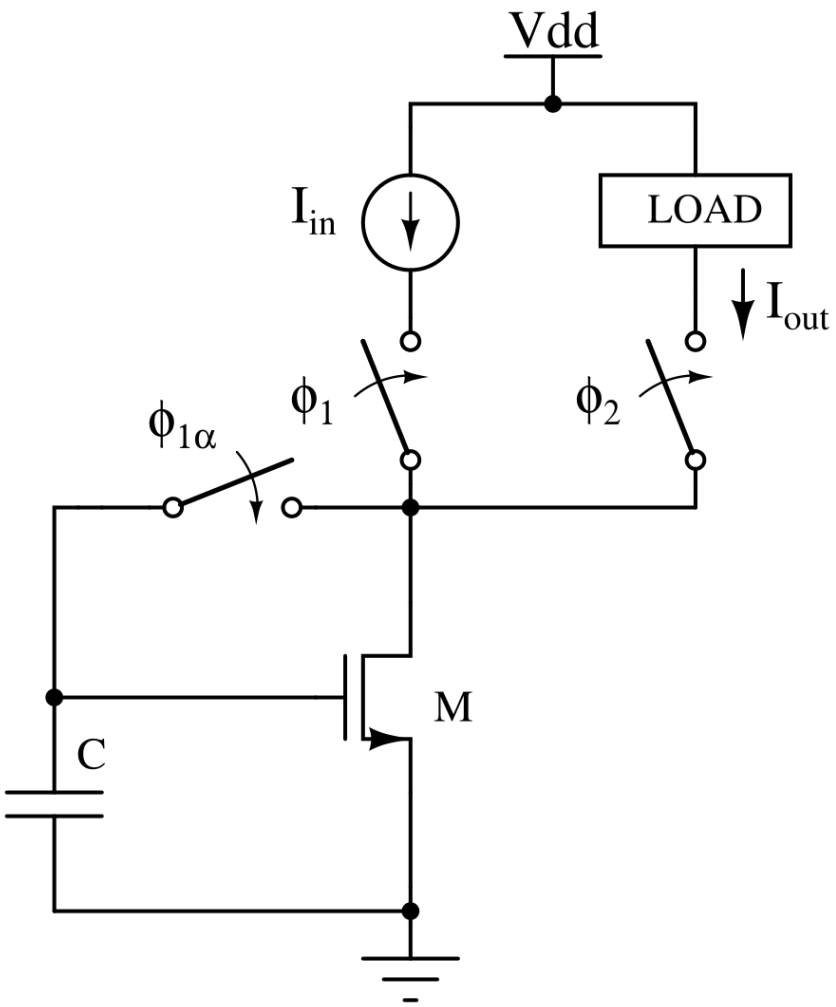


Figure: Basic Current Copier Cell

Phase ϕ_1

- M acts as an **input** device with it's gate and drain connected to input current source
- I_{in} charges C until V_{gs} reaches a value which corresponds to $I_{in} = I_{ds}$

Current Copier Principle

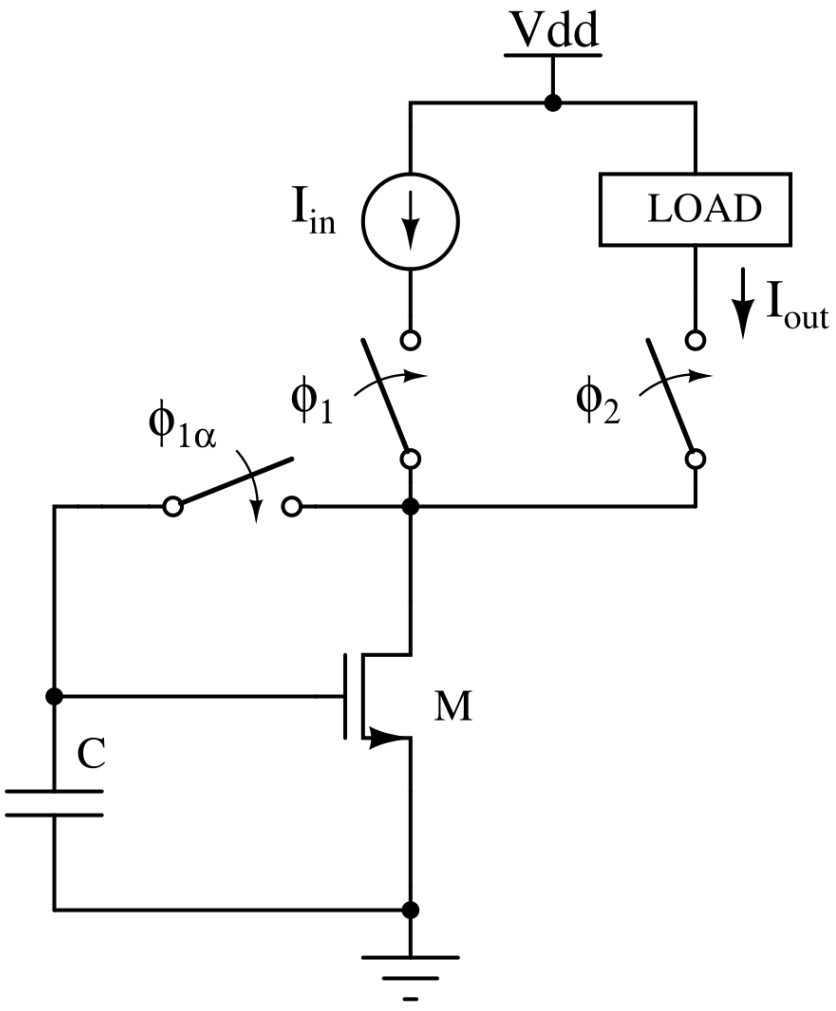


Figure: Basic Current Copier Cell

Phase ϕ_1

- M acts as an **input** device with it's gate and drain connected to input current source
- I_{in} charges C until V_{gs} reaches a value which corresponds to $I_{in} = I_{ds}$

Phase ϕ_2

- M acts as an **output** device with it's drain disconnected from gate and connected to output node
- M sinks I_{out} which is controlled by same V_{gs} and thus equal to I_{in}

