Analog Integrated Circuits and Signal Processing 6, 45-51 (1994)

Pulse-Modulated Analog Neuron Circuits

BOYD FOWLER AND ABBAS EL GAMAL
Department of Electrical Engineering, Stanford University, Stanford, CA 94305

Received October 1, 1992; Revised August 2, 1993

Abstract. Analog neuron circuits based on both frequency modulation and pulse modulation are investigated. The circuits are compared in terms of size, power, performance, and reliability; frequency modulation shows advantages in each area. Test circuits were designed and fabricated in 2 μm CMOS technology. The frequency modulated neuron has an operational frequency of 3.125 MHz and a dynamic range of 17 bits. Our results indicate that this circuit technique may provide substantial advantages in high-performance, low-power neural systems.

1. Introduction

A neuron circuit implements the function

$$y = f\left(\sum_{i=1}^{n} w_i x_i\right) \tag{1}$$

where f is a bounded odd real function, y is the output of the neuron, x_i is a component of the input vector X, and w_i is a component of the weight vector W. We assume that y, x_i, and w_i are bounded real numbers.

Several analog neuron circuits have been developed based on unmodulated signaling techniques. An example is the neuron in the Intel 80170NX [1] depicted in figure 1. In such implementations the signals are communicated as voltage or current levels. This results in low signal-to-noise ratios (SNR) especially for long distance on-chip and on- and off-chip (I/O) communication. A widely employed technique for maintaining SNR over long distances is signal modulation.

Several researchers have suggested the use of pulse modulation [2, 3]. Pratt [2] discussed the theory of

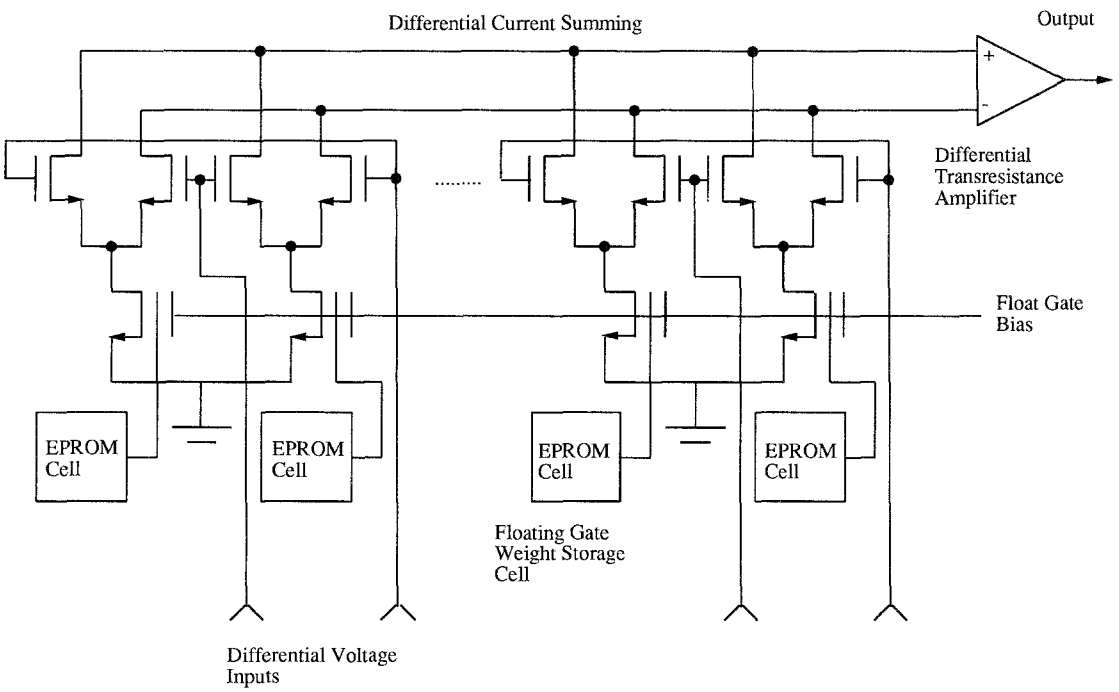

Fig. 1. Differential analog neuron circuit with floating-gate weight storage.

