A Means of VLSI Current Controlled Weight Setting in ANNs

S. K. Singh[%], R. W. Newcomb[%], P. Gomez[&] and V. Rodellar[&]
[%]Electrical Engineering Department [&]Facultad de Informatica
University of Maryland Universidad Politecnica de Madrid
College Park, MD 20742 USA Madrid, 28660 Spain
sksingh@eng.umd.edu

ABSTRACT

A means is given for setting the weights in BiCMOS VLSI realizations of analog neural networks of the Hopfield class. This is achieved by the use of a BJT 4-quadrant current multiplier for which one set of the multiplier's differential current inputs comes from a CMOS differential pair giving a neuron's output current from its input capacitor voltage and the other set of the multiplier's differential currents comes from a differential pair with its input being a weight setting voltage.

Introduction

Since the Hopfield neural network structure [5] contains many other artificial neural networks (ANNs) as special cases, it is a favorite for study and implementation. For example, feedforward layered ANNs result by choosing special semidiagonal weight matrices. Following in the footsteps of Hopfield himself, [5, Fig. 2] most of the early implementations used voltage controlled voltage source (VCVS) amplifiers with resistors, R, for weights [1][3][11]. However, if one wishes to construct these ANNs in VLSI, which is the implementation of choice when hundreds of neurons are desired, then VCVS-R circuits are not practical. Consequently there has been a switch to the use of voltage controlled current sources (VCCS) which are more amenable to analog very large scale intergration (VLSI) [6]-[11] since they are readily constructed through differential pairs. However, whenever a multiplier has been needed, as when multiplying by " synaptic " weights, we only find that voltage input multipliers have been used in the ANN literature [8][9]. Since current mode circuits are the most convenient for VLSI it is appropriate to search for current mode multipliers. Indeed we do find that Herpy [4, p. 415] presents a four-quadrant current mode multiplier using bipolar junction transistors (BJTs). Here we use BJTs available in the MOSIS fabrication facility in this circuit to be able to make a more pertinent realization of the Hopfield architecture in analog VLSI.

1. The Basic Configuration

The basic equations we wish to implement are, in matrix form,

$$C\,\frac{dx}{dt} = WI_{out} + I_{bias} + I_{in} \qquad (1)$$

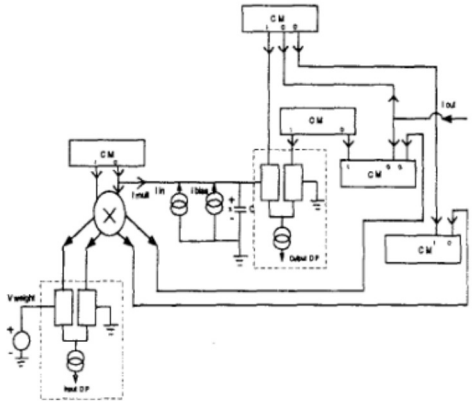
$$I_{out} = g_{act}(x) \qquad (2)$$

where, if there are n neurons, x is the state n-vector of capacitor voltages, C is the diagonal nxn matrix of capacitances, W is the nxn weight matrix, I_{bias} is an n-vector of bias currents and I_{in} is an n-vector of external input s, I_{out} is the n-vector of neuron outputs, which are the states x passed through the activation function, an n-vector function, $g_{act}(.)$. The weights W could be set by voltages that are determined by some external means, as via an adjoint type of network in back-propagation or adaptively in an on-line adaptive control situation, in which case

$$W = g_m * V_{weight} \qquad (3)$$

Figure 1 shows a schematic representation of equations (1), (2) and (3) as important to our current mode realization. In Fig. 1 the multiplier (middle left) takes the weight W, as set by a difference of currents (lower left in Fig. 1) and multiplies it by the difference current I_{out} to give the weight term WI_{out} in equation (1). Since I_{out} and the entries of W are most naturally generated by VCCSs, it is most natural to form WI_{out} in current multipliers. Because the entries of W and of I_{out} may generally take on any sign, the VCCSs should be of differential type, DVCCSs, and the multiplier should be of the four- quadrant type. In actual fact some current mirrors can be saved by individually using the two

0-7803-2768-3/95/$4.00



C M = Current Mirror, I = input to CM, O = output from CM, D P = Differential Pair
Cdx/dt = I in + I bias + I mult
I mult = [gm * Vweight] * I out
I out = g act(x)

Fig. 1: Basic IMode Configuration

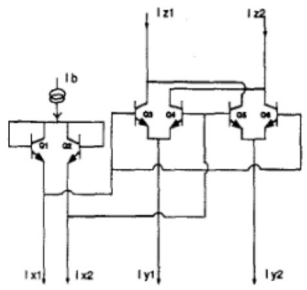


Fig. 2: Four-Quadrant BJT Current Multiplier

differential pair transistor currents rather than their differences formed in current mirrors, so we develop a four-quadrant current multiplier where there are two sets of input current pairs resulting, as we will see in the next section, in a pair of output currents, the difference of which will be taken in a current mirror to form the single multiplier current I_{mult}. Next we turn to the multiplier itself.

