# Low Temperature Performance Of High-Speed Neural Network Circuits

Tuan Duong, Mua Tran, Taher Daud, and Anil Thakoor
Center for Space Microelectronics Technology
Jet Propulsion laboratory, California Institute of Technology
Pasadena, California 91109

## Introduction:

Artificial neural networks, derived from their biological counterparts, offer a new and enabling computing paradigm specially suitable for such tasks as image and signal processing with feature cl ossification/object recognition, global opti mizati on, and adaptive control[1 ]. When implemented in fully parallel electronic hardware, it offers orders of magnitude speed advantage[2]. Basic building blocks of the ncw architecture arc the processing elements called neurons implemented as nonlinear operational amplifiers with sigmoidal transfer function, interconnected through weighted connections called synapses implemented using circuitry for weight storage and multiply functions either in an analog, digital, or hybrid scheme[3]. An important architectural aspect of the biological system is its full integration between the various biological sensors and the processing wetware which establishes a highly parallel, direct, and therefore, very effi cient communi cation Ii nks amongst them. To enhance the overall effectiveness of the neural network hardware, therefore, an innovative architecture for a tight coupling between a sensor and the processing chips has been conceptualized which, among other benefits, offers the high throughput by omitting the 1/0 bottlenecks,

## Innovative Architecture:

A technology opportunity that lends itself to artificial vision related applications utilizing the fully parallel sensor-processor interface is the 3-D Artificial Neural Network (3 DANN). 3DANN would consist of a stack of thinned VLSI neural network chips directly mated to an infrared focal plane array (IR FPA). Offering high density and massively parallel "focal plane", this smart "sugar cube" could be used for a variety of applications such as a smart "eyeball" for industrial robotics and autonomous space exploration. Another similar challenging application requiring extremely intensive computation and high speed processing is the automatic target recognition for BMDO's fast frame seeker function of an interceptor with its real time image acquisition, recognition, tracking, and homing requirements for the

ballistic missile defense during a total mission time of merely a few seconds.

An integrated system such as this must not only operate at high speed, but also be fully functional at low temperatures because of its mechanical attachment to the IR sensor array operating below 100°K. For example, as a 1000 frames/second image processing fast frame seeker, the neural circuits are required to complete a "single-pass" of signals through a synapse-neuron composite in less than 200 nanoseconds, and must be operational below 100°K; the 64-chip-cube consuming less than 2 watts of power during its data processing operation[4].

## Electronic Device Designs:

We have designed, fabricated, and tested high speed and low power synapse and neuron circuits on test chips for their evaluation at low temperatures. Of the various synapse designs based on EEPROM, DRAM, and SRAM for weight storage, we have chosen the latter design because it offers the unique blend of advantages: high speed digital loading of weight storage, especially of importance for fast on-line "training" of hardware; analog multiplier for compact, high speed, and low power design; voltage-input/current-output for ease of input signal distribution to a large number of synapses and ease of summation of currents in the output circuit. Synapse circuit consists of a multiplying digital to analog converter. It (Figure 1) has 7-bit digital memory that can be randomly accessed by a host computer, a 6-bit digital-to-analog converter using seal ed current mirrors, a circuit to convert the input voltage to a current in order to drive the converter's current mirror network, and a programmable current-steering network such that the synapse can be programmed to be excitatory or inhibitory. Each synapse circuit is 120x 120μm in a 1.2pm CMOS fabrication process.

The neuron circuit is basically an operational amplifier as shown in Figure 2. To avoid a speed penalty resulting from having to charge and discharge large summing-node capacitances (especially if these nodes are routed between chips), the potential of the current-summing "net" node is held constant by the corresponding neuron circuit. Programmable neuron gain is useful for normalizing the neuron's response for the number of input synapse connections. The signal conversion from current to voltage is performed by a poly resistor at the signal input terminal. This design offers a wide range, variable gain neuron with a compact geometry and a sigmoidal transfer curve whose slope can be varied with gain (a feature of importance for simulated annealing).

## Results:

The test circuits were simulated using PSIM version IV modeling tool for circuits operating at low temperatures (a PC version of PSPICE simulator specially suitable for low temperature circuit modeling) and the design optimization was performed. Final designs were fabricated on testchip as a 48-pin package using metal oxide semiconductor implementation service (MOSIS). The testing consisted of inserting the package plus socket into a liquid-nitrogen Dewar flask and the leads were taken out to an interface board for computer-assisted data acquisition and analysis. In addition, for a measurement of the timing of its operation (input to synapse, output from the connected neuron), an oscilloscope trace was obtained to record the input and output signal traces. Data was also collected at room temperature for comparison.