Table 1. Pulse modulation techniques.

Pulse modulation techniques	
PAM	Pulse amplitude modulation
PWM	Pulse width modulation
PFM	Pulse frequency modulation
PPM	Pulse phase modulation

neural information encoding in biological systems and explained how this knowledge can be used to optimize artificial neural networks. Specifically, he investigated pulse frequency modulation and binary information encoding techniques, the latter first proposed in [4]. Pratt concludes that biological systems use pulse-modulated information encoding because it offers unique computational advantages over other techniques. Murray et al. [3] investigated four pulse modulation techniques shown in table 1, and described the theoretical advantages and disadvantages. They also compared two pulse-modulated VLSI designs.

Many researchers have also explored pulse-modulated analog neuron circuit techniques [5–8]. Murray et al. [5, 6] investigated pulse frequency signaling and current summation in biologically realistic neuron circuits. Fabricated neuron circuits achieved an operational frequency of 1 MHz using 3 μm CMOS. Tomberg and Kaski [7] investigated pulse-density modulation techniques by implementing switched-capacitor and digital neuron circuits. The switched capacitor neuron circuits operated in simulation at 300 kHz using 2 μm CMOS. Meador et al. [8] investigated floating-gate weight storage in pulse neuron circuits. These circuits operated in simulation at 3 MHz using 2 μm CMOS.

These results suggest that PWM, PFM, and PPM techniques have substantial theoretical and practical advantages over unmodulated system. In this paper we introduce a new analog neuron circuit using FM modulation. FM signals encode information as the frequency of a 50% duty cycle square wave, while PFM signals encode information in the number of small pulses per unit time.

We compare our PWM and FM neuron circuits and conclude that FM circuits have superior operational characteristics.

Section 2 introduces the two neuron circuits considered in this paper. Operation of the pulse-modulated neuron circuit is described in Section 2.1, and the frequency-coded neuron circuit is described in Section 2.2. Results from a fabricated test chip are described in Section 3. These results include data for both the frequency-modulated neuron circuit and the pulse-width-modulated neuron circuit.

2. Implementation

2.1. Pulse-Width-Modulated Neuron

Pulse-width-modulated neurons consist of two separate sections, synaptic weights and neuron summers. Figure 2 shows a pulse modulated neuron with four synapses, two positive and two negative. The input to each synapse is a voltage with fixed amplitude, variable pulse width, and constant period. The output of each neuron circuit has the same signal form as the input to each synapse.

Each synapse is implemented using a single NMOS transistor. The polarity of the synapse is determined by whether it connects to the negative op-amp terminal through a current mirror, whereas the size of the NMOS transistor determines the magnitude of the synaptic weight. Fixed-magnitude synapses prohibit learning in the test circuits but allow an unbiased comparison of the two neuron circuits. Floating-gate transistors could be used to facilitate adjustable synapses in future designs. The input to each synapse is a pulse of fixed voltage amplitude and variable time duration, where the pulse duration can vary between zero and the period of one clock cycle. Voltage pulses at each synapse are converted into currents, which are summed at the negative op-amp input and stored on C_1.

The output voltage of the op-amp at the end of one clock cycle is approximated by equations (2) and (3), with the approximation based on the assumption that the output swing of the op-amp is linear between $V_{\max}$ and $V_{\min}$. The voltage on the feedback capacitor is reset at the end of each clock cycle.

$$V_{\text{output}} = f\left(\frac{(\Sigma I_j)\Delta t}{C_1} + V_{\text{bias}}\right) \tag{2}$$

$$f(x) = \begin{cases} V_{\min}, & x < V_{\min} \\ x, & V_{\min} < x < V_{\max} \\ V_{\max}, & V_{\max} < x \end{cases} \tag{3}$$

V_{bias} is the voltage on the positive terminal of the op-amp, ΣI_j is the current into the negative terminal of the op-amp, and Δt is the clock period. We use a folded cascode op-amp as shown in figure 3 optimized for minimum power dissipation and settling time. Assuming each synaptic transistor is operating above threshold in saturation each I_j in (2) can be expressed by

$$I_j = \pm \frac{W_j}{L_j} C_{\text{ox}} \mu_n (V_{\text{pulse}} - V_t)^2 (1 + V_{ds}\lambda_n) \frac{\text{pulse width}}{\Delta t} \tag{4}$$

