Mixed-signal neuron-synapse implementation for large scale neural network

I. S. Han,
Department of Electronic and Electrical Engineering
University of Sheffield, Sheffield, S1 3JD, U.K.

This paper describes a new mixed-signal VLSI implementation of neural networks for low power and asynchronous operation. The linearised transconductance produces the synaptic function of multiplication, weight programming, and summation of synaptic currents for the neuron. The synapse circuit is designed with 8 transistors, by compensating the non-linearity of MOSFET resistance in the triode region.

The flexible configuration of synapse accommodates either the pulse-based implementation or any analogue synapse with multiplication or summation. The operation speed of individual synapse is up to 300 Mega operations with the power consumption of less than 33μW, from the chip design using 0.18μm(3.3V) CMOS process. The accuracy is extendable in modular structure, though the current design is based on 8-bit accuracy. The overall power consumption can be less in practice, as individual synapse only demands the power when there is an active signal.

The advantages of proposed VLSI implementation are the large scale implementation with low power consumption and its adaptable features to various requirements from different paradigm of neural network architecture, depending on the demand of asynchronous architecture, pulse/spike-based architecture, or the accuracy requirements.
Key words : Analogue-mixed VLSI neural network, pulse/spike-based neural computation, asynchronous operation, MOSFET resistance , voltage-controlled linear resistance

I. Overview

VLSI neural network has been continuously developed either in digital or analogue, as both methods have different advantages. The digital VLSI has advantages of design flexibility or current technology of 0.18μm or 0.13μm CMOS. The advantage of analogue is low power consumption or large network integration. Issues in an analogue neural network VLSI can be the accuracy problem, or the complexity, in comparison with the digital. In some special applications, the analogue utilised or developed better its non-ideal characteristic or complex design[1-2]. There has been a new expectation in brain science, such as large scale general neural network of cognition[3]. Such demand in large capacity or high performance has been driving the analogue-mixed neural network VLSI, as the large scale as well as the accuracy can be achieved. There is the need for improvements in analogue-mixed implementation, as the fabrication technology advances towards 0.18μm or 0.13μm. An example of analogue-mixed VLSI is in Fig.1, which was developed for real-time packet control[4]. Though the circuit has the accuracy, it demands the complex supply voltages which limit the level of mixed integration in low supply voltage.
This paper introduces the new development in analogue-mixed network VLSI with the advantage of the advanced fabrication technology.

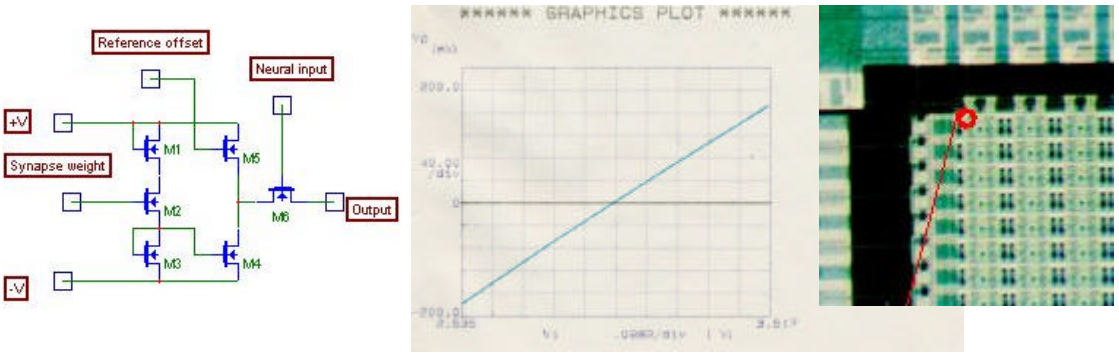

Fig. 1. Electronic synapse circuit with compensated linear MOSFET resistance in the triode region, and neural networks VLSI for packet switch controller

II. New analogue-mixed neural network VLSI for low power and asynchronous operation

The current-voltage(I-V) relationship of MOSFET in the triode region could be clearly adopted to implement electronic synapses, as it provides the multiplication function. The equation of interest is that the drain-source current I_{DS} for a MOSFET in the linear or triode region:

$$I_{DS} = \alpha \{ (V_{GS} - V_T) \bullet V_{DS} - V_{DS}^2/2 \} \qquad (1)$$

Here, α is the MOSFET process parameter, V_{GS}, V_{DS}, V_T, the transistor gate-source, threshold and drain-source voltage respectively. To achieve a linear voltage(V_{GS})-to-voltage(V_{DS}) multiplier, the balanced circuit with an operational amplifier can be used to remove the second order term of $V_{DS}^2/2$ in eq.1. Though various methods for suppressing or controlling the nonlinearity have been developed, there are limitations in solving the nonlinear problems. The synapse circuit of Fig. 1 still has the drawback of demanding bipolar supply voltages to remove the second order term in eq.1.