2. The Four-Quadrant BJT Current Multiplier

The BJT four-quadrant current multiplier we propose to use is that of Herpy [4, p. 414] shown in Fig. 2 where the two input current pairs are I_{x1}, I_{x2} and I_{y1}, I_{y2}; the output current pair is I_{z1}, I_{z2}. The circuit of Fig. 2 yields

$$I_z = \frac{-KI_x * I_y}{I_b} \tag{4}$$

where

$$I_x = I_{x1} - I_{x2} \tag{5}$$

$$I_y = I_{y1} - I_{y2} \tag{6}$$

$$I_z = I_{z1} - I_{z2} \tag{7}$$

and in theory K=1. Since the outputs of the two differential pairs, whose currents are to be multiplied, already have the two currents I_1 and I_2, they can be directly connected to the input leads of the multiplier. However, the current feeding the capacitor needs to be "single-ended", as the difference, $I_z = I_{z1} - I_{z2}$, in which case a PMOS current mirror can be used to form the difference. Figure 3 shows I_z (as a voltage Vout measured across a resistor R=1000 Ohms, Iz = Vout/R) as simulated in PSpice using transistors suitable for BiCMOS VLSI fabrication and Ib = 1milliA, the models being the MOSIS BN2X2 of run data for Oct. 16, 1991, the key parameters being BF=130, VAF=43, IS=3.463E-16.

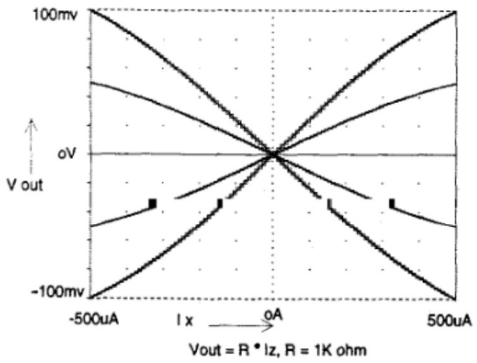


Fig. 3: Multiplication Curves of Circuit in Fig.2

3. An Example

In order to illustrate the nature of a transistorized ANN realized using the concepts presented we show in Fig. 4 a PSpice schematic for a single neuron ANN fully realized in terms of BiCMOS transistors. Any ANN with more neurons and more weights simply replicates the components shown. In Figure 4 the capacitor giving the derivative on the left of equation (1) is in the bottom middle, labeled Cneuron, with the voltage across it being the state x, referenced to ground at its bottom. The PMOS transistors at the top and the NMOS ones at the bottom give the current mirrors and current sources; those labeled Mm... are the current mirrors and those labeled Mwcs, Macs, and Mb are current sources for the weight, W of equation 1, the activation function, $g_{act}(.)$ of equation 2, and I_b of the multiplier, equation 4.

Associated with these latter are the diode connected transistors of M.s. which act as voltage dividers on the supply voltages, Vdd and Vss, to give the respective gate voltages of the current sources. The 4-quadrant current multiplier is in the upper middle left, transistors Qx1 – Qx6, the weight setting differential pair is Mw1 and Mw2, lower left, with weight setting voltage source Vweight (the sign of Vweight is the sign of the weight), and the activation function differential pair Mact1 and Mact2, lower middle right. The output differential current is fed into Rload on the far left, this going through an ideal current controlled current source in our simulations. To check out the concepts, Fig. 4 was simulated on PSpice using the MOSIS parameters given above in Section 2 for the BJTs and the corresponding MOS level 2 ones supplied by MOSIS from run N21H of 04/28/93 for which the key parameters are KPn=5.048E-5, KPp=1.908E-5, VTOn=0.858, VTOp=0.889, LAMBDAn=1.844E-2, LAMBDAp=5.012E-2. For simulation purposes all lengths and widths were chosen as 10 microns except where noted on the figure; for proper operation the largest W/L ratio needed is 4. Space precludes including the eventually satisfactory simulation results.

4. Design Considerations

As seen by equation (4) above, the key design parameter available for setting the weights is the bias current I_b. It should be noted that the tail current, set by Mwcs, for the differential pair Mw1-Mw2 must also be of this bias value; since Ix1-Ix2=Vweight* $(KP*Ib)^{1/2}$, for small enough weight voltages [2, p. 432], the weight itself is given by Vweight*$(KP/Ib)^{1/2}$. The size of the activation function is also determined by its tail current, which is set by the transistor Macs. Although MOSIS recommends the use of the BN2X2 BJT layout we found that satisfactory multiplication will work with their BN1X1 layout which uses much less area.

5. Discussion

Here we have presented a modification to the means of implementing the Hopfield class of ANNs in VLSI. This uses current mode circuits everywhere, except that the voltages on the capacitors giving the dynamics must be sampled and converted to currents. Likewise the weights are normally set via voltages, so DVCCSs are used to convert the weight setting voltages to currents. Since differences of currents are used, the weights may be of either sign as may the neural network outputs, necessitating a four-quadrant current multiplier. The circuit we chose to perform this multiplication is a BJT one due to Herpy which gives a wide range of almost ideal multiplication. The MOS version of the same circuit unfortunately gives multiplication by the square roots of the difference currents, though this may be corrected by additional diode connected MOS transistors. Because of the larger size of the BJT transistors, this MOS version should be investigated since even with the added diode connected transistors to correct for the square roots it may be much smaller in overall size in a VLSI layout. Since the multiplier takes inputs as currents, it needs to be fed by good current sources. Those in Fig. 4 may be satisfactory but could be improved by using their cascode configuration.

