Low Voltage-Power-Area FGMOS Neural Classifier Circuit for VLSI Analog BIST

Ganesh Lakshmana Kumar.M*,
Deptt. Of ECE,
M.A.N.I.T, Bhopal, India.

Kavita Khare,
Deptt. Of ECE,
M.A.N.I.T, Bhopal, India.

Priyanka Sharma,
Deptt. Of ECE,
M.A.N.I.T, Bhopal, India.

Abstract

This paper presents a novel Neural Classifier using FGMOS (Floating Gate MOSFET) . Basic reason for using FGMOS in Neural Classifier instead of classical MOSFET based Neural Classifier is to get significant reduction in area and power. Additional advantage of FGMOS based Neural Classifier is the simple circuitry as compared to classical MOSFET based Neural Classifier. 51.037% reduction in area is achieved in FGMOS based Neural Classifier (in .12µm technology). Along with this Neural classifier, 100 synapse and 10 Neurons reconfigurable network also implemented.

Keywords- FGMOS, Synapse, Neuron, Reconfigurable Network, Floating Gate.

1. Introduction

The use of neural classifier in the BIST architecture is illustrated in Fig. 1. The neural classifier compares the result of test stimulus and applied stimulas and is classified as a valid or invalid code-word pointing to a functional or faulty operation respectively. The focus of this paper is the neural classifier, which is a generic BIST component independent of the circuit under test.

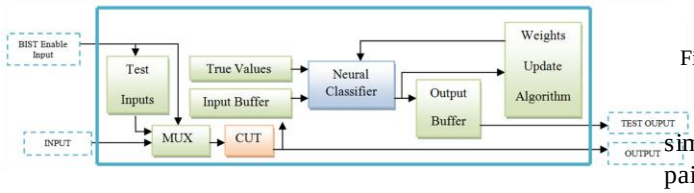

Fig. 1. BIST architecture

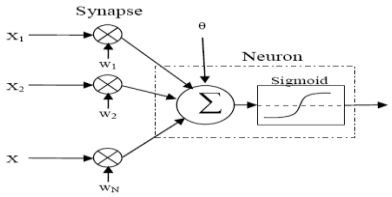

Fig. 2 Neuron and synapse models[1]

Synapse can be considered as a multiplier of an input signal value by the stored weight value. A ***Neuron*** sums the output values of the connected synapses and passes the resulting sum through a nonlinear activation function, as shown in fig. 2.

2. MOSFET Based Synapses Network design

The basic function of a synapse is multiplication. Linear multiplication, as dictated by the mathematical model, is area expensive in analog ICs.

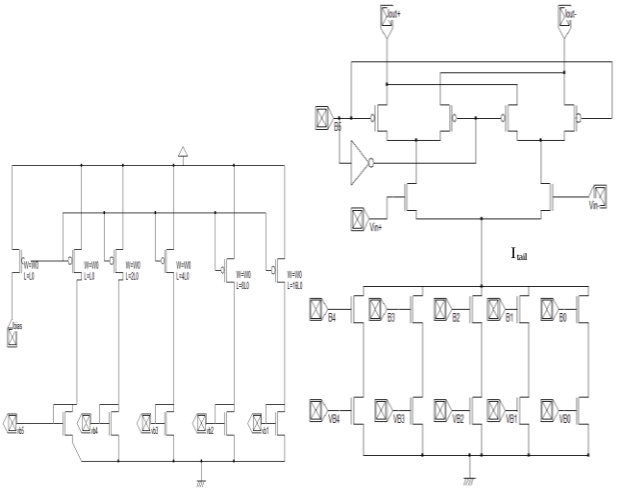

Fig. 3.(a) Current sources control circuit.(b)Synapse circuit schematic

The synapse circuit chosen for this design is a simple multiplying DAC, which represents a differential pair with programmable tail current (Fig. 3(b)) and current ratio shown in fig 3(a). A differential input voltage produces a differential output current which is collected on the summing nodes common to all synapses connected to one neuron.