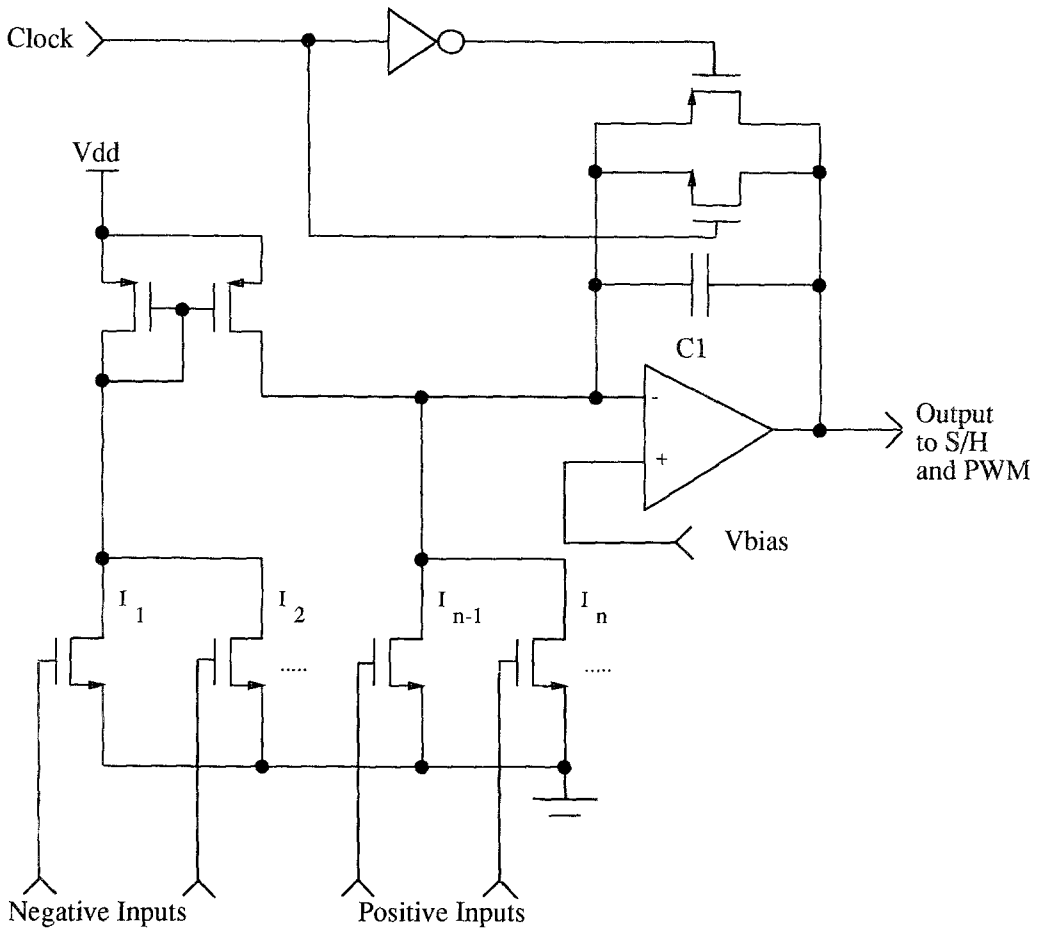

Fig. 2. Pulse-width-modulated neuron circuit.