Current Copier Based Multiplier

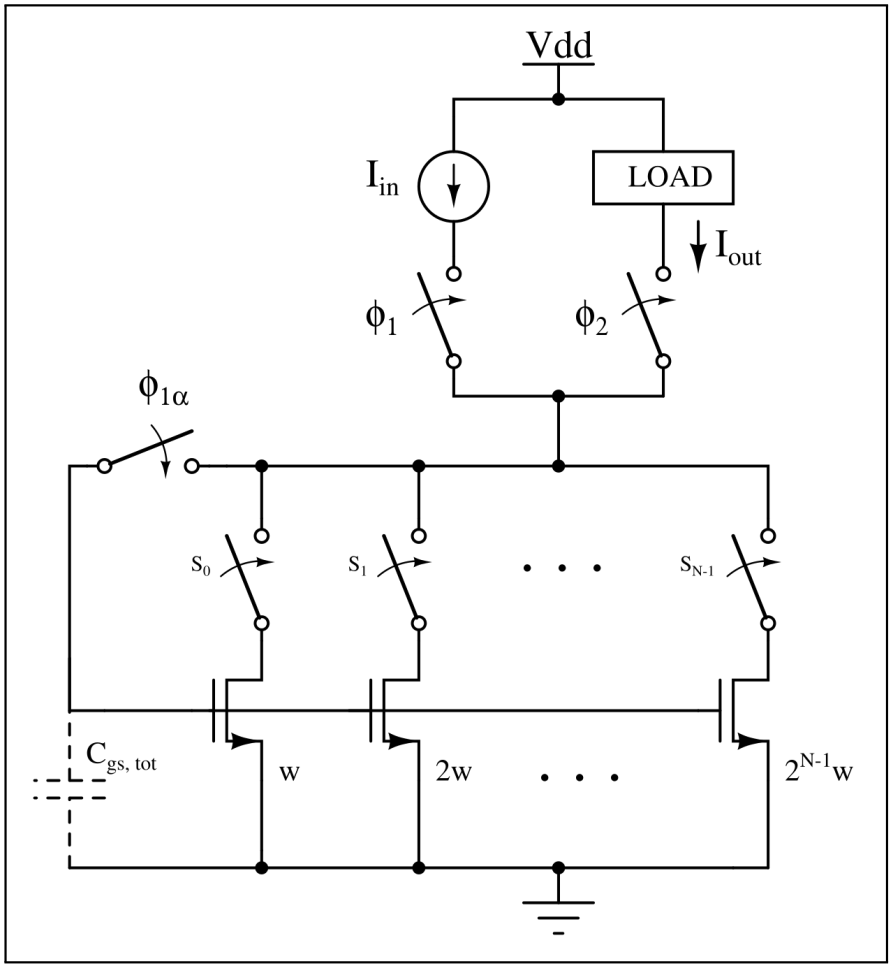


Figure: Multiplier Cell

Assuming V_{gs} remains constant across the 2 phases,

$$I_{out,\phi_2} = w \times I_{in,\phi_1}$$

where,

$$w := \frac{2^0.S_{0,\phi_2} + 2^1.S_{1,\phi_2} + \cdots + 2^{N-1}.S_{N-1,\phi_2}}{2^0.S_{0,\phi_1} + 2^1.S_{1,\phi_1} + \cdots + 2^{N-1}.S_{N-1,\phi_1}}$$

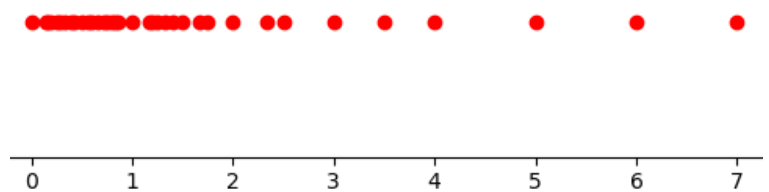


Figure: Possible values of w for $N = 3$

Current Copier Based Multiplier - Alternate Architecture

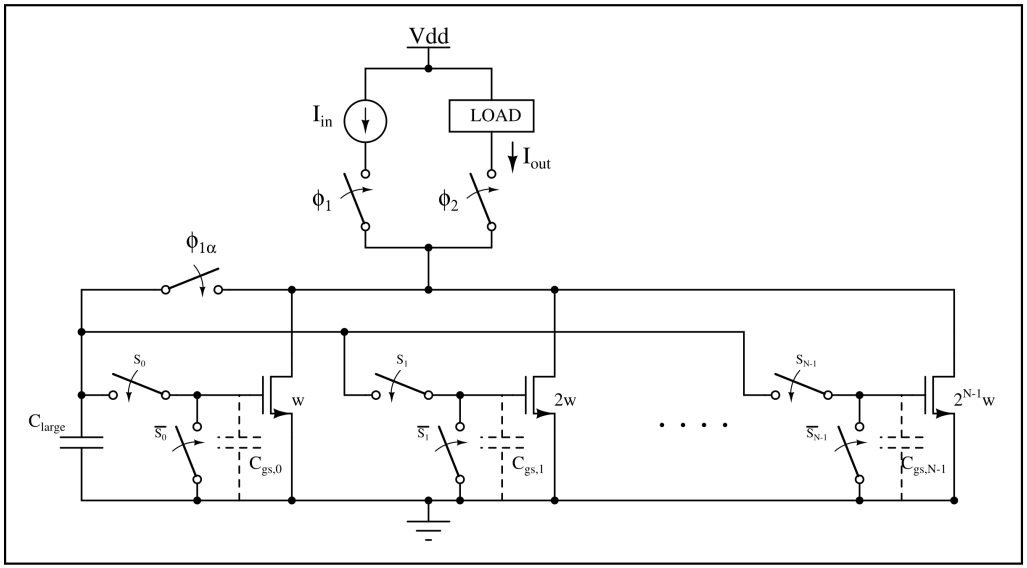
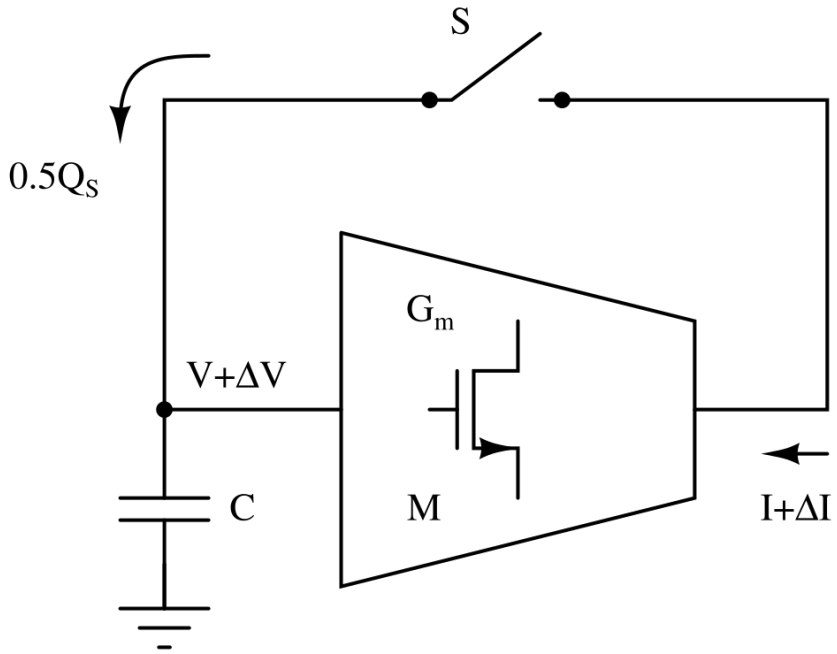


Figure: Alternate architecture of Multiplier Cell

- While avoiding switches in current path, this architecture brings other problems
- The net value of gate capacitor depends heavily on the configuration of switches
- A large external capacitor needs to be connected at the gate which results in more charging time and hence lesser speed

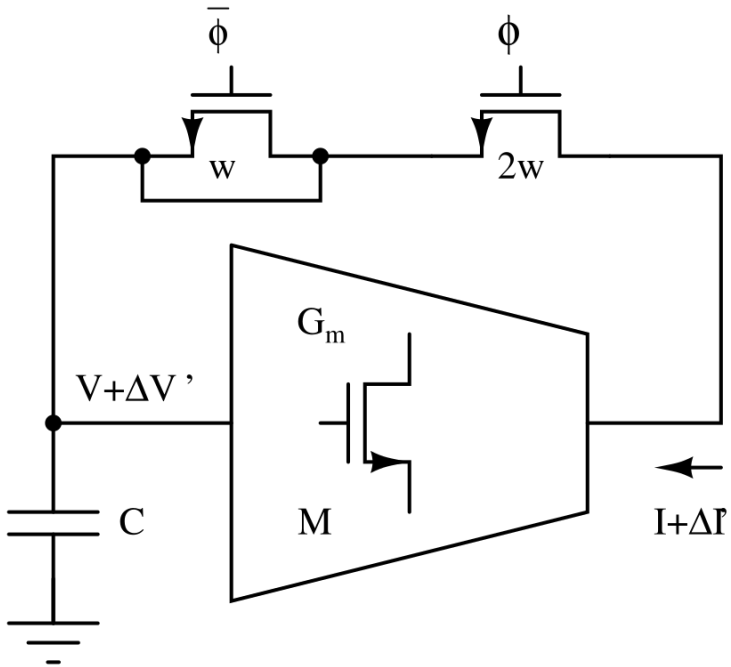
Sources of Error

- Gate Switch Charge Injection



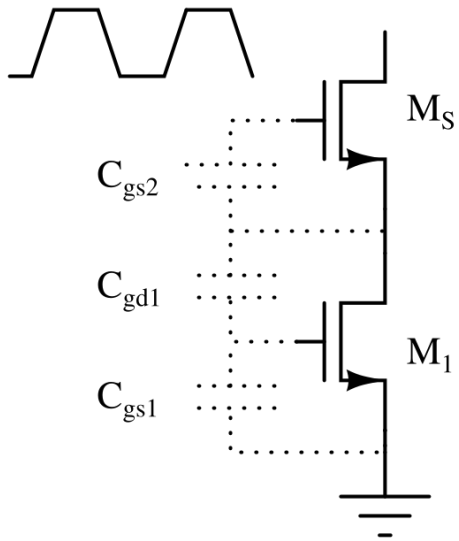
Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation



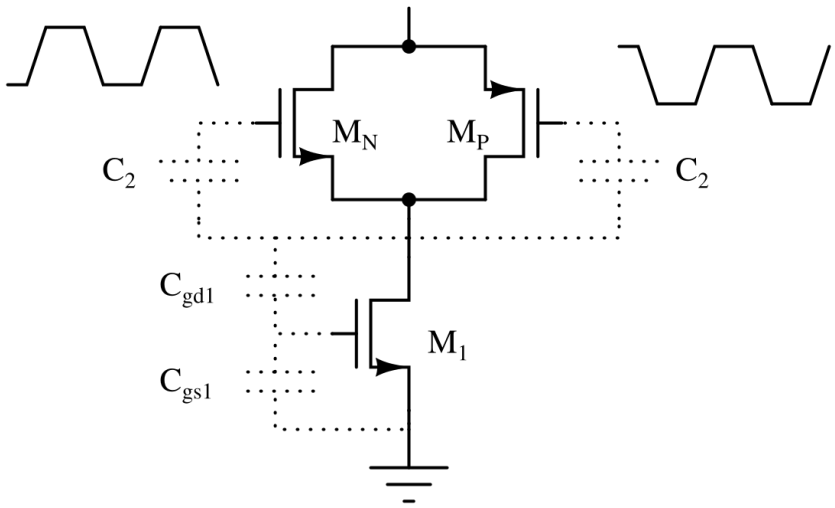
Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough



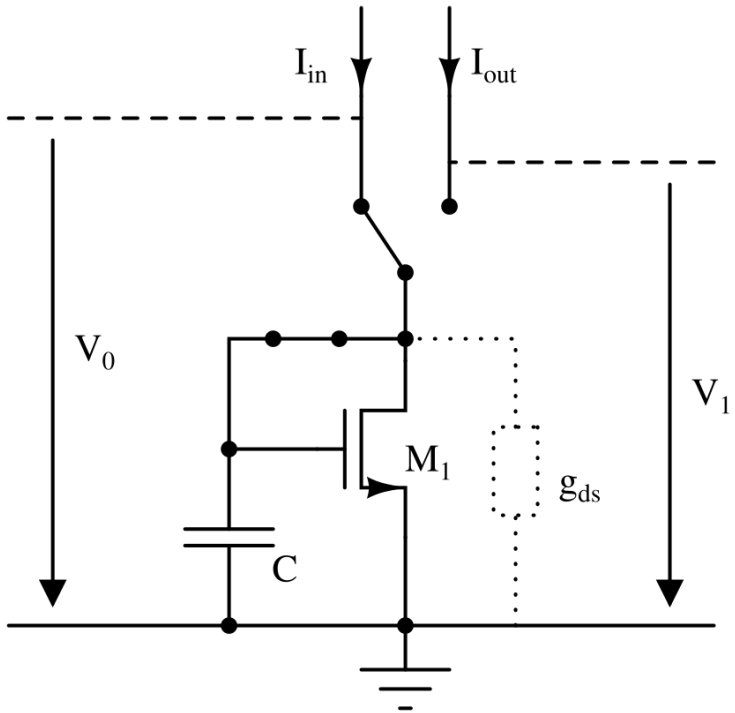
Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough
  - Complementary switches



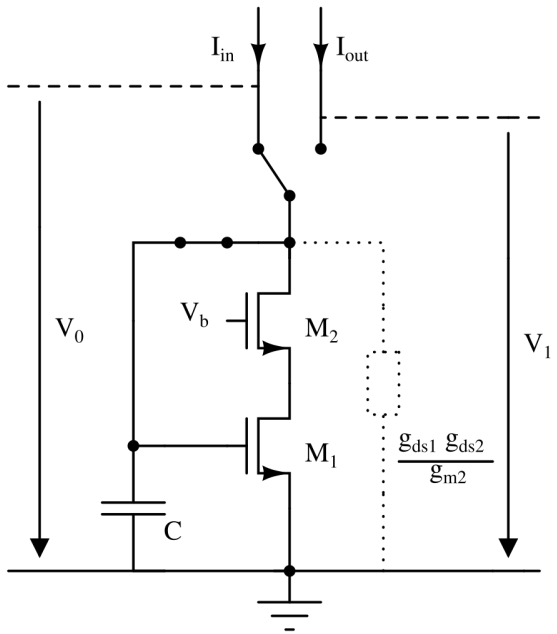
Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough
  - Complementary switches
- Channel Length Modulation



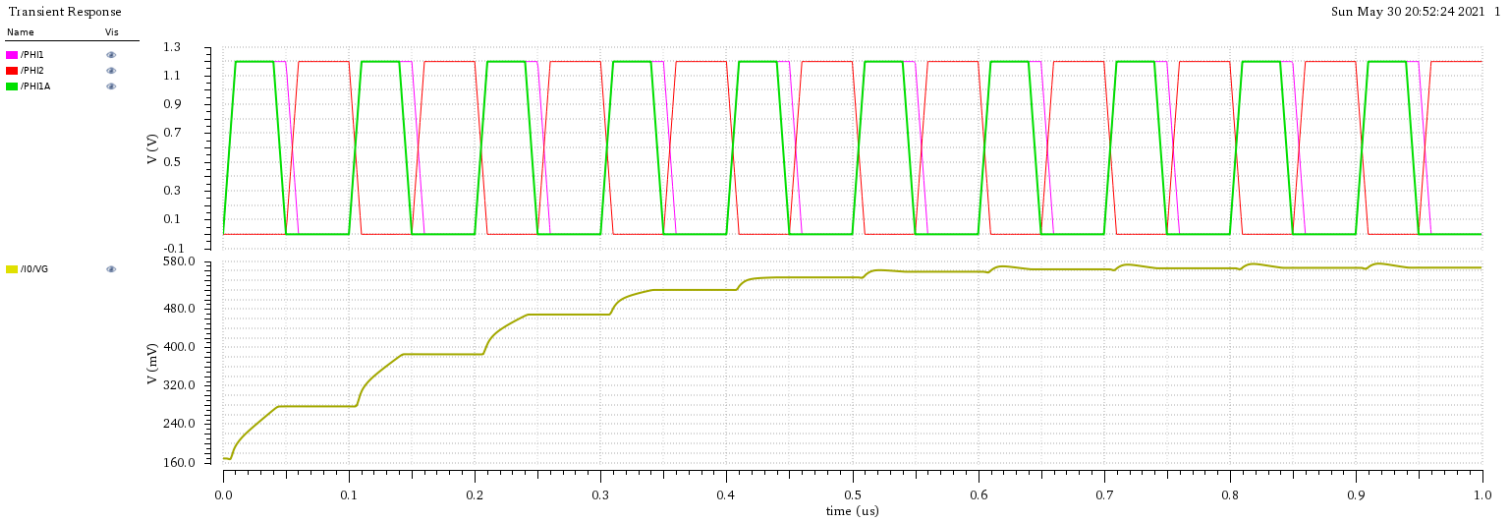
Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough
  - Complementary switches
- Channel Length Modulation
  - Cascode configuration and larger channel length



Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough
  - Complementary switches
- Channel Length Modulation
  - Cascode configuration and larger channel length
- Incomplete settling of Gate Voltage

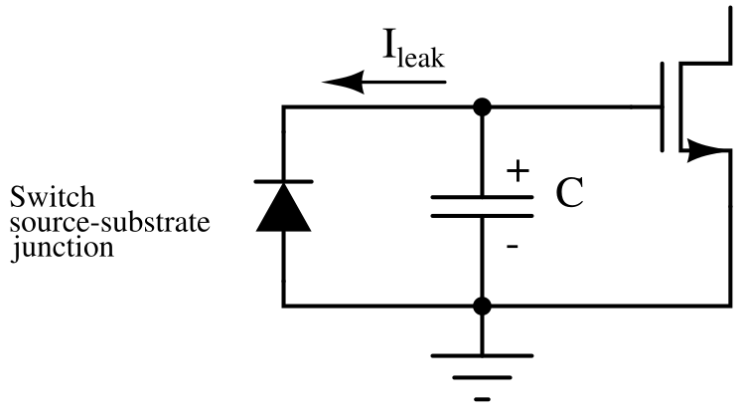


Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough
  - Complementary switches
- Channel Length Modulation
  - Cascode configuration and larger channel length
- Incomplete settling of Gate Voltage
  - Either increase input current range or decrease operating frequency

Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough
  - Complementary switches
- Channel Length Modulation
  - Cascode configuration and larger channel length
- Incomplete settling of Gate Voltage
  - Either increase input current range or decrease operating frequency
- Junction Leakage



Sources of Error

- Gate Switch Charge Injection
  - Dummy switch compensation
- Drain Switch Charge Feedthrough
  - Complementary switches
- Channel Length Modulation
  - Cascode configuration and larger channel length
- Incomplete settling of Gate Voltage
  - Either increase input current range or decrease operating frequency
- Junction Leakage
  - Don't leaving gate capacitor floating for a long time