3. Floating Gate Based Synapse And Weight Storage Circuits

3.1 FGMOS Weights Update Method

Fig.4 shows dynamic memory which is work idea taken from [2], [3] and Table.I gives behavior of weight update circuit. There are two modes of operation in FGMOS update method

1. Hold Mode
2. Updating Mode
   (i). Charging Mode(fig 5(b))
   (ii). Discharging Mode(fig 5(a))

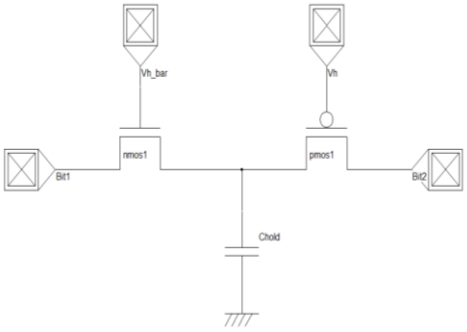

Fig.4.Weight Updating Circuit

Table I. Mode of operation

S.No	B1	B2	Mode
1	1	1	Hold
2	0	1	Charging
3	1	0	Discharging

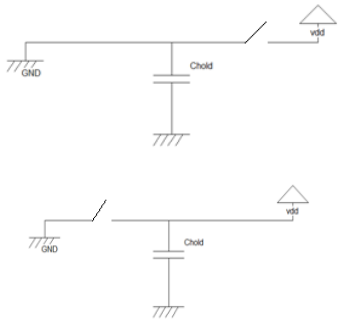

Fig.5.(a) equvalent discharging circuit (b)charging circuit

Bit2 and Bit1 are both logic signals that are either at V_{DD} or ground. V_h is either at V_{DD} or about one $|V_{TP}|$ below V_{DD} . V_{hbar} is either at ground or about one V_{TN} above ground. Only one of V_h and V_{h_bar} is active at a time.

In **Hold mode** Bit1Bit2 are at V_{DD} so both transistors are off state. In **Updating Mode,** *Charging Mode* ,Bit2 to V_{DD}, Bit1 to ground and V_h to active and equivalent mode is shown in fig.5(b). *Discharging Mode* , Bit2 to V_{DD}, Bit1 to ground and V_{h_bar} to active and equivalent mode is shown in fig.5(a). If the current is small, and the logic signals are pulsed for a short time, very small charge packets can be added to or removed from C_{hold} . Further, if the current is proportional to the change in the network error, and the logic pulse time is proportional to the learning rate, this circuit allows a very natural implementation

3.2. Synapse Based On FGMOS

Basically, this work idea is taken from reference [4], Low Power and Low Voltage FGMOS characteristics are described in reference [5]. A FGMOS based schematic of the synapse circuit with nonvolatile weight storage, which is taken from [1] as a future expansion, is shown in Fig. 6. It is a nonlinear four-quadrant multiplier with floating gate device current sources as in the electrically trainable artificial neural network (ETANN) chip [6]. Qualitatively, the circuit computes the product of the differential input voltage and the weight, which is proportional to the difference of the two floating gate device currents.

In a network, the I+ and I- nodes of many synapse circuits are tied together to perform the summing function.

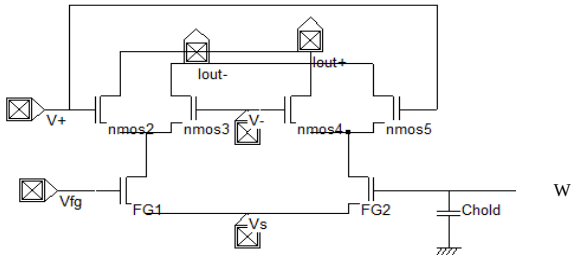

Fig.6 . FGMOS based schematic of the synapse circuit and nonvolatile weight storage

FGMOS based synapse takes inputs from the neurons in the previous layer. Weight currents always changes with FG2 and the constant current from FG1. Using a balanced weight would make the circuit behave more like an ideal multiplier, but would complicate the

weight increment circuit and increase the area of the synapse. The current in a floating gate device can be altered either by varying the control-gate voltage or by varying the charge on the floating gate. Changing the charge on the floating gate requires high-voltage pulses and may take hundreds of μs. The dynamic memory is used to store the control-gate voltage during learning. Dynamic memory allows fast alterability but requires periodic refresh cycles. After learning has converged, long-term storage is provided by storing the weight as charge on the floating gate. A more quantitative analysis of the synapse circuit follows. It can be shown that the differential output current of the synapse circuit is

$$I_{\text{diff}} = \frac{K}{2} V_{\text{in}} \left(\sqrt{\frac{4I_{F2}}{K} - V_{\text{in}}^2} - \sqrt{\frac{4I_{F1}}{K} - V_{\text{in}}^2} \right) \quad \text{for } |V_{\text{in}}| < \min\left(\sqrt{\frac{2I_{F1}}{K}}, \sqrt{\frac{2I_{F2}}{K}} \right) \tag{1}$$