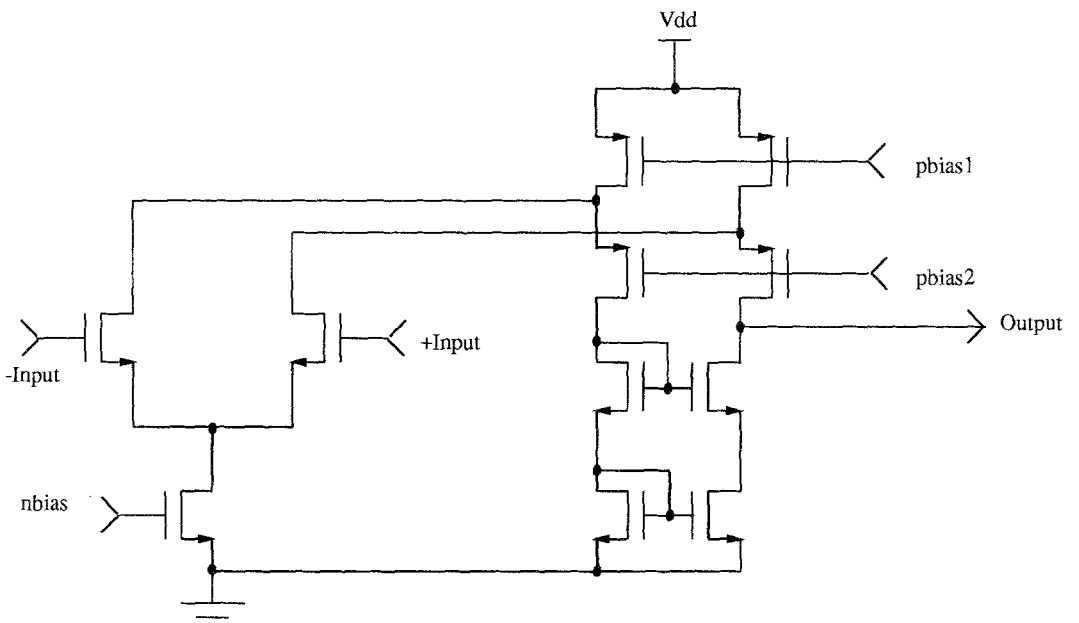

Fig. 3. Op-amp circuit.

where W_j is the width of the input transistor, L_j is the length of the input transistor, C_{ox} is the oxide capacitance, and μ is the channel mobility. Additionally, V_{pulse} is the voltage level of the input pulse, V_t is the threshold of the input transistor, V_{ds} is the drain-to-source voltage of the input transistor, and λ_n is the channel length modulation constant of the input transistor. The plus or minus in (4) represents the polarity of the synaptic weight.

At the end of each clock cycle the output of the op-amp is sampled and then converted into a pulse, the duration of which is proportional to the magnitude of the op-amp voltage. The output of the pulse width modulation circuit as a function of op-amp voltage is described by

$$\text{Pulse width} = \frac{V_{\text{output}} - V_{\text{min}}}{V_{\text{max}} - V_{\text{min}}} \Delta t \qquad (5)$$

where V_{min} and V_{max} are defined by the operating limits of the neuron circuit.

2.2. *Frequency-Modulated Neuron*

Frequency-modulated systems also consist of two separate parts, synaptic weights and neuron summers. Figure 4 shows a frequency-modulated neuron with four synapses, two positive and two negative. The input to each synapse is a PFM signal, but for the output of each neuron circuit we use a 50% duty cycle FM signal. The FM signal is used for long-distance communication, because its SNR characteristics are superior to those of PFM signals. The PFM signal needed for each synapse is locally decoded from an FM signal using the circuit in figure 5. The input in figure 5 is a 50% duty cycle square wave of a given frequency, and the output is a PFM signal. This circuit uses local feedback and time integration to reduce the effects of device offsets and random noise.

Each synapse consists of one NMOS transistor. The polarity of the synapse is determined by the transistor's connection to the neuron. Positive synapses are connected to the negative feedback loop transconductance

