A FULLY ANALOG NEURAL NETWORK FOR ON-CHIP INFERENCE

STUDENTS: JOHN LOUIS ADOLFO, EDWARD DUONG, SAM DECKER, NEHA KRISHNAKUMAR, WILLIAM HENRY, JEREMY CHEN

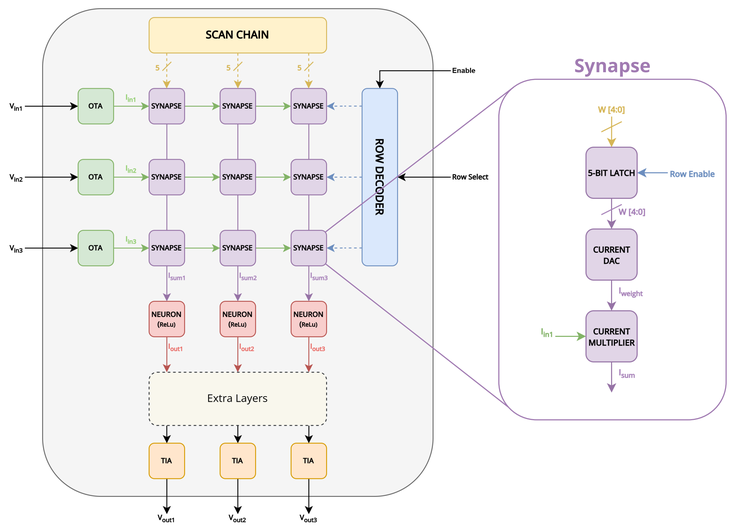
Introduction

- Modern neural networks are implemented digitally, requiring real-world signals to be converted before processing, introducing latency, power consumption, and hardware overhead.
- Analog neural networks offer a path forward by processing signals directly in the physical domain, eliminating the need for conversion entirely.
- This project presents a fully connected, feedforward analog neural network ASIC in TSMC 180nm
- Computations are performed in the current domain with programmable weights that can be configured through off-chip training

Targeted Specifications

Specifications	Value
Signal Bandwidth	>1 MHz
Network Size	[16, 16, 8, 4]
Weight Resolution	5-bit signed
Circuit Area	< 2 mm^2

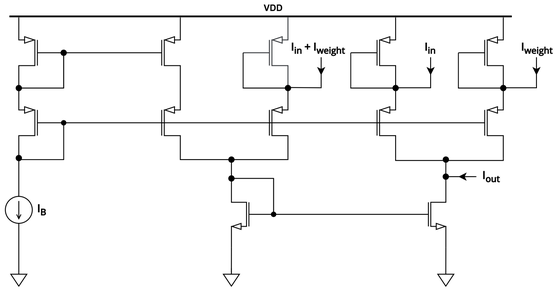
System Architecture



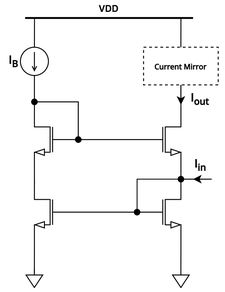
*Note: Network shown at reduced scale for readability

Circuit Implementation

Current Multiplier



Neuron

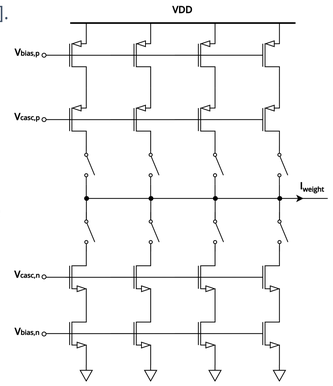


- **Neuron** takes the summed current of several synapses and rectifies it to output only a positive current (ReLu) [1].
- **Current Multiplier** takes two bidirectional currents and their sum to perform a multiplication with the following relationship [2].

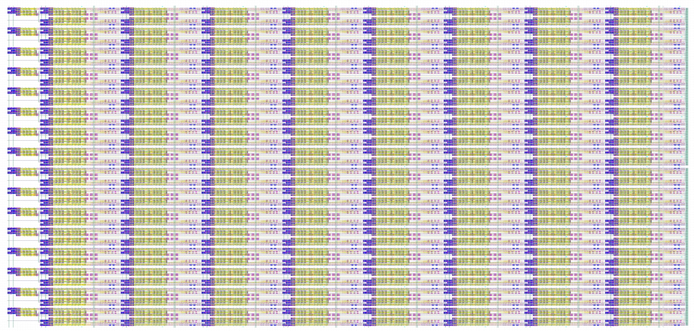
$$I_{out} = \frac{I_{in} \cdot I_{weight}}{8I_B}$$

- Weights are implemented via a **Current DAC** capable of both sinking and sourcing current, with a programmable range of −15 µA to +15 µA.
- Weight bits are serially loaded onto the chip through a scan chain, with a row decoder selecting which row of synapses to update

Current DAC

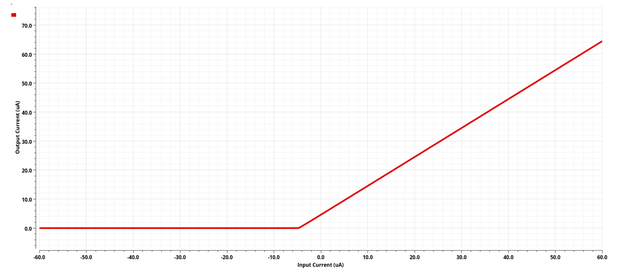


Preliminary Layout

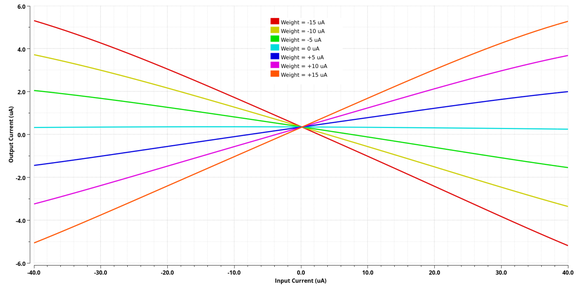


Results

Neuron



Synapse



Input Range (I_B = 15 µA)	-40 µA to 40 µA
Output Swing (I_B = 15 µA)	-5 µA to 5 µA
Bandwidth	3 MHz
Static Power Consumption (V_{DD} = 1.8V)	135 µW

Future Work, References, and Acknowledgments

- Tapeout in TSMC 180nm CMOS in later in June
- Design test PCB for interfacing with chip
- Develop a hardware-accurate software model to train and configure weights directly on the chip for testing and verification

Faculty: Jacques C. Rudell, UW ECE
Teaching Assistants: Ahmed R. Aboulsaad, Elpida Karapepera
Industry Mentors: Frank O'Mahony, Apple

[1] S. Khucharoensin and V. Kasemsuwan, "High performance CMOS current-mode precision full-wave rectifier (PFWR)," Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03., Bangkok, Thailand, 2003.
[2] K. Tanno, O. Ishizuka and Zheng Tang, "Four-quadrant CMOS current-mode multiplier independent of device parameters," in IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol. 47, no. 5, pp. 473-477, May 2000.

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

ADVISERS: Jacques C. Rudell

SPONSORS: TSMC, Apple Inc.