FEEDFORWARD ARTIFICIAL NEURAL NETWORK DESIGN UTILISING SUBTHRESHOLD MODE CMOS DEVICES

COUE, Dominique Xavier Henri Leon

http://hdl.handle.net/10026.1/1633

http://dx.doi.org/10.24382/4960

University of Plymouth

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.
FEEDFORWARD ARTIFICIAL NEURAL NETWORK DESIGN UTILISING SUBTHRESHOLD MODE CMOS DEVICES

by

Dominique Xavier Henri Leon COUÉ

A thesis submitted to the University of Plymouth
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of electronic, Communication and Electrical Engineering
Faculty of Technology

June 1997
Abstract

This thesis reviews various previously reported techniques for simulating artificial neural networks and investigates the design of fully-connected feedforward networks based on MOS transistors operating in the subthreshold mode of conduction as they are suitable for performing compact, low power, implantable pattern recognition systems. The principal objective is to demonstrate that the transfer characteristic of the devices can be fully exploited to design basic processing modules which overcome the linearity range, weight resolution, processing speed, noise and mismatch of components problems associated with weak inversion conduction, and so be used to implement networks which can be trained to perform practical tasks.

A new four-quadrant analogue multiplier, one of the most important cells in the design of artificial neural networks, is developed. Analytical as well as simulation results suggest that the new scheme can efficiently be used to emulate both the synaptic and thresholding functions. To complement this thresholding-synapse, a novel current-to-voltage converter is also introduced. The characteristics of the well known sample-and-hold circuit as a weight memory scheme are analytically derived and simulation results suggest that a dummy compensated technique is required to obtain the required minimum of 8 bits weight resolution. Performance of the combined load and thresholding-synapse arrangement as well as an on-chip update/refresh mechanism are analytically evaluated and simulation studies on the Exclusive OR network as a benchmark problem are provided and indicate a useful level of functionality.

Experimental results on the Exclusive OR network and a 'QRS' complex detector based on a 10:6:3 multilayer perceptron are also presented and demonstrate the potential of the proposed design techniques in emulating feedforward neural networks.
Human intelligence is defined concretely as the ability to construct a new, unique, accurate response to each new, unique experience which confronts each human at each moment of his/her existence.

Harvey Jackins  
The Human side of Human Beings (1978)
Abbreviations

AN  Artificial Neuron
ANN  Artificial Neural Network
CMOS Complementary Metal Oxide Semiconductor
DAC Digital-to-Analogue Converter
EBP Error Back-Propagation
ECG electrocardiogram
EEPROM Electrical Erasable Programmable Read Only Memory
FGMOS Floating Gate Metal Oxide Semiconductor
FIR Finite Impulse Response
HRes Horizontal Resistance
LSB Least Significant Bit
MLP Multi-Layer Perceptron
NETSIM NETwork SIMulator
op amp operational amplifier
ON Output Neuron
PE Processing Element
RAM Random Access Memory
RMS Root Mean Square
SPICE Simulation Program with Integrated Circuit Emphasis
TS Thresholding-Synapse
TTS Transconductance-Thresholding-Synapse
VCO Voltage Control Oscillator
VLSI Very Large Scale Integration
WP Weight Perturbation
XOR Exclusive OR
Contents

1 Introduction ................................. 1
   1.1 Biological to artificial neural networks: an overview .................. 1
       1.1.1 The brain .................................. 2
       1.1.2 Sensory systems .............................. 4
       1.1.3 The effectors ................................. 5
       1.1.4 Artificial neural networks ...................... 5
   1.2 Artificial neural network models .................................. 9
       1.2.1 Synapse .................................... 9
       1.2.2 Neuron ..................................... 10
       1.2.3 A neuron and its synaptic connections ..................... 13
       1.2.4 A perceptron ................................ 14
   1.3 Neural network topologies .................................. 14
       1.3.1 Feedforward neural networks .................... 15
       1.3.2 Feedback neural networks ....................... 16
   1.4 Brief history of artificial neural networks ....................... 17
   1.5 Overview of simulation of artificial neural networks ............. 20
   1.6 Motivation for the study .................................. 25

2 VLSI implementation of artificial neural networks .................. 41
   2.1 Synapse designs .................................. 43
       2.1.1 Resistor-based synapse .......................... 43
       2.1.2 Transistor-based synapse ....................... 44
       2.1.3 Hybrid transistor/capacitor-based synapse .............. 58
   2.2 Adder designs .................................. 59
       2.2.1 Current-based summer ........................... 59
       2.2.2 Charge-based summer ............................ 60
   2.3 Activation function generator ............................... 60
       2.3.1 Squashing function based on a transimpedance amplifier .... 61
       2.3.2 Diode-based current-to-voltage converter ............... 62
       2.3.3 Thresholding generator for pulse-stream .................. 64
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4</td>
<td>Weight storage</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Digital weight storage</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Capacitive weight storage</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>EEPROM technology for weight storage</td>
<td>68</td>
</tr>
<tr>
<td>2.5</td>
<td>Summary</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>Subthreshold design for feedforward neural networks</td>
<td>92</td>
</tr>
<tr>
<td>3.1</td>
<td>Statement of problems</td>
<td>93</td>
</tr>
<tr>
<td>3.2</td>
<td>Design method</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Possible circuit implementation</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Current memory circuit</td>
<td>97</td>
</tr>
<tr>
<td>3.3</td>
<td>The conceptual building blocks of the Neural Network</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>A four-quadrant multiplier</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>The load</td>
<td>106</td>
</tr>
<tr>
<td>3.4</td>
<td>Circuit Implementation</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Combined load and transconductance-thresholding-synapse</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Speed of the network</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>Weight range</td>
<td>114</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary</td>
<td>114</td>
</tr>
<tr>
<td>4</td>
<td>Design and implementation of a neural network chip</td>
<td>128</td>
</tr>
<tr>
<td>4.1</td>
<td>The complete circuits of the neural network</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>The transconductance-thresholding-synapse</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>The load</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>The pre-processing cell</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>The output layer activation cell</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>The core of a feedforward neural network</td>
<td>137</td>
</tr>
<tr>
<td>4.2</td>
<td>Weight storage scheme</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>Capacitive storage</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>The refresh mechanism</td>
<td>141</td>
</tr>
<tr>
<td>4.3</td>
<td>Simulation of a neural network</td>
<td>145</td>
</tr>
<tr>
<td>4.4</td>
<td>The neural network chip</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>Exclusive OR network</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>QRS complex detector</td>
<td>148</td>
</tr>
<tr>
<td>4.5</td>
<td>Layout techniques</td>
<td>149</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary</td>
<td>150</td>
</tr>
</tbody>
</table>
D Analysis of a diode-based current-to-voltage converter incorporating a feed-back mechanism 252