The test results have shown that the synapse performs the 7-bit weighting function at 77°K with good linearity. Similarly, neuron circuit has given a variable gain sigmoidal output with excellent monotonicity which is required for neural network circuits to be able to "learn" when presented with examples. Further, operation of a synapse-neuron pair has been measured as the time delay between the signal input to the synapse and the corresponding out put from the neuron. This is compared with our design goal of <200 ns (required for a 1000 -frames/second image processing speed) based on the circuit simulation performed using PSIM. Figures 3 and 4 show the oscilloscope traces at 77°K, where the rise time of 94 ns and the fall time of 81 ns, respectively, is measured. The PSIM circuit simulation has predicted that a 64x64 synapse-neuron circuit on a chip, operated at 77°K would consume 30 mW of power per chip thereby ensuring that a 64-chip 3-D stack would consume less than 2 watts of power, an important constraint.

## Conclusions:

This is the first time that such high speed analog synapse and neuron circuits have been designed and tested successfully for high speed processing at low temperatures. These results project that when implemented as a fast frame seeker 'sugar cube', the neural network processor will perform the image acquisition, discrimination, and homing functions approaching 1000 frames per second with a 64x64 IR FPA operating at 90°K.

**Acknowledgment:**

The research described herein was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology and was jointly sponsored by the Ballistic Missile Defense Organization/ Innovative Science and Technology Office (BMDO/IST), the Office of Naval Research (ONR), the Advanced Research Projects Agency (ARPA), and the National Aeronautics and Space Administration (NASA). The authors thank H. Langenbacher for technical assistance and useful discussions.

**References:**

1, R.P. Lipprnann, "An introduction to computing with neural nets," IEEE ASSP Magazine, April 1987, pp. 4-22.

2. S.P. Eberhardt, T. Daud, D.A. Kerns, T.X Brown, and A.P. Thakoor, "Competitive neural architecture for hardware solution to the assignment problem, " (Invited) Neural Networks, Vol. 4, pp. 431-442, 1991.

3. P.K. Simpson, "Foundations of neural networks," In: *Artificial Neural Networks, Paradigm, Applications, and Hardware Implementations*, Ed: E. Sanchez-Sinencio and C. Lau, IEEE Press, 1992, pp. 3-24.

4. T. Duong, et al., "I ow power analog neurosynapse chips for a 3-D "sugarcube" neuroprocessor," Proceedings of IEEE International Conference on Neural Networks, WCCI, Vol. III, June 28-July 2, 1994, Orlando, FL, pp. 1907-1911.

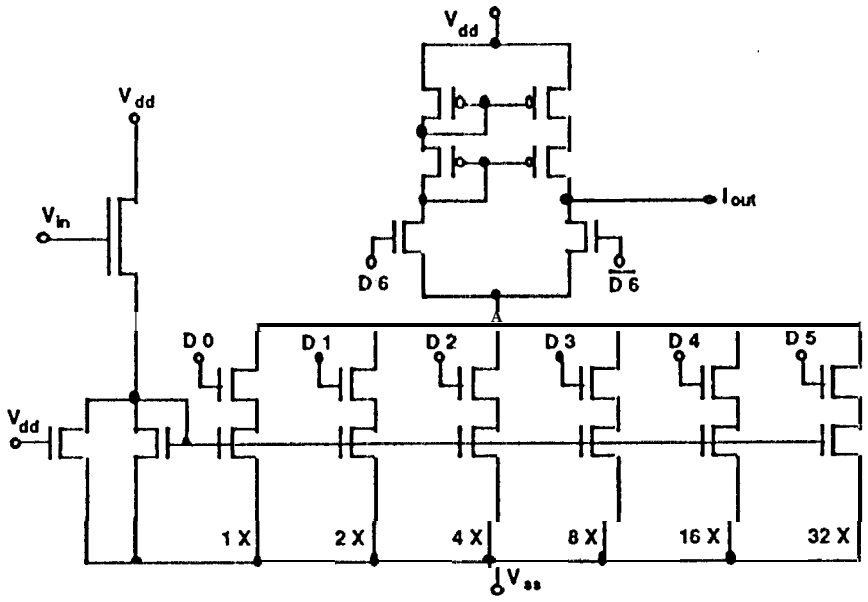


Figure 1. Circuit diagram of a 7-bit (±63 levels) digital-analog hybrid synapse cell with voltage-in/current-out mode of operation. Current summing ona wire from multiples synapses at the output node in analog domain is simple.