Proposed scheme: The synapse circuit proposed in this paper makes simple use of the MOSFET in the triode region. It realises the high speed and a small size analogue multiplier without an amplifier or dual supply voltages. The synapse circuit of the MOSFET resistance-based analogue multiplier is shown in Fig. 2. For an efficient analogue multiplication, two terms, $V_T \bullet V_{DS}$ and $V_{DS}^2/2$, from eq. 1 should be eliminated from the output. Based on eq. 1 for the MOSFET triode region, the currents of transistor M1 and transistor M2 are;

$$I_{M1} = \alpha \{ (V_{2DC} - V_T) \bullet V_1 - V_1^2/2 \} \qquad (2)$$

$$I_{M2} = \alpha \{ (V_2 - V_T) \bullet V_1 - V_1^2/2 \} \qquad (3)$$

Here, V_2 is one of two inputs for the multiplication and V_{2DC} is the offset component of V_2 to keep transistors M1 and M2 in the triode region. V_1 is the other of the two inputs involved in the multiplication. Combining eqns. 2 and 3, the output difference current may be shown to be

$$I_{OUT} = IM1 - IM2$$
$$= \alpha(V_2 - V_{2DC}) \bullet V_1$$

$$= \alpha \bullet V_{2SIG} \bullet V_1 \qquad (4)$$

where V_{2SIG} is the signal component of V_2. From eq. 3, the multiplication function is attained with MOSFETs in the triode region and the pairs of (M3, M8) and (M4, M7) each acts as a pair of current mirror and source. The output transistors of M6 and M8 source or sink the post synaptic current I_{OUT} of eq. 4. The summation function of neuron is embedded in each synapse's wired-OR capability.
The new VLSI synapse is designed using 0.18μm standard digital CMOS technology and tested by HSPICE simulation. The power supply voltage is 3.3V and the low power consumption as well as the accurate operation are design objectives. The linearity of synaptic computation is illustrated in Fig.3 and well comparable to analogue mixed-mode neural computation. The overall power is less than 33mW per active cell and the new circuit is suitable for large scale VLSI neural network implementation. As observed in the circuit of Fig.2, there is no current or power consumption with an inactive neural input.

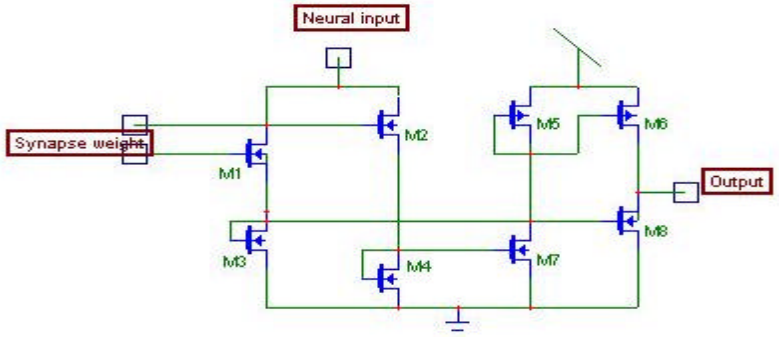

Fig. 2. New synapse circuit design by the voltage-controlled linear resistance of two MOSFETs in the triode region

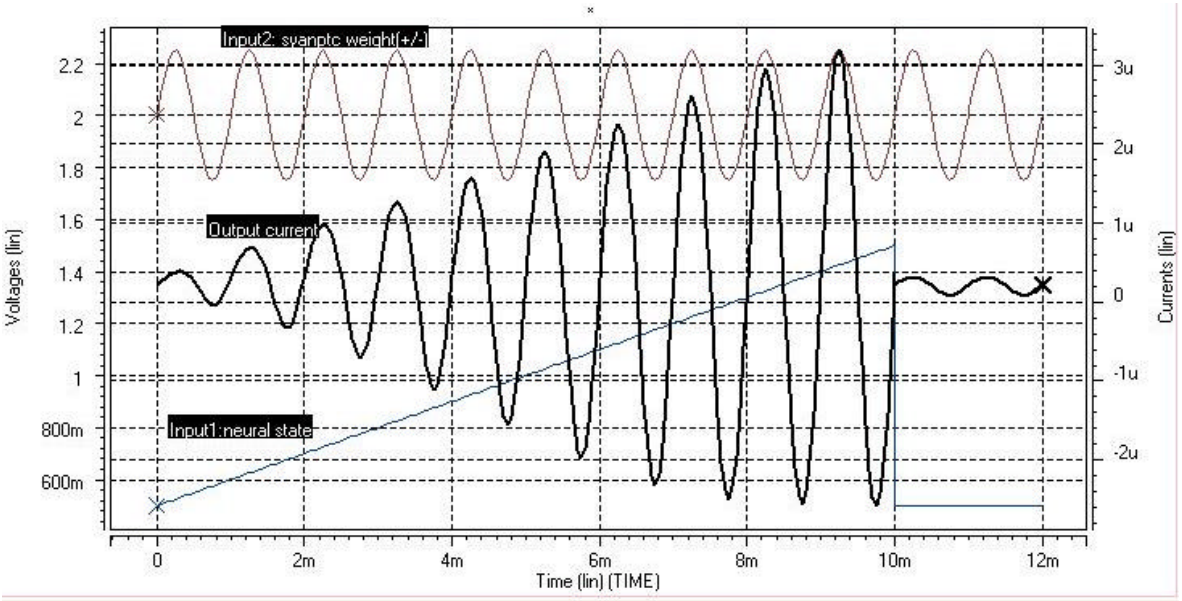

Fig.3. V-I characteristics of new synapse circuit, with multiplication function of input1 (neural state)X input2(synaptic weight) for output as current