E Switch-induced error voltage on an elementary PMOS-based sample-and-hold circuit 255

F Switch-induced error voltage on a dummy-compensated PMOS-based sample-and-hold circuit 260

G Yield analysis of the digital-to-analogue converter 265

H Published papers 268
List of Figures

1.1 Block diagram representation of the nervous system ............... 28
1.2 Typical biological neuron cell .................................... 29
1.3 Stained section of brain taken from the visual cortex of a rat ....... 30
1.4 Reconstitution of a section of a vertebrate retina .................. 31
1.5 Partial representation of an artificial neural network signal processor ...... 32
1.6 Electronic-circuit representation of a model neuron and its synaptic connections .............................................. 33
1.7 Silicon retina of Mead and Mahowald ................................ 33
1.8 Symbolic notation of a synapse .................................... 34
1.9 Symbolic notation of a neuron ..................................... 34
1.10 Activation functions ................................................. 35
1.11 A neuron and its synaptic connections ............................. 36
1.12 A perceptron .......................................................... 36
1.13 Single-layer feedforward neural network ............................ 37
1.14 Multi-layer feedforward neural network ............................. 38
1.15 Single-layer recurrent network ..................................... 39
1.16 McCulloch-Pitts model of a neuron .................................. 40
1.17 NETSIM neurocomputer .............................................. 40

2.1 Resistor-based synapse ................................................ 72
2.2 An MOS transistor synapse .......................................... 73
2.3 Static characteristics of an MOS transistor synapse ................... 73
2.4 Dual-transistor synapse .............................................. 74
2.5 Static characteristics of a dual-transistor synapse ..................... 74
2.6 Cross-coupled quad synapse ........................................... 75
2.7 Static characteristics of the cross-coupled quad ....................... 75
2.8 Differential-pair multiplier ............................................ 76
2.9 Static characteristics of the differential-pair in saturation MOS ......... 77
2.10 Static characteristics of the differential-pair in subthreshold MOS ...... 77
2.11 Sign switching cell .................................................... 78
2.12 Modified Gilbert cell .................................................. 78
5.9 Transient characteristics of the output neuron when loaded by its associated sensing scheme ........................................ 203

5.10 Variation of the TTS differential output current during the charge leakage process ................................................. 204

5.11 Effect of clock feedthrough and capacitive coupling onto the output of the TTS during weight refreshing ................................................................. 205

5.12 Transient characteristic of the refreshing mechanism ......................................................................................................................... 206

5.13 Signal flow of the error back-propagation algorithm ................................................................................................................... 207

5.14 Signal flow of the weight perturbation algorithm ................................................................................................................................. 208

5.15 ECG training patterns ................................................................................................................................................................................. 209

5.16 Learning curve of the computer simulated QRS complex detector .......................................................... 209

5.17 Experimental set-up for training and testing the analogue ANNs ................................................................................................................. 210

5.18 Learning curve of the analogue XOR neural network using the weight perturbation algorithm ................................................................. 211

5.19 Performances of the XOR network when trained using stepwise mode ................................................................................................. 212

5.20 Transient characteristics of the XOR network ................................................................................................................................. 212

5.21 Learning curve of the analogue QRS detector network using the weight perturbation algorithm ................................................................. 213

5.22 Frequency response of the linear phase FIR filter .............................................................................................................................. 214

5.23 Example of QRS detection for medium quality ECG ............................................................................................................................. 214

6.1 Transconductance-thresholding-synapse incorporating an FGMOS-based memory device ............................................................. 225

A.1 An n-channel MOS transistor .................................................................................................................................................. 244

E.1 Equivalent lumped models of the basic PMOS-based sample-and-hold circuit .................................................................................. 259

F.1 Equivalent lumped models of the PMOS-based dummy-compensated sample-and-hold circuit ............................................................................. 264
# List of Tables

4.1 The chip features ......................................................... 171

5.1 Output current offset of the individual TTSs ....................... 215
5.2 Nominal resistance of the I/V converter for various bias currents ................................................................. 216
5.3 Output voltage offset of the individual I/V converters .......... 216
5.4 Output voltage offset of the individual HResS ..................... 217
5.5 Output current offset of the individual ONs ....................... 217
5.6 DACs bit current sources ............................................... 218
5.7 Performances of the analogue and digitally simulated QRS complex detectors ...................................................... 218

A.1 List of Symbols .......................................................... 245
Acknowledgements

I would very much indeed like to thank the following people for their help and support during the course of this research:

Dr G. Wilson for suggesting the field of study, introducing me to research, and his encouragement over the last three years.

F. Fuchs who has freely given his time to many hours of discussion on the subjects of MOS transistor and clock feedthrough phenomenon and his constant stream of assistance.

All the technical and administrative staff within the department; particularly for Mr K. Jackson, Mr D. Groom and Mr A. Jerram for their support with the CAD facilities.

Dr P. Culverhouse for his valuable help in ANN and for critically reading this thesis and making suggestions for improvements.

Y. Tcha and G. Novello of the Institut Universitaire de Technologie D' Angers, Maine-et-Loire, France, for building the experimental set-ups.

Finally, I am deeply grateful to my wife Heather for reading the manuscript in full, and for making many helpful suggestions to improve the English.
Author's declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

This study was financed with the aid of a studentship from the University of Plymouth.

A relevant scientific conference was attended and the following papers published:


Signed

Date
Chapter 1

Introduction

1.1 Biological to artificial neural networks: an overview

Creativity has been an important fulfilling and driving element throughout the evolution of humankind and has, for example, progressed development of the wheel, through to modern vehicles etc. Today, we would identify the Von Neuman computer as a key 20th century invention. Although state of the art computers can now outperform biological neuron systems in executing tasks such as repetitive arithmetic operations, their abilities to perform complex functions such as image processing in real time are inadequate due to their limited processing power. In contrast to electronic computers, human (biological) neural systems seem to offer enormous computing power combined with an amazing flexibility. The desire to duplicate human computing performances artificially has been one of the principal driving forces in the quest for a greater understanding of the human neural system. The computational abilities of artificial neural systems are based on the collective operations of a dense network of inter-connected Artificial...
Neurons (ANs). As with biological systems, networks of artificial neurons draw their powerful computing abilities from the fact that they can be taught to perform desired tasks. These characteristics have led to the increasing use of artificial neural systems as a solution for a wide range of complex applications, for example, automatic speech recognition and image processing.