The example presented of course needs to be extended into an ANN which does something. That is a straightforward matter of repeating the standard components in Fig. 4. Nevertheless Fig. 4 does illustrate that any current mode realization of ANNs will undoubtedly use considerable chip area. The resultant advantage appears to be a relatively wide dynamic range as well as the suitability for handling weights of any sign.

References

[1] S. Y. Foo, L. R. Anderson, and Y. Takefuji, "Analog Components for the VLSI of Neural Networks," *IEEE Circuits and Devices Magazine*, vol. 6, no. 4, pp. 18–26, Jul. 1990.

[2] R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*. NewYork, NY: McGraw-Hill, 1990.

[3] H. P. Graf and L. D. Jackel, "Analog Electronic Neural Network Circuits," *IEEE Circuits and Devices Magazine*, vol. 5, no. 4, pp. 44–55, Jul. 1989.

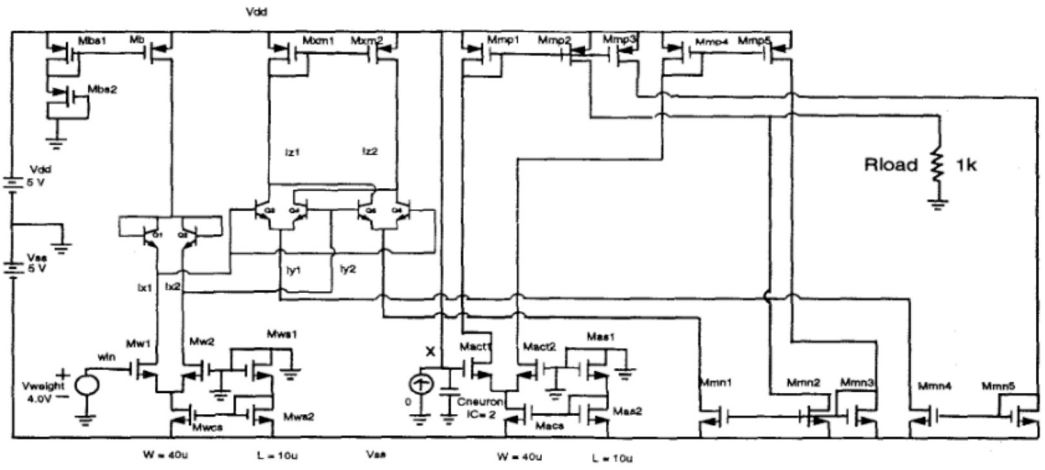


Fig. 4: Single Neuron Example ANN

[4] M. Herpy, *Analog Integrated Circuits*. Chichester: John Wiley and Sons, 1980.

[5] J. J. Hopfield, "Neurons with graded response have collective computational p roperties like those of two-state neurons," *Proc. of the National Academy of Sciences of the United States of America, Biological Sciences*, vol. 81, no. 10, pp. 3088–3092, May 1984.

[6] B. Linares-Barranco, E. Sanchez-Sinencio, and A. Rodriguez-Vasquez, "CMOS Circuit Implementations for Neuron Models," *Proc. of the 1990 IEEE International Symposium on Circuits and Systems, New Orleans*, pp. 2421–2423, May 1990.

[7] B. Linares-Barranco, E. Sanchez-Sinencio, and A. Rodriguez-Vasquez, and J. L. Huertas, "VLSI Implementation of a Transconductance Mode Continuous BAM with on Chip Learning and Dynamic Analog Memory," *Proc. of the 1991 IEEE Intern'l Symp. on Circuits and Systems, Singapore*, pp. 1283–1286, Jun. 1991.

[8] R. W. Newcomb and J. D. Lohn, "Analog VLSI for Neural Networks," *Handbook of the brain, edited by M. Arbib*. MIT Press, pp. 86–90, 1995.

[9] F. M. A. Salam, "New Artificial Neural Net Models: Basic Theory and Characteristics," *Proc. of the 1990 IEEE International l Symposium on Circuits and Systems, New Orleans*, pp. 200–203, May 1990.

[10] S. Satyanarayana, Y. Tsividis, and H. P. Graf, "A Reconfigurable VLSI Neural Network," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 1, pp. 67–81, Jan. 1992.

[11] M. Sivilotti, M. Emerling, and C. Mead, "A Novel Associative Memory Implemented Using Collective Computation," *Proc. of the 1985 Chapel Hill Conference on Very Large Scale Integration, edited by H. Fuchs*. Computer Science Press, pp. 329-342, 1985.

[12] M. J. S. Smith and C. L. Portmann, "Practical Design and Analysis of a Simple "Neural" Optimization Circuit," *IEEE trans. on Circuits and Systems*, vol. 36, no. 1, pp. 42–50, Jan. 1989.