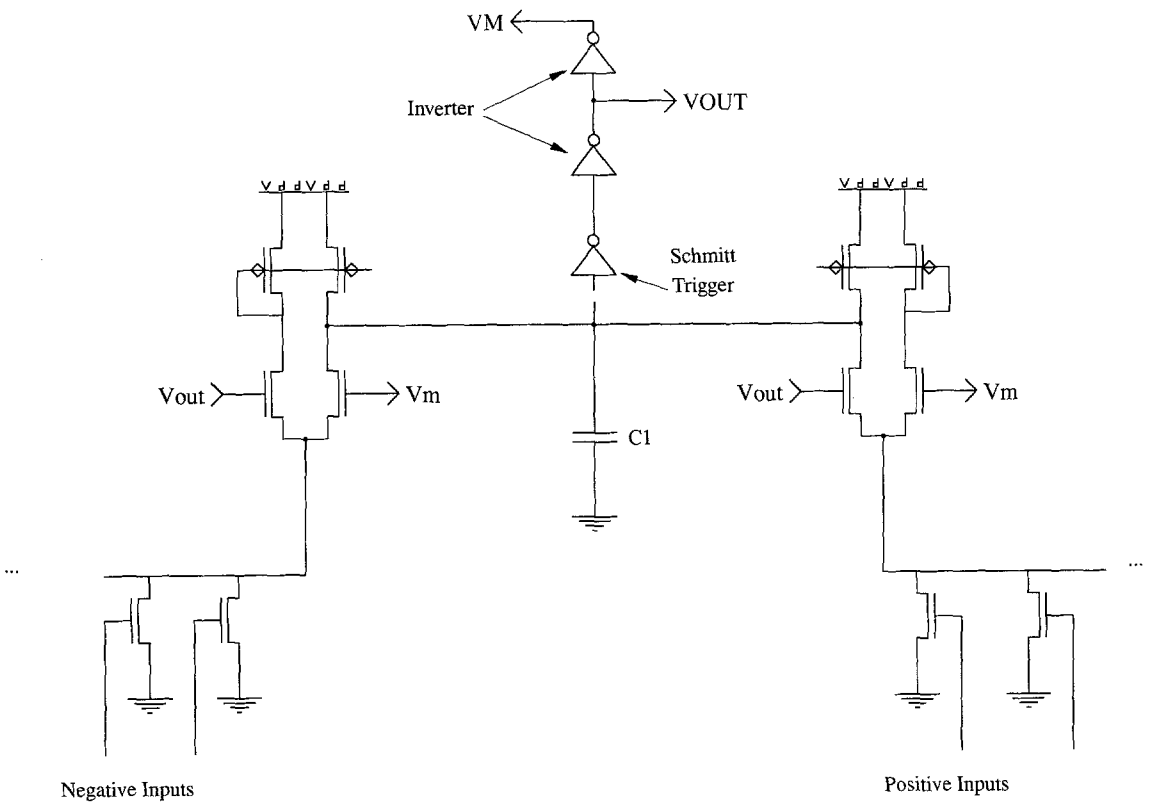

Fig. 4. Frequency-modulated neuron circuit.

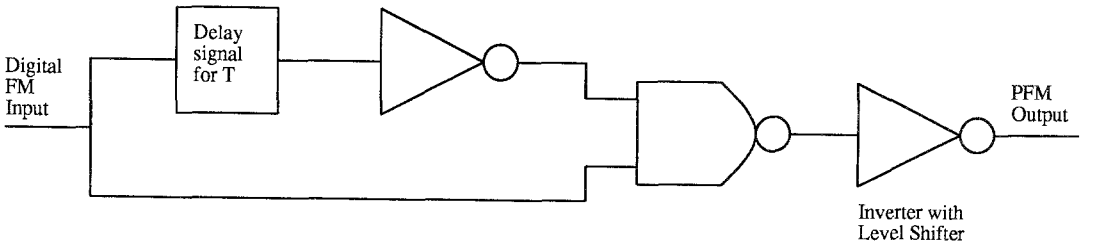

Fig. 5. FM to PFM decoder circuit.

amplifier, and negative synapses are connected to the positive feedback loop transconductance amplifier. The weight of each synapse is determined by the size of the NMOS transistor.

Synapses convert PFM signals into charge packets which are integrated on C_1 and either enhance or retard (depending on the polarity of the synapse) the output frequency. The time-varying voltage on C_1 causes the output to oscillate. Assuming the rise time of the inverters is sufficiently small, the output frequency of the neuron is given by

$$f_{\text{output}} = g\left(\frac{\Sigma I_j}{2C_1(V_{th} - V_{tl})}\right) \tag{6}$$

where

$$g(x) = \begin{cases} 0, & x < 0 \\ x, & 0 < x < f_{\max} \\ f_{\max}, & f_{\max} < x \end{cases} \tag{7}$$

ΣI_j is the sum of the synaptic inputs to the neuron, and V_{th} and V_{tl} are the high-voltage trip point and low-voltage trip point of the Schmitt trigger.