Schematics

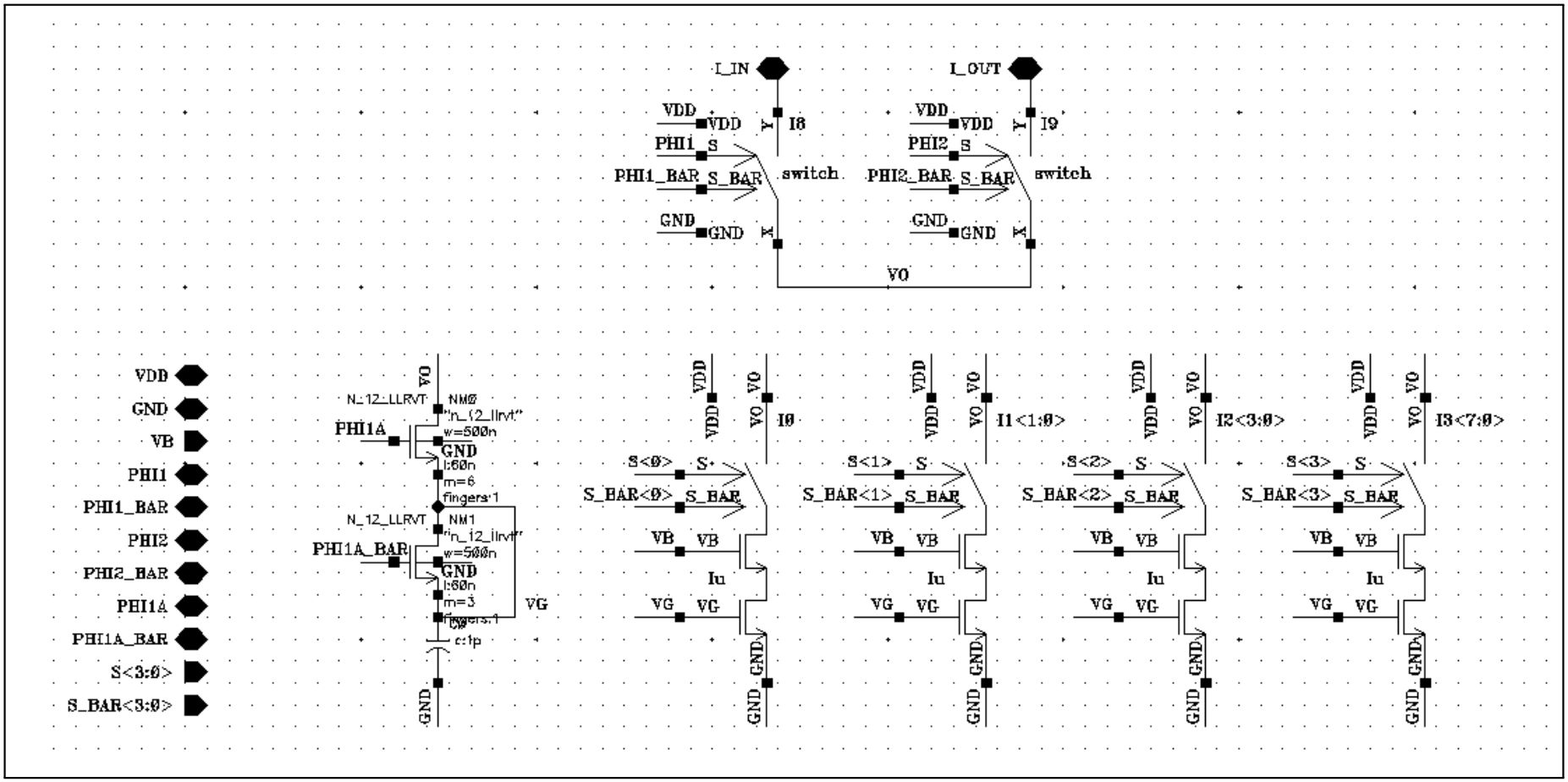


Figure: Final schematic of 4-bit Current Copier Multiplier

Schematics

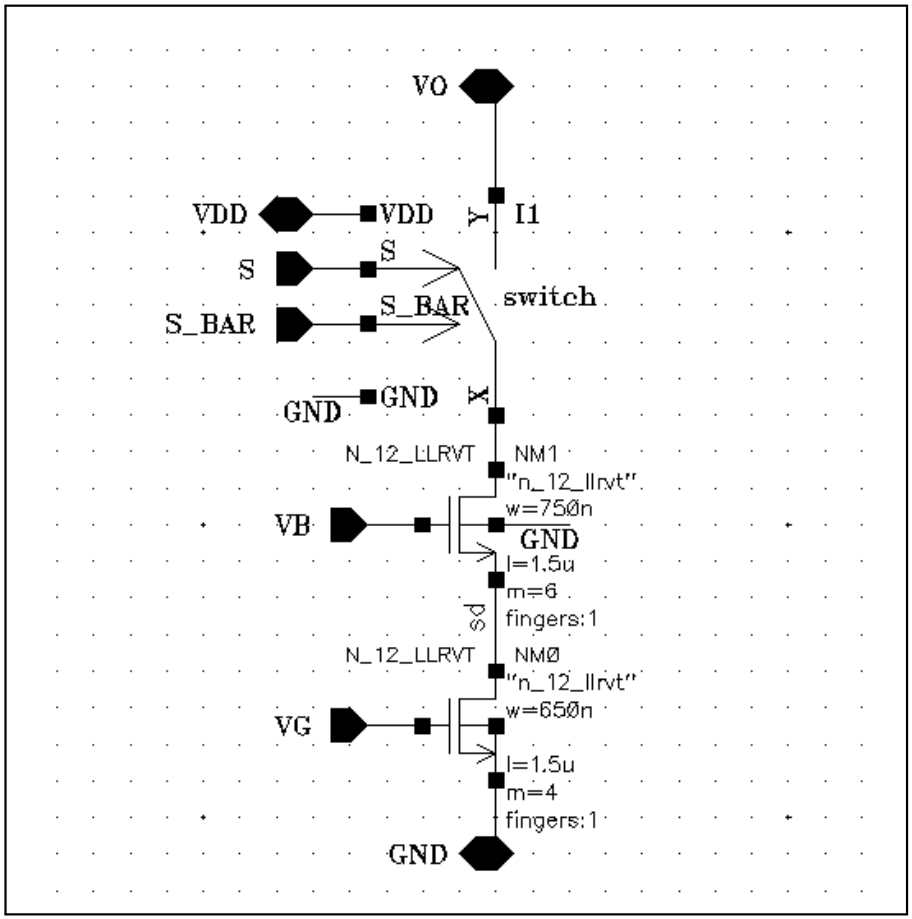


Figure: Unit Cell

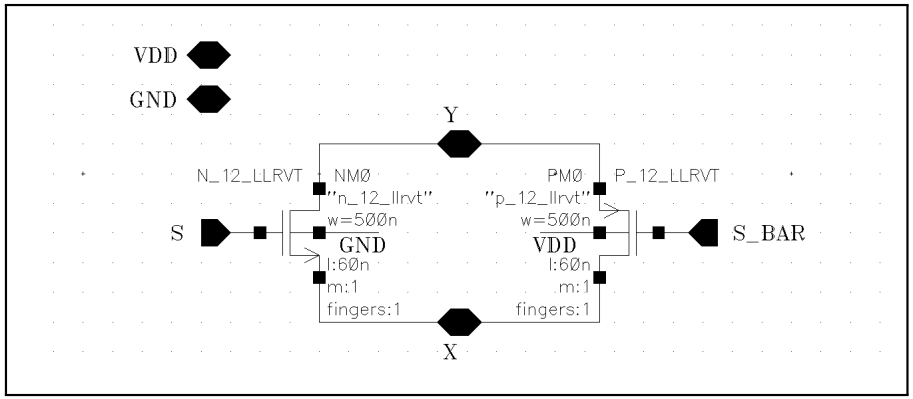


Figure: Complementary Switch

Testbench

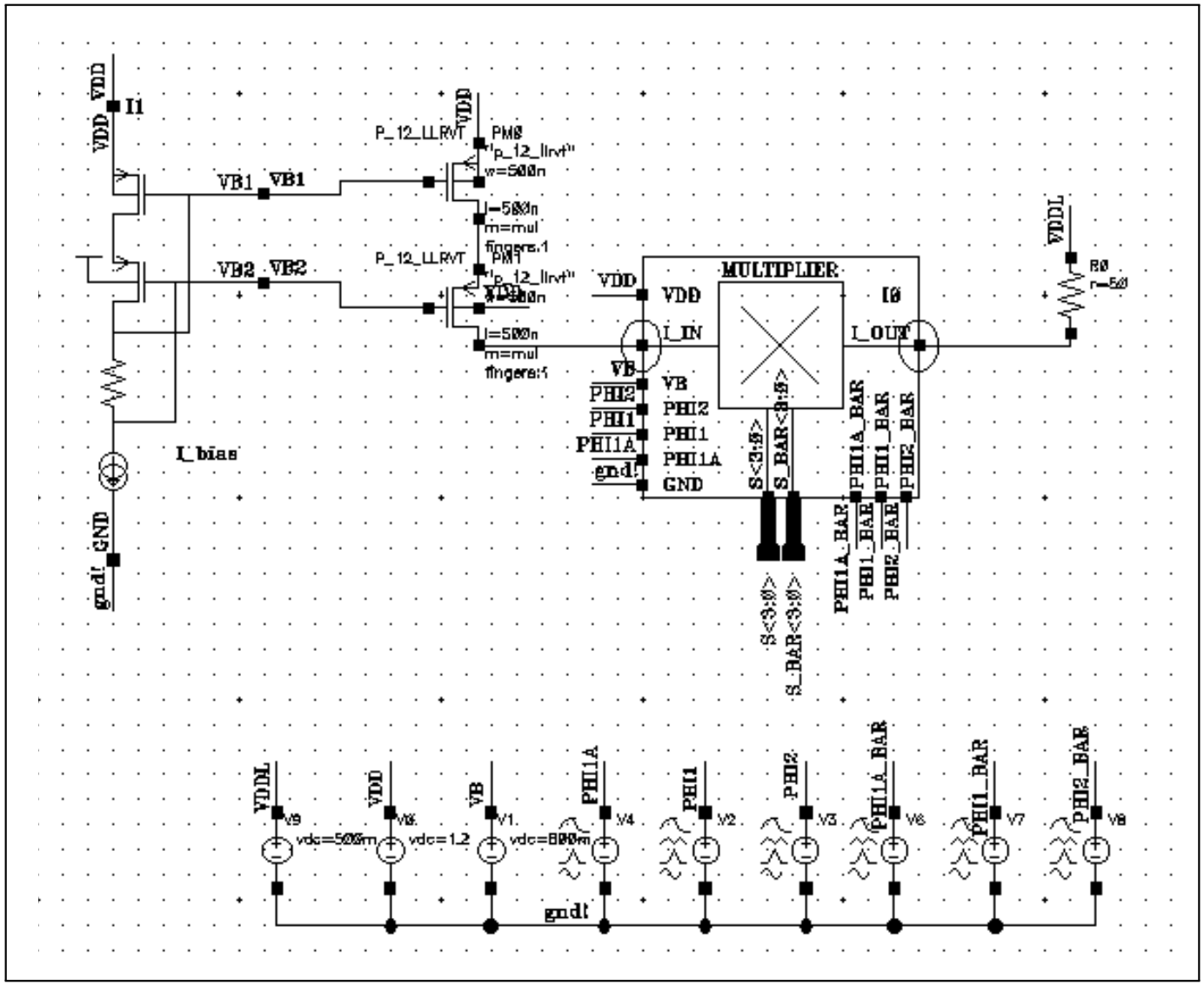