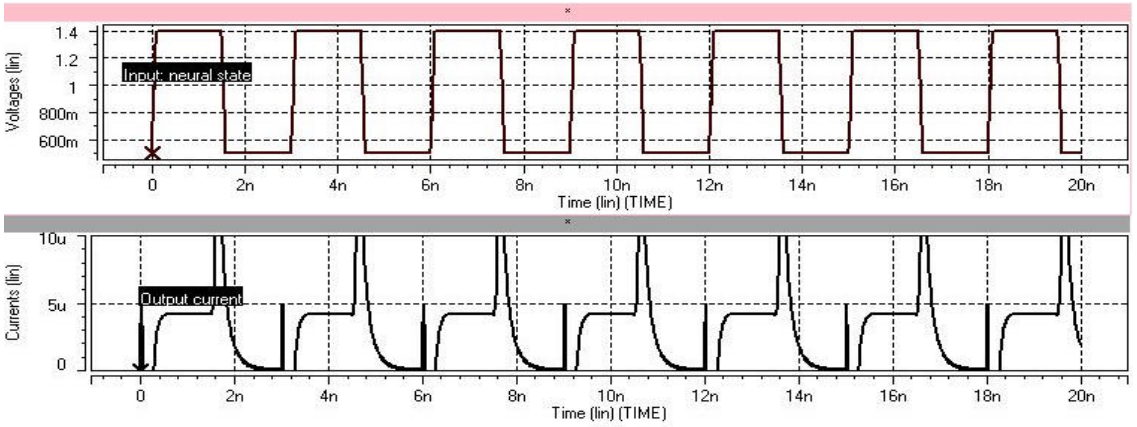

Fig.4. Analogue-digital mixed operation over 300 Mega computation per second per unit synapse cell for pulse/spike implementation

The speed of more than 300Mega connection per second inputs is tested as in Fig. 4. A flexibility in power consumption can increase the operation speed further, as it accelerates charging or discharging rate of output current.

III. Conclusion

The new analogue-mixed synapse demonstrates the feasibility of implementing large scale, low power and asynchronous neural network VLSI, because of its simpler circuit, speed and accuracy . It is flexible to adapt its characteristics of speed, power, accuracy on demand, as those can be tailored by controlling the current.
The core of proposed electronic synapse is within the size of 2μmX2μm and the VLSI in 0.18μm digital CMOS technology provides the advantage of analogue-digital mixed neural network system on chip with small power consumption.

Reference

1. Yoshihiko Horio, Kazuyuki Aihara, and O. Yamamoto, “Neuron–Synapse IC Chip-Set for Large-Scale Chaotic Neural Networks, IEEE Trans on Neural Networks, vol. 14, no 5, Sept, pp1393-1404, 2003
2. Momchil Milev and Marin Hristov, “Analog Implementation of ANN With Inherent Quadratic Nonlinearity of the Synapses”, IEEE Trans on Neural Networks, vol. 14, no 5, Sept,pp1187-1200, 2003
3. John G. Taylor, ”Paying attention to consiousness,”TREMDS in Cognitive Sciences, vol.6. No. 5, May, pp206-210, 2002
4. Il-Song Han and R. Webb, "Neural Network Switch Controller with Analogue-Digital Mixed Neural Network VLSI," Proc. of EANN'97, pp. 299-302, 1997
5. I. S. Han, K. C. Kim and H. S. Lee, "Implementation of Neural Hardware with the Neural VLSI of URAN in Applications of Reduced Representation", NIPS, Nov, 1994
6. Il-Song Han and Ki-Hwan Ahn, "Neural Network VLSI Chip Implementation of Analogue-Digital Mixed Operation for more than 100,000 Connections" MicroNeuro'93, pp.159-162, 1993

7. Il-Song Han, "Analogue Circuit for Neural Network," ICANN'91, Vol.2, pp.1577-1579, 1991
8. M. Brownlow, L. Tarassenko, A. F. Murray, A. Hamilton, Il-Song Han, H. M. Reekie, "Pulse Firing Neural Chips Implementing Hundreds of Neurons" NIPS2, pp.785-792, 1990
9. Il S. Han and Song B. Park, "Voltage-Controlled Resistors by MOS Transistors and their application to Active RC Filter MOS Integration," Proceedings of IEEE, pp.1655-1657, Nov., 1984