Modeling an Analog VLSI Integrated Feed-Forward Artificial Neural Network for Off-Chip Training Purposes

Axel Germann Ansgar Wego Frank Stüpmann

Silicann Electronics Gmbh Rostock www.silicann.com

Abstract

To get off-chip training techniques applicable, a good model of the VLSI hardware is required. Two methods of modeling the VLSI circuits, that is a numeric and an analytic model, are presented. The fields of application of both versions are discussed. It is shown how fabrication tolerances influence the network performance.

1 Introduction

The benefits of hardware integrated artificial neural networks (ANN) in terms of multilayer perceptrons (MLP) are well known. Particularly VLSI ANNs easily outclass software implemented neural Networks regarding response speed and compactness. But the most challenging task designing analog hardware ANNs is to get the synaptic weights adapted to the desired result (training). A fast way of training can be done in a computer. After the computer has found a set of network parameters, they can be transferred into the hardware network. This method is known as off-chip training. The scope of this article is how the performance of a hardware network can be described in a model that can be used for off-chip training.

The network consists of neurons and multipliers. The input neurons perform a simple signal conditioning. They transform the input voltages into currents in a range that can be processed by the network. The neurons in the following layers sum up the output signals of the synapses that are connected to their inputs and apply a nonlinear function to their input signals. The neurons that are connected to a chip output are called output neurons, neurons that are connected neither to an input nor to an output are called hidden neurons. The multipliers multiply the neuron outputs by synaptic weights that are stored in analogue memory cells inside the synapses.

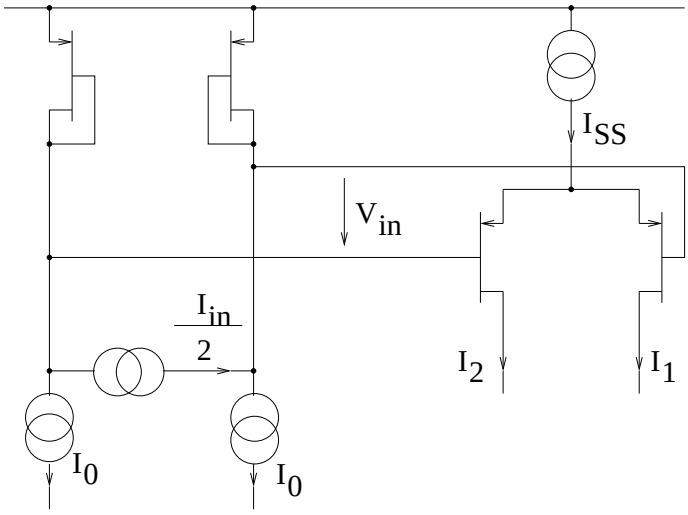

Figure 1: The hidden or output neuron (simplified)

2 Analytical model

In the beginning of the design phase, the analytical model is used to determine the ranges of the signals, to choose the types of circuits to be used, and to calculate the dimensions of the transistors in the circuits. The model is not precise and does not account for second order effects. It is fast, contains only a small amount of data, gives a good approximation of the circuit behaviour, and faciliates the choice of design parameters. It also gives an estimate for circuit tolerances.

After the design is complete, the model including the chosen parameters is transformed into a description of the circuit blocks for fast mixed-signal analysis of the entire chip.

In the input neurons, a voltage has to be converted into a current. A differential amplifier is used for this conversion. If the input voltage is V_{in} and the tail current of the differential amplifier is I_{SS}, the differential output current ΔI is [1, p. 247]

$$\Delta I = k\frac{W}{2L}V_{in}\sqrt{\left(\frac{2I_0}{kW/2L}\right) - {V_{in}}^2} \text{ with } k = \frac{\mu\varepsilon_{ox}}{t_{ox}} \tag{1}$$

if the input voltage is smaller than

$$V_{in} \leq \sqrt{\frac{2I_0}{kW/2L}} \tag{2}$$

μ is the carrier mobility, ε_{ox} and t_{ox} the isolation permittivity and thickness and W and L are width and length of a transistor.