Figure: 4-bit Multiplier Testbench

Simulation Results

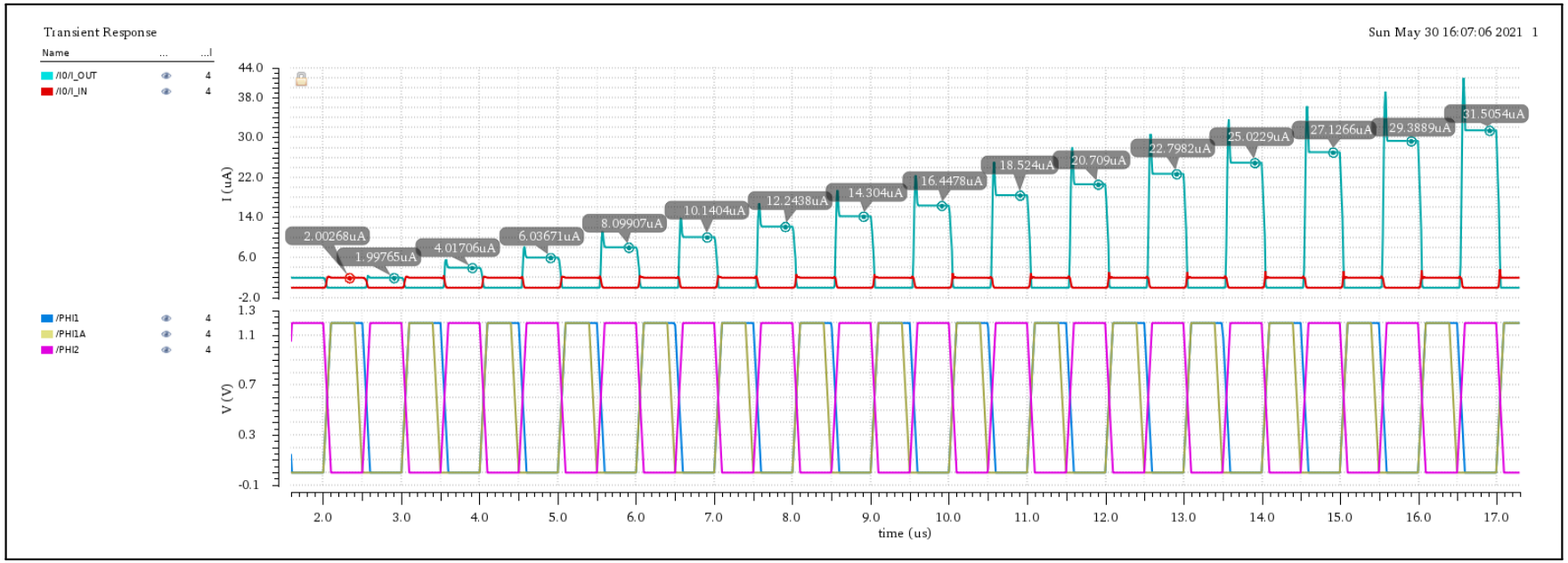


Figure: Transient Analysis with w varying from 1 to 15

Simulation Results

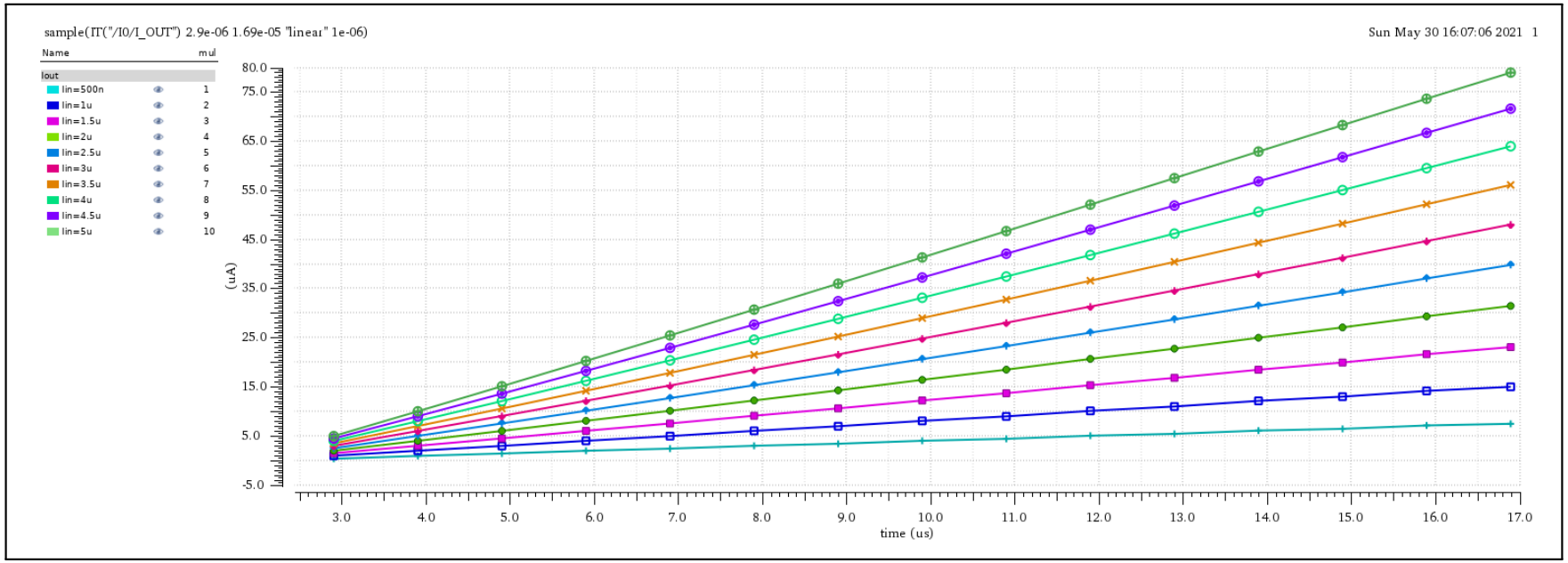


Figure: I_{out} value in phase ϕ_2 with w varying from 1 to 15 for different I_{in}