Assuming each synapse is operating above threshold in saturation, each of the synaptic currents is given by

$$I_j = \pm \frac{W_j}{L_j} C_{\text{ox}} \mu_n (V_{\text{pulse}} - V_t)^2 (1 + \lambda_n V_{ds}) \Delta t f_{\text{input}_j} \tag{8}$$

where W_j is the width of the transistor, L_j is the length of the transistor, V_{pulse} is the voltage of the PFM signal, V_t is the threshold voltage, λ_n is the channel length modulation constant, V_{ds} is the drain-to-source voltage of the transistor, Δt is the duration of a pulse in the PFM signal, and f_{input_j} is the frequency of the PFM signal at the jth synapse. The polarity of the synaptic weight is determined by the sign in (8).

3. Test Chip Results

Two PWM neurons and two FM neurons with four synapses each have been fabricated in 2 μm p-well CMOS technology with two layers of metal and two layers of poly. The fabrication was performed at ORBIT through MOSIS. Figure 6 is a CIF plot of the four neurons, synapses and associated multiplexers, which occupy an area of 0.00531 mm^2 and operate from a single 5-V power supply.

The performance of the neurons was determined by driving the individual circuits with either PWM, FM, or dc signals (as needed) from off-chip function generators. The output of each neuron was observed with frequency counters and analog oscilloscopes. The tests performed included I/O transfer characteristic measure-

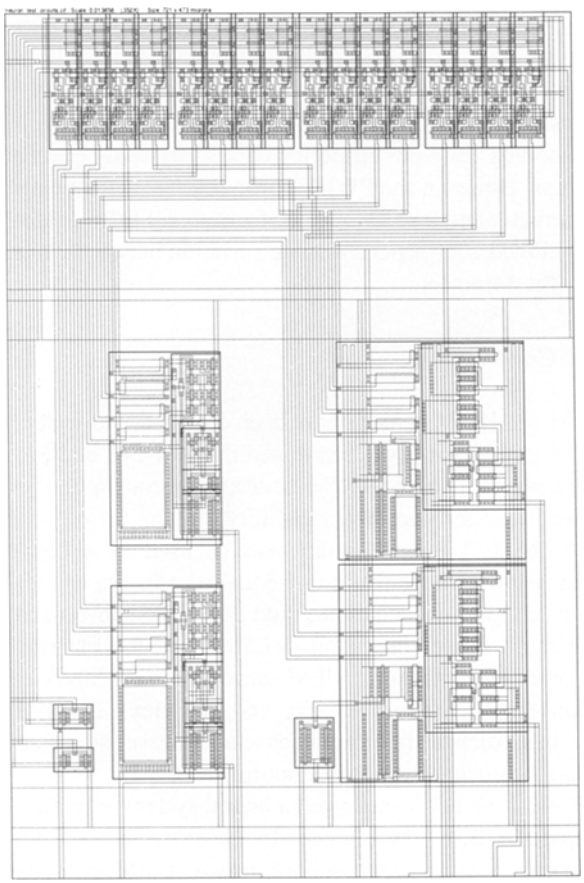

Fig. 6. Layout of neuron circuits.

Table 2. Neuron test results.

Neuron test results		
	PWM neuron	FM neuron
Maximum operating frequency	500 kHz	3.125 MHz
Dynamic range	6 bits	17 bits
Addition linearity error	±1.5 LSB	±0.5 LSB
Supply voltage	5 V	5 V
Layout area without synapses	2.9e-4 mm^2	1.49e-4 mm^2
Technology	2 μm CMOS	2 μm CMOS

ment, frequency response measurement, and weighted addition calculations. The results of these tests are summarized in table 2; note that addition linearity tests were conducted using circuits with only 5 bits of accuracy.

Neuron output waveforms are shown in figures 7 and 8. These graphs were obtained by stimulating one synapse in a neuron of each type, where the synapse sizes were the same. These data were sampled from an oscilloscope. I/O transfer characteristics for both neurons are shown in figures 9 and 10. These graphs were obtained by stimulating one synapse of a given neuron (again, the synapse size is the same for both neurons).