A description of the structure and function of the human nervous system and of the development of artificial neural networks is given in the following sections. The human nervous system has received a great deal of attention during the past century, providing an ever better understanding of its operation [1–11]. This outcome has been enhanced by the combined efforts of researchers involved in cognate disciplines, such as neurobiology, neuropsychology, physics, mathematics, computer science and electrical engineering. Due to its extreme complexity, the human nervous system may be viewed, in a rough caricature, as a three-stage system depicted in the block diagram of Fig. 1.1 [1], [12]; the brain, the receptors and the effectors.

1.1.1 The brain

The central part of the nervous system is the brain, described in Fig. 1.1 as the neural network. The term neural network has been derived from the fact that the human brain tissue is made up of a network of nerve cells also referred to as neurons. The human brain is thought to contain around $10^{11}$ of these neurons. Although, structural features may vary from cell to cell, the arrangement of a biological neuron, shown in Fig 1.2 [12], consists of three major elements:

- A cell body, called the soma;
- A long transmission-line like structure, called the axon; and
• A branching structure, comprising of what are called dendrites, where the neuron picks up signals from other neurons. In general terms, the cell body of a neuron collects incoming signals from its dendrites and aggregates them. If the sum reaches a certain level, referred to as the thresholding level, a pulse is sent (or "fired") down the axon which abuts other neurons and dendrites. The signal generated by the neuron and transported along its axon is an electrical impulse. This electrical information is subsequently passed on to other nerve cells via connecting links which are also commonly referred to as synapses. A synapse is believed to be an electrochemical transmitter, in the sense that it converts a presynaptic electrical signal into a chemical signal and then back into a postsynaptic electrical signal [2], [6]. Some synapses are excitatory in that they tend to promote firing, whereas others are inhibitory and so are capable of cancelling signals that otherwise would excite a neuron to fire. A typical neuron may receive information from anywhere between hundreds and thousands of adjacent nerve cells and in turn feeds a similar number of other neurons. This may give a rough indication of the number of synapses that make up the human brain which is understood to be in the order of $10^{14}$. Fig. 1.3 shows a stained section of brain taken from a cat cortex [1]. This section indicates that the cell bodies may be arranged in a layered fashion. Although, this arrangement suggests that the information flows principally in a forward direction, i.e. from receptors to effectors, it appears that there are also large numbers of lateral and feedback connections.

A biological neuron may be viewed as an input/output Processing Element (PE); in the sense that it collects input signals from either receptors or neighbouring neurons and produces an output impulse, depending upon both the incoming signals and the state of the synaptic connections. Based on this assumption, the brain could be viewed as a
massive parallel computer that enables human beings to understand, generate speech, recognise images, generate movements and make complex decisions involving reasoning. However, this analogy is probably too simplistic in the sense that brain mechanisms involve complex electrical and chemical relationships that are not yet fully understood [6] and that the central nervous system cannot be viewed as a whole. The human brain is a wonderfully well-organised and structured system. Investigations into the organisation of the brain have established that specific areas of the cerebral cortex are dedicated to elementary sensor and motor functions such as the Broca's area in the left frontal lobe associated with the ability to speak and Wernicke's area related to the ability to understand natural language [4], [9].

1.1.2 Sensory systems

As mentioned earlier, the central nervous system continuously receives information, processes it, and makes appropriate decisions. The principal sources of information for the brain are the receptors, which are modified nerve cells that are specialised to transforming into electrical signals the stimuli generated externally. Some sensors respond to light, others to chemicals (taste and smell) and still others to mechanical deformation (touch and hearing). It may be pointed out that our knowledge of the structure and function of these sensory neurons is more developed than that of the cells making up the central nervous system, since they are more readily accessible [1]. Fig. 1.4 shows a reconstituted section of a vertebrate retina. The structure of the retina is similar to that found in the visual cortex, see Fig. 1.3, in the sense that the inner nerve cells are organised in layers. The outer layer of the retina is made up of photoreceptors that transduce light into electrical signals.
1.1.3 The effectors

An effector may be best described as an electrochemical transmitter, in the sense that it transforms an electrical signal into a chemical signal that activates a neuromuscular junction and subsequently generates a muscular contraction [8]. The electrical signal that activates few or many effectors, depending on the muscle function, is generated by a motor-neuron whose cell body is located in the spinal cord. A motor-neuron has a similar structure to that of a neuron located in the cortex. However, some of its characteristics, such as size of the axon and velocity of nerve-impulse, may vary from one to another, depending on the muscles innervated.

1.1.4 Artificial neural networks

Research that has led to today's knowledge of the human nervous system has not only been driven by medical [10] and psychological [11] purposes but also by the fact that the brain has the ability to perform complex functions, such as thinking and reasoning, that are not yet achievable by any other means. The fact that the nervous system is able to perform such functions with ease, precision and at remarkable speed has fascinated human-kind, in much the same way that Copernicus was by the universe four centuries ago. This interest combined with that of the revolution of electricity since its discovery by the British physicist Faraday, less than two centuries ago, and the discovery of the fundamental processing element, referred to as the transistor, by the American physicist Shockley and his co-workers, no more than half a century ago, have led to the possibility of duplicating some neuronal functions artificially. The aims of the research into artificial neural systems in the past five decades have been to understand the operation of nervous functions, as well as to model their behaviour using mathematical
tools and to try to simulate those using either software on a digital computer or
dedicated integrated circuits.

The field of artificial neural system research is vast and diverse and, from an
engineering viewpoint it may be divided into three distinct areas, in much the same
way as the descriptive approach of the biological nervous system was adopted formerly;
pattern/signal processing, adaptive sensors and motor control. Although this thesis
concentrates mainly on the simulation of a specific type of pattern/signal processor, the
next sections present a succinct description of a number of important areas.