Simulation Results

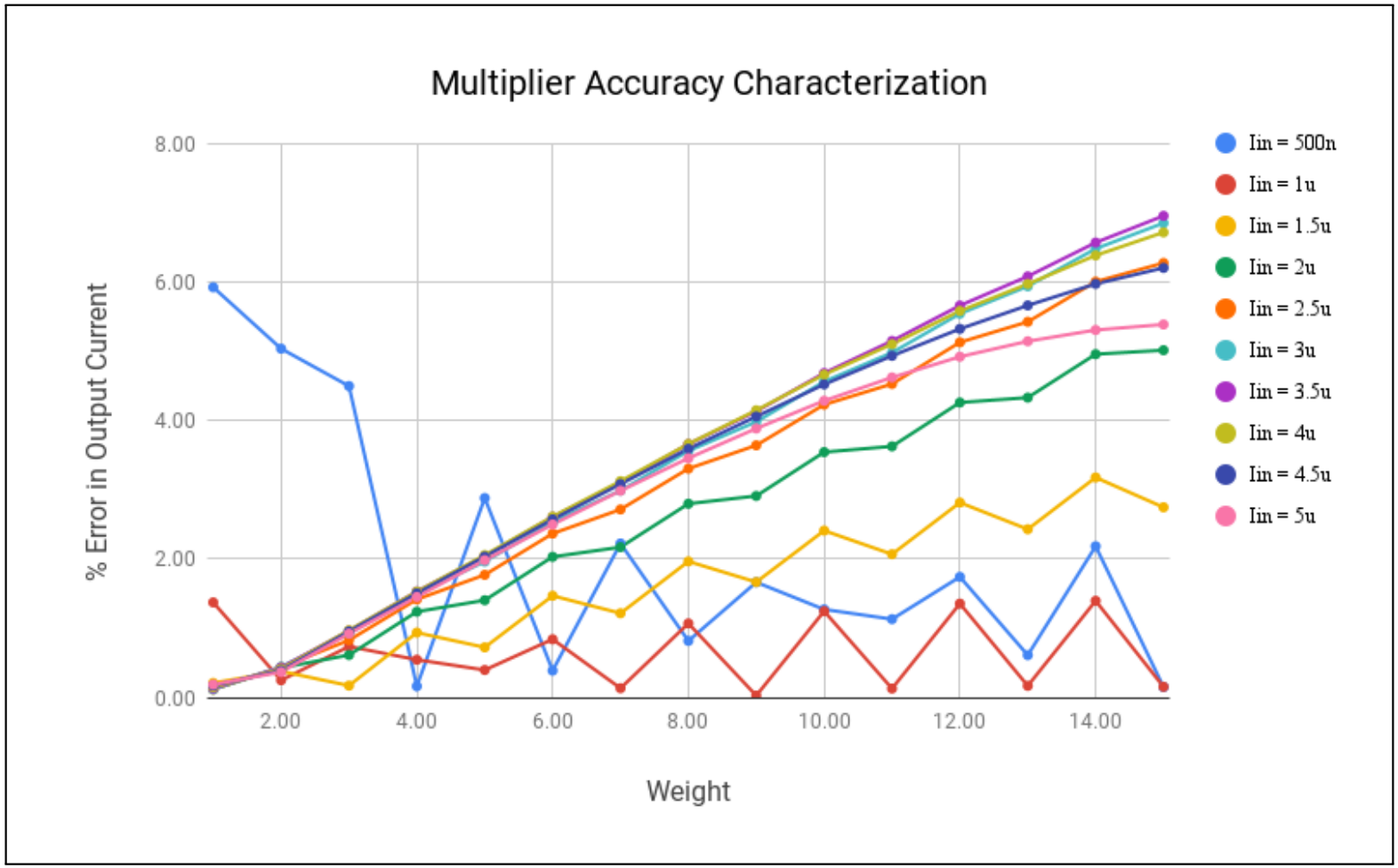


Figure: Error Analysis

Simulation Results

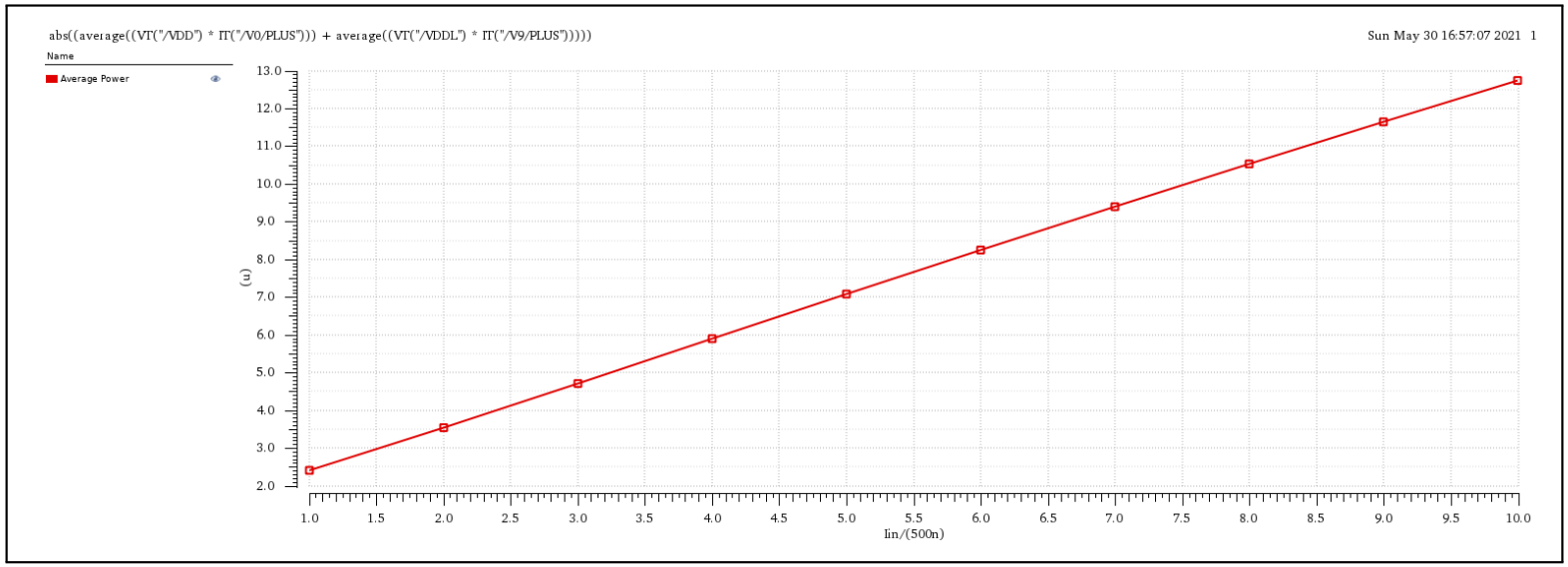


Figure: Average Power vs Input Current

For input current range of 500nA to 5μA, average power consumed by multiplier plus biasing circuit is **2.5 μW** to **13 μW**

Simulation Results

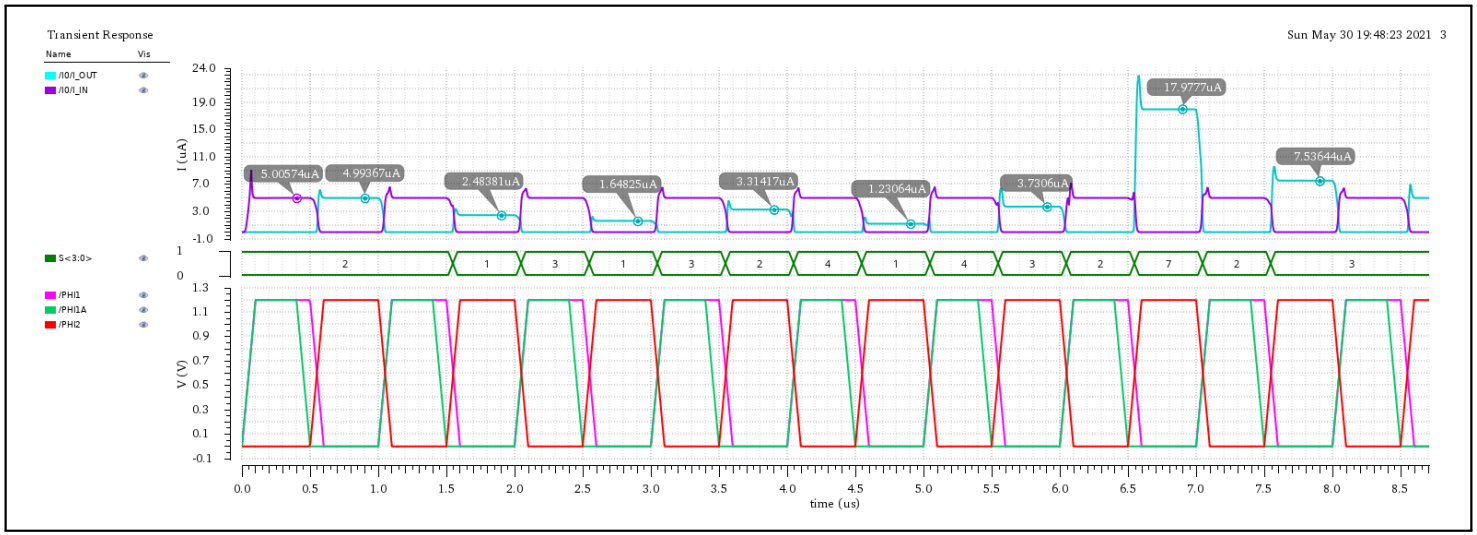


Figure: Transient Analysis with fractional w

Weight	$I_{out,actual}$ (μA)	$I_{out,ideal}$ (μA)	**% Error**
$2/2 = 1$	4.994	5	0.12
$1/2 = 0.5$	2.484	2.5	0.64
$1/3 = 0.333$	1.648	1.667	1.14
$2/3 = 0.667$	3.314	3.333	0.58
$1/4 = 0.25$	1.231	1.25	1.52
$3/4 = 0.75$	3.731	3.75	0.51
$7/2 = 3.5$	17.978	17.5	2.73