where V_{in} is V_+-V_-,$K=\mu_o C_{ox} W/L$ and I_{Fi} is the current of floating-gate device . For larger input voltages, the output current saturates to the difference of the two floating gate device currents times the sign of the input voltage. The
drain current of a floating-gate device is

$$I_{ds} = \frac{K}{2}(V_{fg,s} - V_T)^2 \tag{2}$$

where Vfg,s is the floating-gate to source potential of the floating-gate device and V_T is the threshold voltage. By charge conservation, the floating-gate potential is given by

$$Vfg = \frac{Cfg,cg\,Vcg + Cfg,d\,Vd - Qf}{Cfg,cg + Cfg,b + Cfg,d + Cfg,s} \tag{3}$$

$$= Vcg\,\alpha g + Vd\,\alpha d - \frac{Qf}{Ct}$$

where $C_{fg,cg}$is the floating gate to control gate capacitance $C_{fg,d}$is the floating gate to drain capacitance, $Q_f$is the charge on the floating gate, $C_{fg,b}$ is the floating gate to substrate capacitance, $C_{fg,s}$ is the floating gate to source capacitance,α_g is the gate coupling ratio, $\alpha_d$is the drain coupling ratio, and C_T is the total capacitance seen by the floating gate.

Once learning is complete, the current of FG_2 is stored temporarily in a sample-and-hold circuit in the periphery. High-voltage pulses are applied to FG_2 until its current is equal to the target current stored in the sample-and-hold. In practice, it is difficult to achieve very high precision when programming floating-gate devices although, in theory, one electron resolution is possible. High-precision programming requires very short high-voltage pulses and a very precise sample and hold to store the target current.

The complete synapse circuit is shown in Fig.7. During recall and learn modes, the V_s node is held at ground. Weights are perturbed by pulsing V_{pert}. The change in the weight is

$$\Delta w = \Delta V_{\text{pert}} \frac{C_{\text{pert}}}{C_{\text{pert}} + C_{\text{hold}}} \frac{g_{mF}}{I_{F1}} \tag{5}$$

where g_{mF} is the transconductance of the floating-gate transistor. $C_{pert}$is a very small capacitor about 15 fF so that very small perturbations can be applied. As long as V_{pert} returns to its original level after the perturbation, the stored weight is unaffected.

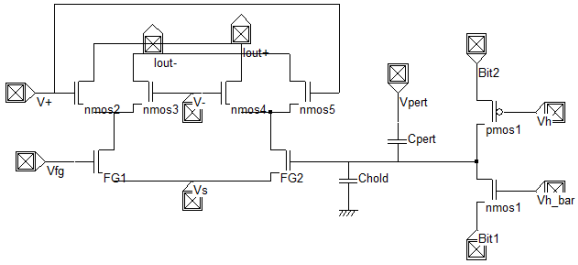

Fig. 7. Complete synapse circuit.

4. Reconfigrable Architecture

It is based on a single cascadable basic module chip which contains synapses, neuron and MUX shown in fig.8. As for the MLP (multi layer perceptron), we will map the synapse inputs and outputs, respectively, on voltages and currents. Consequently, the neuron inputs and outputs are mapped, respectively, on currents and voltages.

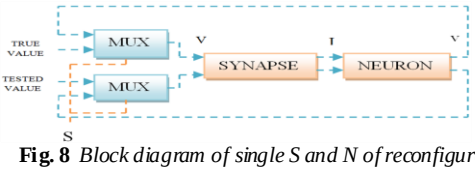

Fig. 8 *Block diagram of single S and N of reconfigurable n/w*

In the above figure select(S) pin of MUX must be high to select the input voltages. The o/p current of S is fed to N and its o/p voltages is fed back to MUX. Total 100 synapse and 10 neuron are reconfigurable as shown in fig.10.

5. Results

Fig. 9 shows that FGMOS gain area factor when compared with MOSFET based synapse circuit. With

FGMOS fewer current branches are required, and therefore the power consumption also decreases. Complete improvement is shown in Table.II. Finally, as shown reference[1] 100 synapse and 10 neurons circuit is shown in fig.10 which is a generic BIST component independent of the circuit under test (CUT).