1.1.4.1 Pattern/signal processing

A pattern/signal processor based on an artificial neural system is a computational unit
that consists of a network of interconnected ANs. In the literature, such an artificial
neural system is not only referred to as an Artificial Neural Network (ANN) but also as
neurocomputing, network computation, connectionism, parallel distributed processing,
layered adaptive system, self-organizing network, neuromorphic system or network. The
properties of such ANN systems mainly depend on the adopted neuron model, the
number of interconnected PEs and the pattern of interconnection (namely, the topology).
A partial representation of such an ANN system is shown in Fig. 1.5. Furthermore, the
function that is performed by a complete ANN is also determined by the interneuron
connection strengths known as the synaptic weights. Similar to its biological
counterpart, ANN signal processors acquire the ability to perform complex tasks such as
pattern recognition [14], speech processing [15], image processing [16], etc., through
experience commonly referred to as learning. The learning procedure consists of
modifying the synaptic weights of the ANN in an orderly manner so as to attain the
desired function. Although a reasonable idea of the functions of biological neurons and
synapses is acquired, duplicating them fully may appear to be an impractical task. Many different neuron and synaptic models have been suggested but each of them either embodies a limited number of features compared with that of their biological equivalents or are highly speculative [12], [17-18]. One of the most commonly used models simulates the synaptic connection as an arithmetic multiplier weighting the neural signal passing through it. The neuron aggregates its post-synaptic signals and generates a single thresholded output. A more detailed explanation of concepts and terminology of neural network models and topologies is provided in paragraph 1.2 and paragraph 1.3, respectively.

Considerable advances in Very Large Scale Integration (VLSI) technologies have facilitated the simulation of reasonable size ANN systems [12], [18-19]. A straightforward approach consists of substituting a neuron with an analogue operational amplifier (op amp) configured as a summator. The interneuron connections are made using silicon resistors [20-21]. Fig. 1.6 shows an electronic-circuit representation of a model neuron and its synaptic connections. However, such electronic-circuit simulators suffer from some drawbacks:

- Communication between neurons is done in a synchronous manner that is not a characteristic exhibited by biological neural systems [1]; and

- The number of interconnection links that can be associated to an AN is limited due to the wiring restrictions of VLSI technologies.

On the other hand, the time response of such an electronic simulator is measured in the order of a microsecond which is approximately three orders of magnitude faster than that of its biological equivalent.
1.1.4.2 Adaptive sensors

Modelling and simulating adaptive sensory systems such as those found in the retina [22-24] and the cochlea (sense organ of hearing) [22], [25-26] have received an incredible amount of attention during the past decades. Although both of these biological stimulus transducers have distinct sensory characteristics, they are largely regarded as parallel systems. Research into artificial sensory systems has mainly consisted of developing biologically inspired models and simulating those using either software on a digital computer or a dedicated analogue VLSI chip [22-26]. However, real time simulation of artificial vision and auditory systems is computationally intensive, and is not practically achievable using software simulators even by employing today's most powerful Von Neuman machines. However this technique may still be used as a means of correlating the behaviour of a newly developed sensory model with that of its biological counterpart. On the other hand, Carver Mead and his fellow researchers at Caltech [22-23] have pioneered a new approach that consists of developing a set of analogue VLSI subcircuits that correspond to primary vision and hearing functions. At the basis of his ground breaking work Mead has exploited the exponential characteristic of Complementary Metal Oxide Semiconductor (CMOS) transistors operating in the subthreshold mode of conduction. In addition to this fundamental characteristic, power consumption is extremely low since current levels are in the range of $10^{-12}$ to $10^{-7}$ A [22-23]. Fig. 1.7 shows an electronic representation of the silicon retina developed by Mead and Mahowald [22-23].

1.1.4.3 Motor control

Research in this field ranges from restoration of movement in paralysed human limbs [27-28], to mimicking human movements using ANN based manipulators [29].
Although diverse, the basis of movement control investigations consists of determining an optimum mathematical model of motor systems that is based on experimental observation of human mechanisms of movement. However, such an empirical and theoretical model contains dynamic characteristics that are highly non-linear, making motor control systems extremely difficult to control. This computational problem may be overcome by judiciously using an adequate ANN system that can be trained to generate complex patterns of stimulation required by the motor system that produces the desired movements. This technique is usually referred to as Functional Electrical Stimulation (FES) when applied to the restoration of movement to paralysed human limbs.

1.2 Artificial neural network models

In this section, a basic definition of each function making up an ANN and its graphic illustration is given.

1.2.1 Synapse

The function performed by an artificial synapse is the mathematical operation of multiplication. In addition to this feature, a synaptic connection is also characterised by a weight. Hence, the output of a connection link is the product of its input and its weight value. The input of a synapse is either the output of a neighbouring neuron or an input to the network. Fig. 1.8 shows the symbolic notation for a synapse. Although a square box embodying a multiplication sign is an adequate graphic illustration, such a symbol may at times appear to be an encumbrance, especially when the number of synaptic
connections becomes dense. In such conditions, connection links are illustrated as arrows. The input and output of a synapse are related as

\[ U_j = W_{ij} \cdot X_j \] (1.1)

where \( X_j \) is the input signal, \( U_j \) is the output signal and \( W_{ij} \) is the synaptic weight. Since the number of synapses associated with a single neuron is usually greater than one, the weights are identified by a subscript notation. The first subscript, \( i \), refers to the neuron to which the synapse is attached and the second subscript, \( j \), identifies the input of that neuron. The property of a synaptic connection is also determined by the weight value associated to it. Hence, if the weight \( W_{ij} \) is positive the synapse is considered as excitatory, or if it is negative the synapse is regarded as inhibitory.

1.2.2 Neuron

As in its biological equivalent, an artificial neuron has several inputs (represented in number by \( N \)), and one output. The inputs of a neuron may be represented using a vector notation

\[ U \triangleq [U_1, U_2, ..., U_j, ..., U_N] \] (1.2)

where \( U \) represents the input vector and \( U_j \) is the \( j^{th} \) input. A neuron performs two mathematical operations which are, first, a linear summation function and, second, a squashing function, also referred to as an activation function.

1.2.2.1 Summation

Initially, a neuron forms the sum of its inputs expressed as

\[ S_i = \sum_{j=1}^{N} U_j \] (1.3)
This neuronal function is usually symbolically represented as an encircled capital sigma, see Fig. 1.9 (a). In a similar manner to the synaptic weight, the sum of a neuron is also identified using a subscript notation.