The performance of the FM neuron exceeded initial design expectations. Operation is stable, and there is generally little circuit performance variation between chips. However, the PWM neuron is difficult to operate correctly and tended to suffer from signal jitter. It also exhibited a large performance variation between chips in the same lot.

4. Conclusions

FM and PWM neurons have been designed, fabricated, and tested. We have determined that our FM neurons are superior to our PWM neuron in throughput, dynamic range, size, and reliability.

The FM neuron could also be used as a high-precision 17-bit current-to-frequency converter. This data conversion technique could be used to convert and transmit sensor data from VLSI ICs to the external world. Presently most VLSI sensors use on-chip A/D converters, which are large and complex. The FM neuron offers a simple and dense alternative for sensor data acquisition and transmission.

We have not investigated a neural system with multiple PWM or FM neurons. However, we expect that neural systems composed of multiple FM neurons would be more robust than neural systems composed

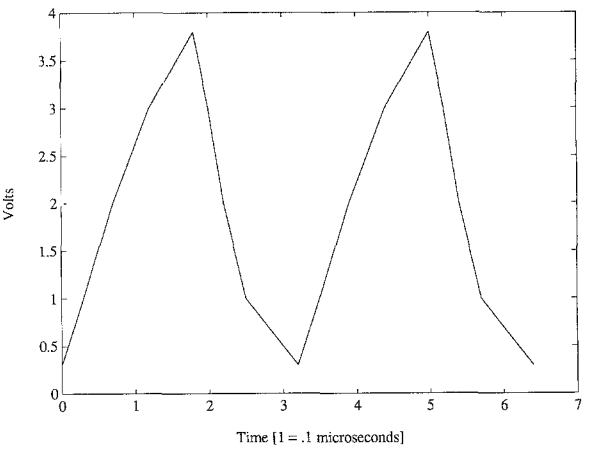

Fig. 7. Output waveform: FM neuron operating at 3.125 MHz.

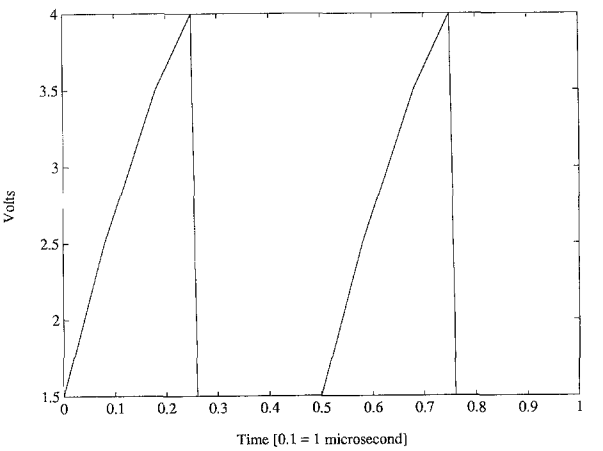

Fig. 8. Output waveform (from op-amp): PWM neuron operating at 200 kHz.

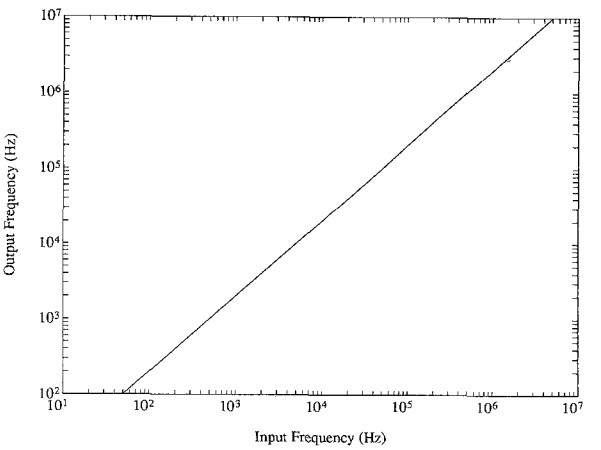

Fig. 9. I/O transfer characteristic for FM neuron.