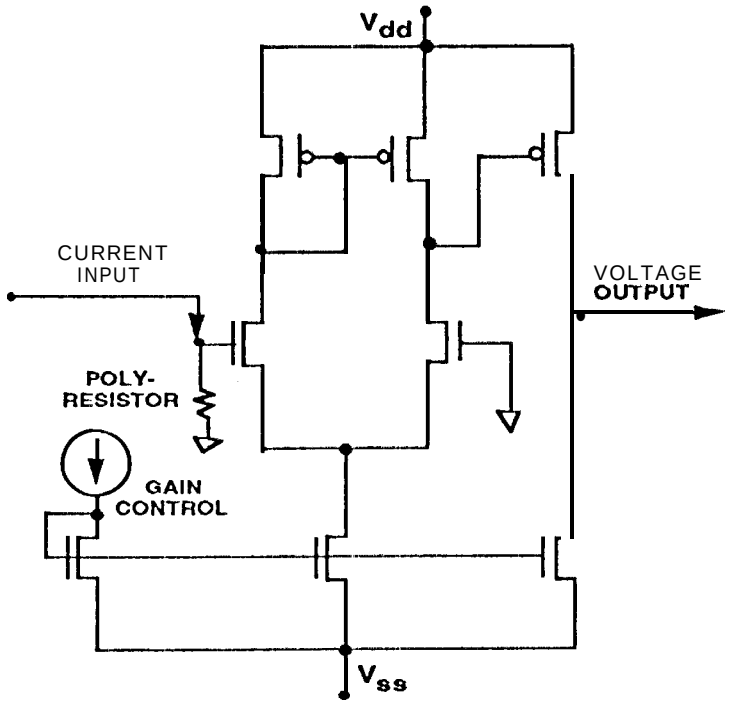


Figure 2. Circuit diagram of a wide range variable gain sigmoidal transfer function neuron cell with current-in/voltage-out mode of operation. Voltage distribution from its output to multiple synapses is easily performed.

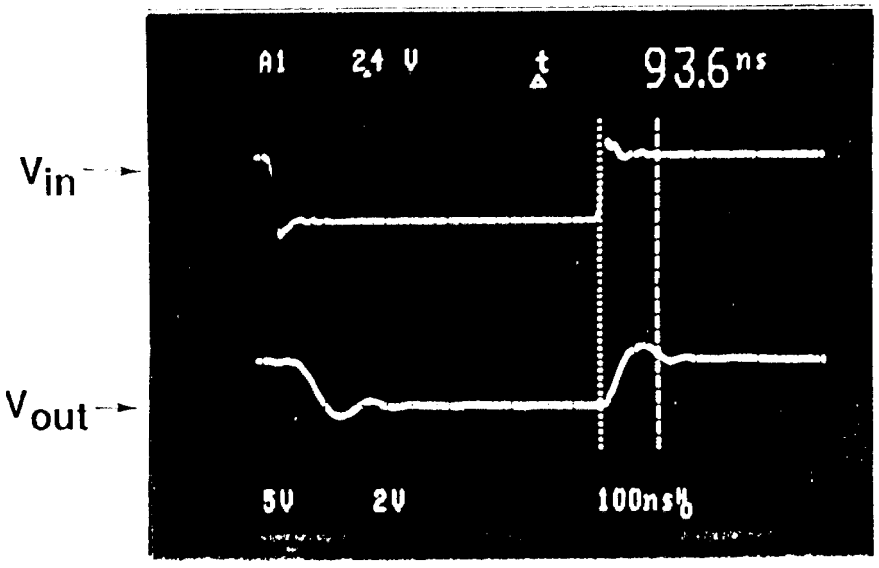


Figure 3. An oscilloscope trace showing the rise time between an input to a synapse and the neuron output at 77°K.

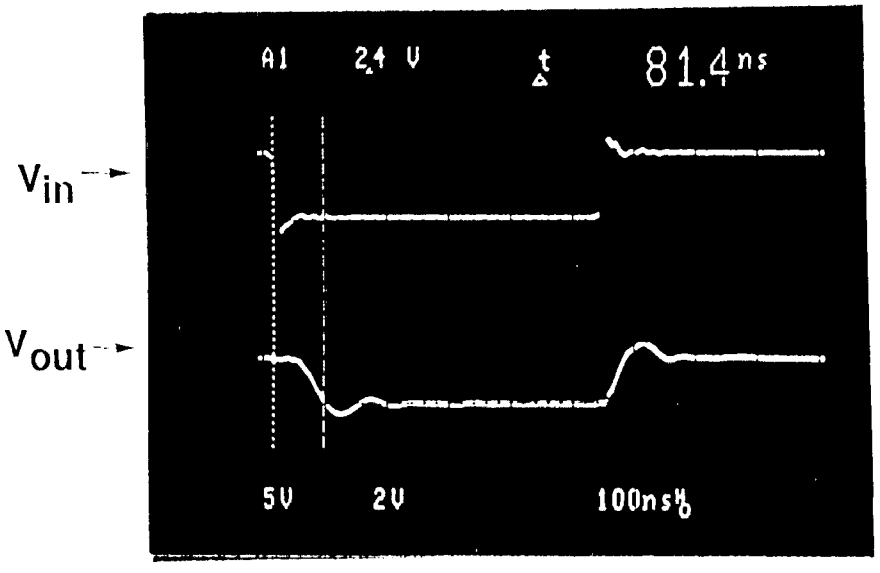


Figure 4. An oscilloscope trace showing the fall time between an input to a synapse and the neuron output at 77°K.