1.2.2.2 Activation function

The role of the activation function is to provide an output neuron signal, depending upon both the type of function and the active level of its input (output of the adder), that is limited (or squashed) between two specified boundaries. The activation function is denoted by $f(.)$ and the output of the neuron is expressed as

$$Y_i = f(S_i)$$  \hspace{1cm} (1.4)

Combining (1.3) and (1.4), the output of a PE is related to its inputs as

$$Y_i = f\left(\sum_{j=1}^{N} U_j\right)$$  \hspace{1cm} (1.5)

The squashing operation of the neuron effectively introduces a non-linearity into the PE input-output relation. This characteristic enables ANNs to model highly complex non-linear systems [12], [18]. Although a literature review demonstrates that a large number of squashing functions have been proposed, a much smaller number have been adopted. The next section introduces the three most frequently used activation functions and their properties. In these descriptions, the activation functions are presented as having a bipolar characteristic, in the sense that the function allows the response of the neuron to be either positive or negative. Although most ANN architecture or applications require bipolar neuron response, a unipolar characteristic is simply obtained by shifting and scaling the bipolar function [18].
The activation function is symbolically represented by its own functional characteristic which is, for the bulk of this thesis, a sigmoid symbol. For reasons of convenience, both summation function and activation function symbols are brought together, as shown in Fig. 1.9 (b).

1.2.2.2.1 Step activation function

The step function, see Fig. 1.10 (a), produces two output values, -1 and 1, in the following fashion:

- If the input of the of the step function $S_i$ is greater or equal to zero, then the output of the function takes on the value 1;
- Otherwise the output of the function is -1.

This type of function is mathematically described as

$$f(S_i) = \begin{cases} 
1 & \text{if } S_i \geq 0 \\
-1 & \text{if } S_i < 0 
\end{cases} \quad (1.6)$$

This type of function is also referred to as a binary function, since it produces a binary type value.

1.2.2.2.2 Ramp activation function

The ramp function, see Fig. 1.10 (b), contains three regions of operation of which two are areas of saturation. Between the saturation areas the ramp function produces an output signal that is a linear function of the active level of its input and is mathematically defined by

$$f(S_i) = \begin{cases} 
1 & \text{if } S_i \geq 1 \\
S_i & \text{if } -1 < S_i < 1 \\
-1 & \text{if } S_i \leq -1 
\end{cases} \quad (1.7)$$

12
For this particular case, within the linear region of operation, the gain of the ramp function is unity. Although the ramp function allows the output of the neuron to take on any value between the two saturation boundaries (in this case 1 and -1), like the step function it contains areas of discontinuity.

1.2.2.2.3 Sigmoid activation function

The sigmoid activation function, see Fig. 1.10 (c), is a continuous function characterised by a shape similar to an "S". It is by far the most widely exploited activation function in the design of ANNs. This is due to the fact that the sigmoid function has a monotonic characteristic which is an activation function feature required by many learning algorithms [18]. Several functions offer such a sigmoid characteristic, however the most commonly exploited is the hyperbolic tangent function, usually written as tanh. Such an activation function is represented mathematically by the following expression

\[
f(S_i) = \tanh(S_i) = \frac{1 - \exp(-2S_i)}{1 + \exp(-2S_i)}
\]  

(1.8)

Despite being a non-linear function, it may be noted that the sigmoid function, within a narrow range around the origin, has a linear approximation.

1.2.3 A neuron and its synaptic connections

In the literature, a neuron is usually represented with its associated synaptic connections as shown in Fig. 1.11. The operation performed by a neuron having N synaptic weighted inputs is obtained by combining (1.1) and (1.5) and is

\[
Y_i = f \left( \sum_{j=1}^{N} W_{ij} \cdot X_j \right)
\]  

(1.9)
1.2.4 A perceptron

The addition of a bias allows the threshold of the activation function to be varied from one neuron to another, see Fig. 1.12. The bias \( b \), is usually obtained by adding an extra synaptic connection to the neuron for which the input is set to a fixed value of 1. Thus the value of the bias is given by the value of the weight associated to the biasing connections, i.e. \( b = W_{ij} \). The operation of a neuron with its synaptic connections and bias is expressed as

\[
Y_i = \left( \sum_{j=1}^{N} W_{ij} \cdot X_j + b \right)
\]

(1.10)

The structure shown in Fig. 1.12 is the basic element that is used to build ANNs, and is usually referred to as a perceptron after Rosenblatt [30].

1.3 Neural network topologies

Neural networks may be constructed as interconnected perceptrons. ANNs are typically organised in layers in a similar way to their biological equivalents. Within a layer, neurons are fed by synaptic connections from sources which are either the outputs of neurons situated within a neighbouring layer or inputs to the network. The choice of architecture is strongly influenced by the task that the neural network has to perform [18]. Although this thesis is primarily focused on the behaviour of a particular ANN topology, namely the class of feedforward networks, a discussion of some of the most common topologies and their characteristics are presented in the following section.
1.3.1 Feedforward neural networks

A feedforward network may be described as a network where signals are exclusively propagated in a forward fashion, through layers of perceptrons until the information reaches the last level, i.e. the output layer. This type of configuration is commonly used in ANN design since it can be trained to perform pattern recognition [16], pattern classification, tasks etc. Feedforward neural networks are said to be fully connected when each neuron within a layer is connected to every neuron in the abutting forward layer. It is said to be partially connected if otherwise.

1.3.1.1 Single-layer feedforward network

The simplest form of feedforward neural network is the single-layer perceptron. The term single-layer is derived from the fact that the inputs of the network (namely, the input nodes) are projected onto one layer of neurons which is also the output layer. Fig. 1.13 shows a single-layer perceptron which contains M perceptrons which have N inputs. The fundamental characteristic of the single-layer feedforward neural network is that it is limited to the classification of linearly separable input patterns [31]. In order to increase the capacities of such networks to non-linear classifiers, one or more additional layers of perceptrons are required.

1.3.1.2 Multi-layer feedforward networks

Multi-layer feedforward configuration is the most extensively used architecture in ANN design. Such a type of architecture is also commonly known as a Multi-Layer Perceptron (MLP). A MLP is a network which combines two or more layers of perceptrons. Fig. 1.14 shows a multi-layer feedforward neural network which includes M layers. For simplicity of notation the network shown in Fig. 1.14 is referred to as a
N:P:Q: ...R MLP, where N represents the number of source nodes, P the number of neurons in the first layer, Q the number of neurons in the second layer and R the number of neurons in the output layer. It may also be noted that an additional superscript number needs to be added to the weight notation in order to associate a set of synaptic weights to a layer. As an example, synaptic weight $W_{ij}^k$ would be the weight associated to the synapse which is connected to the $i^{th}$ neuron of the $k^{th}$ layer and which takes as its source the output of the $j^{th}$ neuron of the antecedent layer. It may also be added that the layers which are located between the input layer of source nodes and the layer of output neurons are referred to as hidden layers.