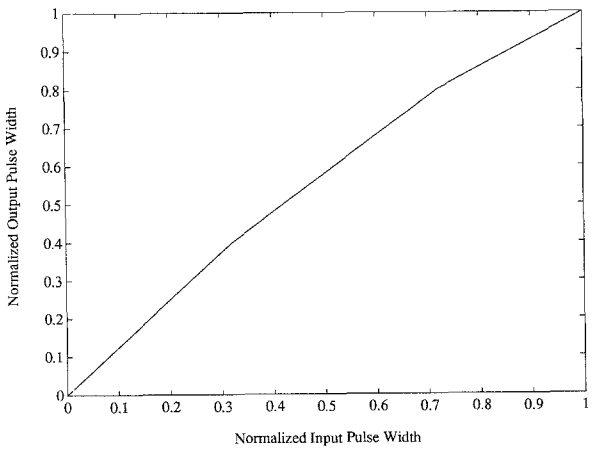

Fig. 10. I/O transfer characteristic for PWM neuron.

of multiple PWM neurons. This expectation is related to the FM neuron's higher tolerance to parameter variations and more robust signal communication technique. These characteristics help to minimize single point failures and error propagation.

Acknowledgments

The authors wish to acknowledge the support of Texas Instruments and the partial support of DARPA under contract J-FBI-89-101. They are also thankful for the encouragement and support of Michael Godfrey and the helpful comments of the reviewers.

References

1. Intel. 80170nx, *Electrically Trainable Analog Neural Network*, June 1991, Order number 290408-002.
2. G.A. Pratt, *Pulse Computation*, Ph.D. thesis, MIT, 1989.
3. A.F. Murry, D. DelCorso, and L. Tarassenko, "Pulse stream VLSI neural networks mixing analog and digital techniques," *IEEE Trans. Neural Networks*, Vol. 2, pp. 193–204, 1991.
4. W.S. McCulloch and W.H. Pitts, "A logical calculus of the ideas imminent in nervous activity," *Bull. Math. Biophys.*, Vol. 5, pp. 115–133, 1943.
5. A.F. Murry and A.V.W. Smith, "Asynchronous VLSI neural networks using pulse-stream arithmetic," *IEEE J. Solid-State Circuits*, Vol. 23, pp. 688–697, 1988.
6. A.F Murry, "Pulse arithmetic in VLSI neural networks," *IEEE Circuits Systems*, pp. 64–74, Dec. 1989.
7. J.E. Tomberg and K.K.K. Kaski, "Pulse-density modulation techniques in VLSI implementations of neural network algorithms," *IEEE J. Solid-State Circuits*, Vol. 25, pp. 1277–1286, 1990.
8. J.L. Meador, A. Wu, C. Cole, N. Nintunze, and P. Chintrakulchai, "Programmable impulse neural circuits," *IEEE Trans. Neural Networks*, Vol. 2, pp. 101–109, 1991.

Boyd Fowler received the B.S. degree summa cum laude in engineering from California State University, Northridge, in 1987, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1990. Presently he is a Ph.D. candidate in the Information Systems Lab at Stanford University. He worked at Hughes Aircraft Company, Canoga Park, CA, during the summer of 1991 on the design and VLSI implementation of high-speed OCR systems, and has been a research assistant in the Information Systems Lab, Stanford University from 1988 to the present. His current research interests include real-time focal plane image compression, analog signal processing, ultra-low-power analog circuits, and ultra-low-power digital circuits. He is a member of Tau Beta Pi, Phi Kappa Phi, and IEEE.

Abbas El Gamal (S'71-M'73-SM'83) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1978. He is currently an associate professor of electrical engineering at Stanford University and Chief Technical Officer of SiArc. From 1978 to 1980 he was an assistant professor of electrical engineering at the University of Southern California. From 1981 to 1984 he was an assistant professor of electrical engineering at Stanford. He was on leave from Stanford from 1984 to 1987, first as director of LSI Logic Research Lab, then as cofounder and chief scientist of Actel Corporation. His research interests include VLSI architectures, design automation and synthesis for VLSI, field-programmable gate arrays and MCMs, and information theory.