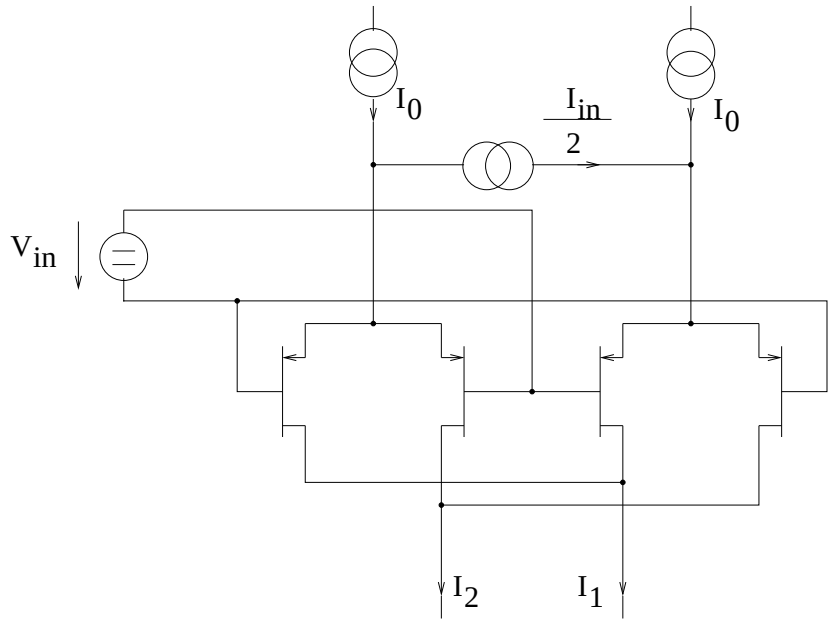

Figure 2: The multiplier (simplified)

The other neurons consist of a transistor pair that convert an input current I_{in} into an input voltage V_{in} and a differential amplifier. Figure 1 shows a hidden neuron or output neuron. The input voltage can be approximated as

$$U_{in} = \frac{I_{in}}{g_m} \text{ with } g_m = \sqrt{2k\frac{W}{L}I_o} \tag{3}$$

The synapses are modified Gilbert multipliers (Figure 2). They multiply a current by a voltage. The output current is

$$\Delta I = I_1 - I_2 = kV_{in}\left(\sqrt{\frac{2(I_0 + I_{in})}{k} - V_{in}^2} - \sqrt{\frac{2(I_0 - I_{in})}{k} - V_{in}^2}\right) \tag{4}$$

2.1 Precision

Between the stages, the signals are transmitted by means of current mirrors. For small drain currents, the offset depends mainly on the mismatch in threshold voltages ΔU_t [2]. The offset I_{off} of a current mirror is approximately

$$I_{off} = \sqrt{2}g_m\Delta U_t \tag{5}$$

It is advisable to use cascode mirrors to maximize the mirrors' output resistances. ΔU_t is inversely proportional to the square root of the transistors' areas.

The offset current of a multiplier is

$$I_{off} = 2g_m \Delta U_t \tag{6}$$

The offset of the transistor pairs at the neuron inputs is

$$U_{off} = \sqrt{2} \Delta U_t \tag{7}$$

The offset of a differential amplifier is

$$I_{off} = \sqrt{2} g_m \Delta U_t \tag{8}$$

3 Numerical Model

The numerical model is used in a network framework for training. After the network design is completed, a simulation is done for all cells and the results including the circuit tolerances are stored in table form. This is done over the whole temperature range of the chip. They are used for the numerical model.

The numerical model contains all the information about second order effects and nonlinearities that would make the analytical model cumbersome. It can be used to execute the network training in a computer and to determine the weights to be loaded into the chip.

4 Results

A chip with 10 input neurons, 6 hidden neurons and 10 output neurons was manufactured in a 0.6 μm CMOS process. As a benchmark, a real world colour classification problem with sensor data and 10 colors was chosen. A weight set that was determined off-line by backpropagation with a network model that was made from the numerical model was loaded into the chip. The hardware network performed as simulated. It is capable of classifying the colours. The maximum deviation from the ideal was 30 % for one output.

References

[1] Paul Gray and Robert Meyer. *Analysis and Design of Analog Integrated Circuits*. Wiley, New York, 3rd edition, 1993.

[2] Eric Vittoz. The design of high-performance analog circuits on digital CMOS chips. *IEEE Journal of Solid-State Circuits*, SC-20(3):656–665, June 1985.