This thesis focuses on multi-layer feedforward networks since they are capable of implementing within a prescribed degree of accuracy, many complex input/output mapping functions of practical interest [32]. This capability is due to the evolution of a suitable learning algorithm which is commonly known as the Error Back-Propagation (EBP). The characteristics of this learning algorithm are introduced in chapter 5. It may also be added that the output signals of such a structure, at any given time, depend entirely on the state of the input pattern and the synaptic weights.

1.3.2 Feedback neural networks

Feedback networks are also commonly known as recurrent neural networks. A recurrent network is differentiated from a feedforward type structure by the fact that one or more output signals are fed back into the network via delay units. A single-layer recurrent neural network is depicted in Fig. 1.15. The distinct characteristic of recurrent networks is that the output signals, at any instant, depend not only on the state of the input pattern and the strength of the synaptic connections but also on the internal state of the network (i.e., the level of output activity of the previous instant). Feeding the outputs back into
the network has also the effect of introducing a memory characteristic into the system [18]. However, unlike the feedforward networks where outputs are "instantaneously" generated, the output signals are only available after the network has reached a steady state. One of the most common feedback neural networks is the Hopfield scheme [33], which has a structure similar to that shown in Fig. 1.15. The characteristics of such a neural network have also attracted a lot of interest within the ANN research community.

1.4 Brief history of artificial neural networks

As mentioned earlier, the field of ANN research is vast and diverse and this has consequently led to a large amount of literature. This historical review concentrates on the most significant research developments. However, if the reader requires more historical details, additional information may be found in the articles of Grossberg [17] and Widrow et al. [34] and the Haykin [12] and Zurada [18] texts. The following summary is presented in chronological order. However, it is worth mentioning that since its early days, research into ANN has followed discontinuous paths.

1943 is usually considered as a key year in terms of ANN research. During that year, McCulloch and Pitts [35], proposed the idea of a logical calculus for modelling the nervous system. The McCulloch and Pitts model of the neuron is shown in Fig. 1.16. Although revolutionary at the time, the McCulloch-Pitts model contains some severe limitations such as:

• The inputs and output of the neuron are limited to binary type values; and

• The synaptic weights are constant and are confined to either $W_i = +1$ for excitatory synapses or $W_i = -1$ for inhibitory synapses.
Although being severely restricted, McCulloch and Pitts have shown that their neuron can perform basic logic operations such as an AND, OR and NOT. It was not long after this pioneering work, in 1949, that Hebb [36] suggested the first simple learning rule which is still in use today. The rule, known as the Hebbian learning algorithm, is formulated by Hebb as follows:

- When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes place in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased [36].

That is to say that the learning rule expresses an update of the synaptic weight as being a proportion of the correlation between its presynaptic signal $X_j$ and its postsynaptic activity $Y_i$ [12]. At that time Hebb's work had a great deal of influence in the world of neural network theory and has since been exploited by many other researchers and evolved in many directions. However, the ANN research community had to wait until 1956 to see the first computer simulation of a neural network. This work was carried out by Rochester and his colleagues [37]. The system simulated 512 neurons and made use of the Hebbian learning rule. Inspired by Hebb's work, in 1958, Rosenblatt [30] presented his theory on the perceptron; Rosenblatt's probabilistic approach, introduced a link between the learning rule of Hebb and the neuron model of McCulloch and Pitts. Subsequently, the perceptron received a considerable amount of attention. However, the excitement generated by the perceptron was soon to disappear with the publication of a book by Minsky and Papert [31] in 1969. Using a rigorous mathematical approach Minsky and Papert demonstrated the computational limits of the perceptron. These results were to leave the research community more or less at a standstill for more than a
decade. In the mean-time, in 1960, Widrow and Hoff suggested a basic neural network building block known as the ADAdaptive LINEar element also referred to as an ADALINE, along with a new powerful but simple learning rule known as the least mean square algorithm also referred to as the Widrow-Hoff learning rule [34]. The major difference between the ADALINE and the perceptron lies in their learning characteristics; for the former the adaptive algorithm requires the knowledge of a target output while in the case of the latter this information is not needed. Here lies the origin of the supervised learning algorithm.

While many researchers abandoned the field of neural network during the 1970's, Anderson [38] and Kohonen [39], published independently, their ground breaking work on a model for associative memories.

Interests into neural network research were revived, in the 1980's, due in part to the publication of Hopfield's paper in 1982 [33]. In his article, he suggested that artificial neural systems draw their computational characteristics from the collective properties of a fully connected network of PEs. His theory put an end to the scepticism brought up by the revelations of Minsky and Papert more than a decade earlier. He also presented a new descriptive approach of associative memory using differential equations. He additionally drew attention to the possibility of implementing artificial PEs using integrated circuits. Since he formulated his theories, artificial neural systems have been extensively studied and have also generated a lot of interest within the engineering community for whom ANN offers a new solution to some complex problems. As an example, Sejnowski and Rosenberg [15] have used an ANN that can be taught to convert a string of characters into a string of phonemes.
In addition to all of these influential developments, the advance of VLSI technologies have made possible the implementation of neural networks in electronic hardware. Amongst the first electronic neural network designs were those suggested by Graf and his collaborators [20] and Hopfield and Tank [21], in 1986. Their designs made use of basic electronic devices such as the resistor and the capacitor. However this was immediately followed by many other designs which were using more computationally advanced electronic devices such as the MOS transistor. Mead's book, published in 1989, presented a new view on the similarity between primary biological functions and circuits and devices based on CMOS transistors operating in the subthreshold mode of conduction [22]. It may be added that Mead's philosophy and work have been a great source of inspiration for much of the work presented in this thesis.

1.5 Overview of simulation of artificial neural networks

Within the last decade, computer scientists and electrical engineers have dedicated time to the design of ANN simulators as a means for complex problem solving. The aim of this section is to provide a broad view of the many different lines of thought that have evolved over the years.