Table: Error Analysis with fractional w for $I_{in} = 5\ \mu A$

MAC Testbench

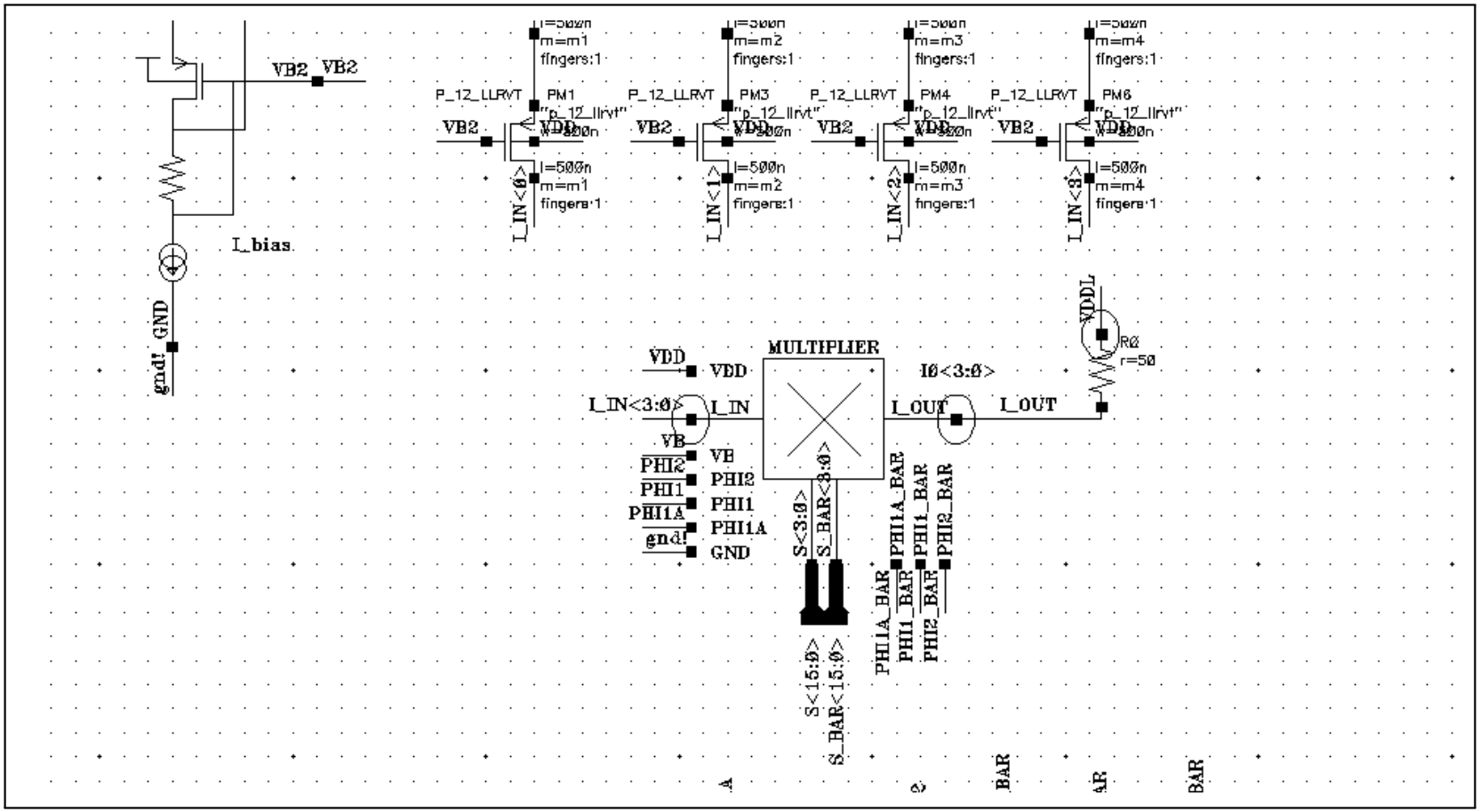


Figure: 4 Multipliers connected in parallel for testing MAC operation

Simulation Results

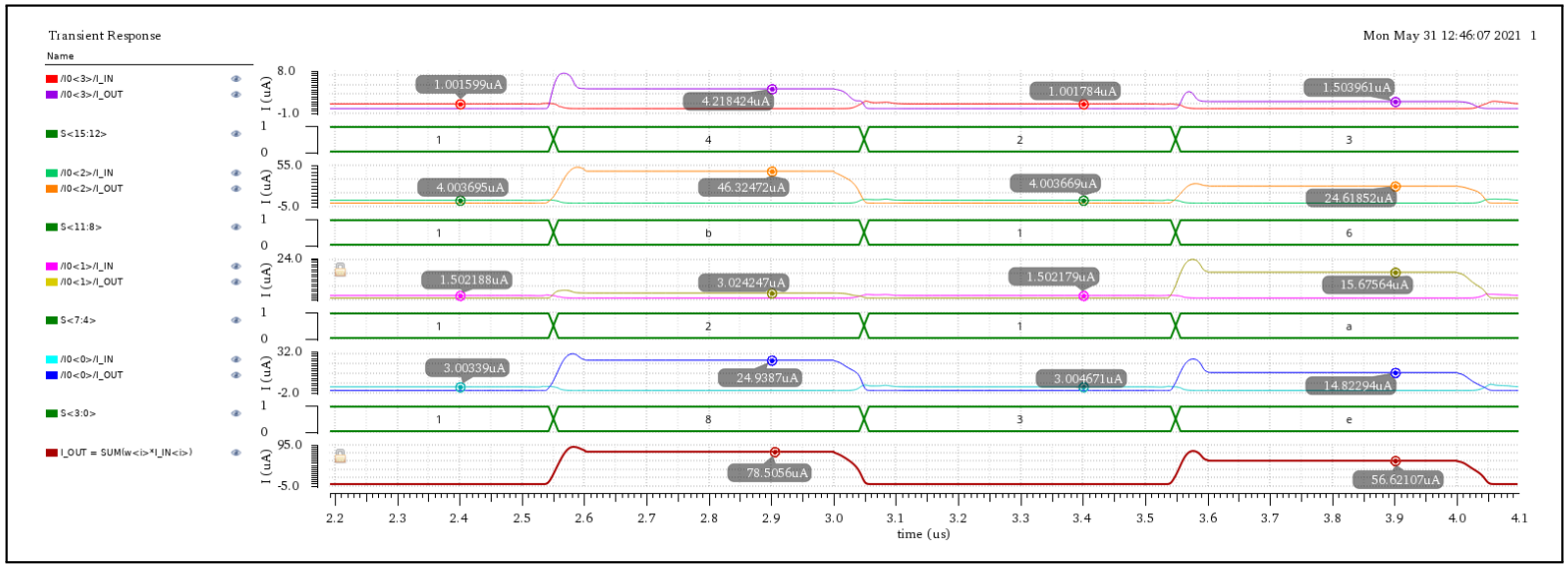


Figure: MAC Transient Simulation

w_1	$I_{in,1}$ (μA)	w_2	$I_{in,2}$ (μA)	w_3	$I_{in,3}$ (μA)	w_4	$I_{in,4}$ (μA)	$I_{out,actual}$ (μA)	$I_{out,ideal}$ (μA)	% Error
8	3	2	1.5	11	4	4	1	78.506	75	4.675
4.667	3	10	1.5	6	4	1.5	1	56.621	54.5	3.892

Table: MAC Error Analysis

Conclusion

- A new architecture of Mixed-Signal Multiplier based on Current-Copier principle was explored
- The 4-bit Multiplier operating at 1 MHz consumes < 13 μW power and has $< 7\%$ output error for input current range of 500 nA - 5 μA
- Achieving high resolution is difficult with this architecture due to input dependent charge feedthrough from drain switches
- Future work includes:
  - Exploring ways to increase multiplier speed, resolution of weights and support for negative weights
  - Incorporating transistor level design in place of behavioural models in the image classifier system
  - Designing storage elements and circuits for non-linear transforms

References

- S.J. Daubert, D. Vallancourt. and Y.P. Tsividis, "**Current Copier Cells**", Electronics Letters, voL 24, no. 25, pp. 1560-1562, 8 Dec. 1988.
- D. Vallancourt, Y. P. Tsividis and S. J. Daubert, "**Sampled-current circuits**," IEEE International Symposium on Circuits and Systems,, 1989, pp. 1592-1595 vol.3, doi: 10.1109/ISCAS.1989.100665.
- W. Groeneveld, H. Schouwenaars and H. Termeer, "**A self calibration technique for monolithic high-resolution D/A converters**," IEEE International Solid-State Circuits Conference, 1989 ISSCC. Digest of Technical Papers, 1989, pp. 22-23, doi: 10.1109/ISSCC.1989.48217.