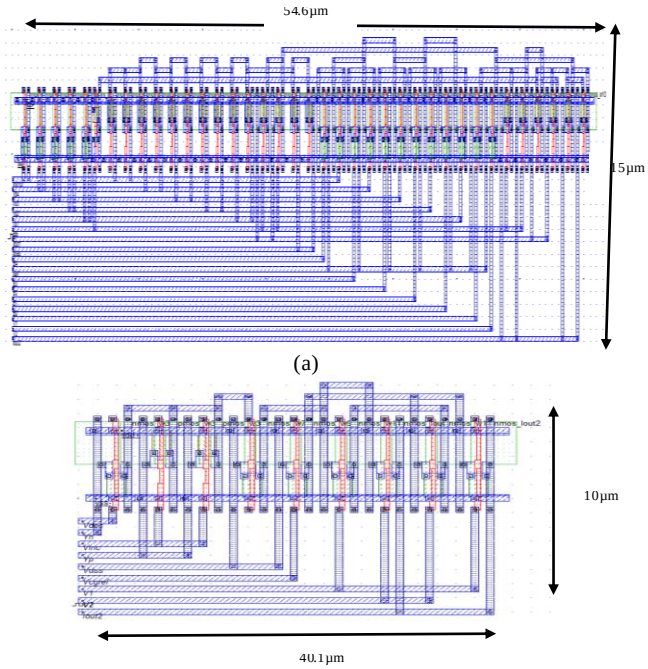

Fig .9.(a)MOSFET based Synapse Layout (b)FGMOS based Synapse

Table.II Comparation table of MOSFET and FGMOS

S.No	Parameter	MOSFET	FGMOS	%Saving using FGMOS
1.	Area	819µm^2	401µm^2	51.037 %
2.	No of elements Used	23-Nmos 4-Pmos	5-Nmos 1-Pmos 2-FGMOS	
3	Power Consumption	2.036mW	0.5102mW	74.941 %

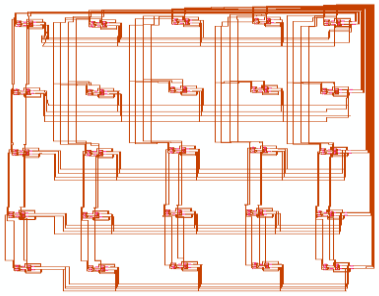

Fig.10..100 Synapse,10 Neuron based Reconfigurable network

6. Conclusions

In this paper circuit complexity of NEURAL CLASSIFIER have been significantly reduced by using FGMOS. With FGMOS fewer current branches are required, and therefore the power consumption also decreases. FGMOS is also provides reduction in area. This NEURAL Classifier has also additional benefits related to the frequency response, since the number of internal nodes is smaller. The devices can be biased in the most appropriate operating region for a wider range of input signals, by shifting the effective threshold voltages accordingly in the FGMOS transistors, hence also facilitating larger operating bandwidth,. All this benefits can be achieved without the need for extra level shifters. Finally, we can conclude FGMOS required less area when compared to the MOSFET based Neural Classifier.

7. References

[1] Dzmitry Maliuk, Haralampos-G. Stratigopoulos and Yiorgos Makris" An Analog VLSI Multilayer Perceptron and its Application Towards Built-In Self-Test in Analog C.ircuits" 2010 IEEE 16th International On-Line Testing Symposium, pp71-76,2010.

[2] C. R. Schneider and H. C. Card, "Analog CMOS deterministic Boltzmann circuits," *IEEE J. Solid-State Circuits*, vol. 28, pp. 907–914, Aug. 1993.

[3] G. Cauwenberghs, "A learning analog neural network chip with continuous-time recurrent dynamics," in *Advances in Neural Information Processing Systems 6*, J. D. Cowan, G. Tesauro, and J. Alspector, Eds. San Mateo, CA: Morgan Kauffmann, 1994, pp. 858–865.

[4] Antonio J. Montalvo, Ronald S. Gyurcsik, and John J. Paulos, "Toward a General-Purpose Analog VLSI Neural Network with On-Chip Learning" IEEE transactions on neural networks, VOL. 8, NO. 2, MARCH 1997.

[5] Esther Rodriguez-Villegas," Low Power and Low Voltage Circuit Design with the FGMOS Transistor" in IET Circuits, Devices and Systems Series 20, first published in 2006.

[6] M. Holler, S. Tam, H. Castro, and R. Benson, "An electrically trainable artificial neural network (ETANN) with 10240 'floating gate' synapses," in *Proc. Int. Joint Conf. Neural Networks*, Washington, D.C., vol. I, June 1989, pp. 191–196.