One of the most inexpensive and readily available solutions, to the simulation of ANN functions, is based on software implemented on conventional single-processor computers. However, the superior computational power of an ANN in performing tasks such as speech processing or image processing in real time results from the parallel operation of a large number of interconnected neurons. Unfortunately, computational
loads such as these require extensive execution times on a conventional computer architecture. Even with the introduction of architecturally advanced processors such as the Digital Signal Processor (DSP) and the Reduced Instruction Set Computer (RISC) processor, leading to speed improvements of an order of magnitude every five to ten years, many real-time applications remain too demanding [18], [40-41].

To overcome this speed trap, Forrest and his colleagues [41] have advocated the use of Multiple Instruction Multiple Data (MIMD) arrays of transputers where the processing load is distributed over many processors. Their simulator consists of 40 transputers, a host acting as an overall controller and a graphics processor used as a display generator. Each transputer includes a 32-bit microprocessor with on-chip memory, inter-processor links and a 32-bit-wide external memory interface. A processor can be programmed, using a dedicated programming language, to compute a share of the processing load and to communicate with up to four of its neighbouring processors. Forest and his collaborators applied their general-purpose parallel computer to the simulation of an Hopfield neural network [21] which was trained to perform image restoration. However, the limited number of links between processors poses a special challenge to the interconnection needs of ANNs. The efficient programming of parallel processors poses further challenges. On the other hand, such a simulator offers some degree of flexibility, in the sense that the topology of the simulated ANN is easily alterable through programming and arithmetic precision is high.

Much attention is now focused on exploiting VLSI technologies for the hardware implementation of dedicated ANN simulators [42-43]. CMOS technologies offer the capability of fabricating chips with tens of millions of transistors on a single silicon die. Garth [44] and Hammerstrom [45] have suggested digital approaches, where arrays of
identical modules, consisting of an arithmetic processor with sufficient local memory to simulate moderate size single layer networks, operate in parallel. For example, Garth's simulator consists of a 3-dimensional array of interconnected NETwork SIMulator (NETSIM) cards and a host computer, see Fig. 1.17. Each NETSIM card is an autonomous element with sufficient local memory and processing power to simulate a single-layer network of 256 neurons with 256 synaptic connections per neuron (i.e., 65,536 synapses) in less than 20 millisecond. A NETSIM card contains:

- A communication chip which enables the fast exchange of messages between cards or the host and a card;
- A local microprocessor;
- The solution engine designed to act as a co-processor which accelerates the computation of the extensive number of multiply-and-add operations required by ANNs; and
- Some memory to store data such as the synaptic weights, the input vectors, etc.

The access of the memory is pipelined, allowing low-cost 120 nano-second Dynamic Random Access Memories (DRAMs) to be used.

This approach results in a low-cost, efficient and fast means of simulating ANNs. The X1 neural network chip of Hammerstrom is also based on a similar approach. However it is worth noting that neither neural network simulators fully exploit the parallel processing and fault tolerance characteristics of ANNs [21] since synaptic and neuronal operations are distributed in time. Alternative digital approaches [43], [46] which do offer such features are based on a systolic array of processing units (i.e., one multiplier per synapse and one adder and activation function generator per neuron). However, due
to the digital complexity and the area of silicon required for the multiplying operation.
The number of processing units which can be implemented on a silicon die is limited.
For example, the GENES IV chip of Lenne and Viredaz [46] integrates a matrix of $2 \times 2$
processing units. This small number is mainly due to the extended arithmetic precision
of a processing unit (e.g., 17 bits). However Murray [43] indicates that even utilising a
reduced precision arithmetic approach, the number of elements which can be integrated
on a silicon die is considerably limited (<100). Furthermore, digital approaches also
involve relatively high power consumption since it is proportional to the square of the
operating frequency. On the positive side however, digital systems have a greater
immunity to noise, operate at high speed (frequency in excess of 100 MHz), offer high
levels of precision and ease of weight storage.

The simulation of large numbers of interconnected neurons, and therefore synaptic
links, in a parallel systolic manner can be achieved by implementing the processing
elements using an analogue approach [19-26], [47-69]. Successful integration of large
numbers of synaptic connections and neuron functions on a single silicon die demands
that each element uses the smallest possible area. One analogue approach [20-21],
requires an op amp for each neuron and simulates synaptic connections using simple
resistors. Unfortunately, such systems cannot easily be programmed. Proposed solutions
[19], [22-26], [47-69] have exploited the ease with which the transconductance of MOS
transistors can be changed by modifying the device's bias point. These implementation
techniques may be classified in distinct categories depending upon the mode of
conduction of the transistors [72-74]:

- Artificial neural functions can be implemented using building blocks
  consisting of MOS devices operating in the above-threshold mode of operation
Here the designers exploit the fundamental quadratic function of the transconductor.

- An alternative approach simulates ANN functions exploiting the exponential characteristics of MOS transistors operating in the subthreshold mode of conduction (also known as weak inversion) [61-69].

Within these two sub-sections of ANN design many implementation techniques have been suggested ranging from the single transistor synapse [47-48] to the MOS version of the Gilbert transconductance multiplier, which utilises six transistors [70]. Although, analogue techniques are well suited to the implementation of artificial neural networks, issues such as dynamic weight storage [71], arithmetic precision [67-68], and noise immunity, pose significant challenges.

A hybrid approach which combines analogue and digital technologies has been suggested by Murray and his colleagues [75-77] and subsequently investigated by other groups [78]. The pulse-stream technique uses digital signals to both carry information and to control analogue circuitry, allowing the compactness of analogue computation to be effectively coupled with the simplicity and robustness of digital signals. However, one such design [76] requires 100 pulses to drive a neuron fully "on" or "off". With a pulse frequency conservatively set at 0.5 MHz, a two-layer perceptron would settle in around 3 ms, which is approximately three orders of magnitude slower than alternative analogue techniques [54], [58].

The capacities of previously published analogue and hybrid designs that emulate ANN functions systolically are further detailed in chapter 2.
1.6 Motivation for the study

The still emerging revelation that ANNs are applicable to a wide range of problems coupled with the growing availability of an already extended range of simulators is leading to an increasing use of artificial neural systems in many different fields and especially in the fertile areas of medicine and healthcare. ANN systems are particularly capturing interest in medical applications which are directly connected to improving the quality of human life. For example, several research groups are working on the problem of restoring movement to paralysed human limbs [27-28] using ANN based functional electrical stimulators. For such medically related problems patients could further benefit from the design of an implantable system. Such an approach has already been suggested for the design of cardioverter defibrillators providing an on-line heart therapy to patients suffering from life-threatening arrhythmia [79] which task is similar to the detection of heart beats, i.e. QRS complexes, of foetal electrocardiograms [80].

In chapter 3, design of multi-layer feedforward neural systems which exploit the low current levels associated with MOS devices operating in subthreshold mode are investigated. Such designs have continued and will continue, to receive much attention, because they satisfy the fundamental characteristics required by implantable neural network systems:

- Low power consumption enables a battery power source;
- Compact analogue circuits allow systolic implementation, hence facilitating pattern classification in real time;
- Computation is carried out in the analogue domain resulting in simplified interfaces to outside systems such as sensors and effectors; and
• Analogue parallel architecture exhibits robust performance in the presence of hardware faults.

Although subthreshold operation exhibits valuable characteristics, it also lacks strength in some other areas:

• Subthreshold currents have reduced abilities to charge/discharge capacitive elements, hence the operating speed of ANN systems is degraded [66], [69], [72];

• MOS devices operating in the weak inversion are affected by mismatching effects [65], [81-83]. These effects are characterised at the circuit level by offsets, hence limiting arithmetic precision;

• High weight storage resolution (at least 8 bits) is unobtainable in practice without a trade-off in silicon area;

• Subthreshold currents are vulnerable to thermal noise; and

• The transconductance of an MOS transistor operating in the weak inversion is strongly sensitive to temperature [72-74] and biasing variations.

To address some of these issues, a novel building block called the thresholding-synapse which exploits the non-linearity associated with the input of transconductance multipliers to perform both the activation function and synaptic multiplication is presented. A new four-quadrant multiplier is developed in order to efficiently simulate this thresholding synapse, complemented by an equally original current-to-voltage converter. Critical discussions on the effectiveness of the proposed schemes to emulate neuronal functions are derived from analysis and simulation results. The unit's potential in terms of speed and ease of programmability is also provided.
Based on these suggested arrangements the implementation of a whole neural network chip which comprises a 2:2:1 feedforward network, to solve the Exclusive OR (XOR) problem as well as to check the AN operations, and a QRS complex detector, founded on a 10:6:3 MLP, is then the subject of chapter 4. An extra two processing cells, namely the horizontal resistance and output neuron,27 are introduced in order to alleviate problems associated with the effect of combining the thresholding and synaptic functions. Structural details of all of these basic building blocks and their vulnerability to process parameter variations are presented. Further discussions are also devoted to the issue of an adequate weight storage scheme and tailoring of an on-chip weight update/refresh mechanism to suit it. In order to evaluate the performance of the proposed system as a whole, simulation studies on the XOR neural network as a benchmark problem have also been conducted.

Experimental results from each of the primitive structural elements are subsequently presented in chapter 5. For each experiment the set-up and chip configuration are provided. The concept of learning is also discussed in this chapter. The suitability of the back-propagation and weight perturbation learning algorithms for chip-in-loop training is then evaluated using illustrative examples. The experimental set-up used for training and testing the networks is then presented. Finally the performance of the two low-power analogue VLSI ANNs is provided and compared with that of a digital simulator.
Figure 1.1: Block diagram representation of the nervous system.
Figure 1.2: Typical biological neuron cell. Adapted from Haykin [12].
Figure 1.3: Stained section of brain taken from the visual cortex of a rat. Adapted from Hubel [1]. The numbers on the right-hand side identify cellular layers; the capital letters label individual neurons.
Figure 1.4: Reconstitution of a section of a vertebrate retina. Adapted from Spooner [13]. There are three layers of cells; the outer receptors, R, the internal bipolar cells, BC, and the inner ganglion cells, G.
Figure 1.5: Partial representation of an artificial neural network signal processor.
Figure 1.6: Electronic-circuit representation of a model neuron and its synaptic connections.

Figure 1.7: Silicon retina of Mead and Mahowald. A single pixel element is illustrated in the circular window.
Figure 1.8: Symbolic notation of a synapse.

Figure 1.9: Symbolic notation of a neuron. (a) expanded form, (b) compacted form.
Figure 1.10: Activation functions. (a) Step function, (b) Ramp function, (c) Sigmoid function.
Figure 1.11: A neuron and its synaptic connections.

Figure 1.12: A perceptron.
Figure 1.13: Single-layer feedforward neural network.
Figure 1.14: Multi-layer feedforward neural network.
Figure 1.15: Single-layer recurrent network.
Figure 1.16: McCulloch-Pitts model of a neuron.

Figure 1.17: NETSIM neurocomputer.
Chapter 2

VLSI implementation of artificial neural networks

Since the beginning of the 1970's the evolution of VLSI technologies has mainly been driven by an increasing demand for more computationally evolved digital systems. During the past decade, growth has, however, been more noticeable due to the maturing of analogue VLSI. Developments in both analogue and digital VLSI and MOS techniques have recently attracted interest in the field of ANN simulation for several reasons:

- State of the art VLSI technologies offer the capability of fabricating chips with tens of millions of transistors on a single silicon die [22];

- An MOS transistor is a powerful computational element. Several of these devices can be combined to form advanced processing circuits such as multipliers and adders;
• Today's VLSI processes allow for both analogue and digital circuits to be integrated on the same silicon chip;

• VLSI technologies are well suited to the highly parallel, regular, and modular architecture of artificial neural systems;

• Computer-Aided-Design (CAD) tools for design and simulation of analogue and digital MOS circuits are well developed; and

• VLSI technologies are more readily accessible to the public.

All these evolving characteristics make VLSI technologies increasingly better candidates for the implementation of artificial neural systems.

As stipulated in the introduction, to possess the fundamental features of parallel processing and fault tolerance, ANN hardware simulators require the use of a design architecture which is based on a systolic approach. To fulfil such a design condition, it is highly desirable, if not essential that each element making up a neural network simulator occupies a minimum silicon area. Such an issue arises if one considers the implementation of Sejnowski-Rosenberg's MLP [15], since 18,629 synaptic connections and 106 neuronal functions are required, or when designing an implantable system [79].

This chapter reviews design techniques that have been developed as a means of simulating ANN systems in a systolic manner. Hence, these include a family of reported schemes based on either analogue [19], [22-26], [47-69] or analogue/digital [75-78] (i.e. hybrid) design methods. These building blocks are presented in the following sections depending upon which of the three basic operations it is simulating namely, synaptic weighting (multiplication), signal summation and activation functions. These various cells are implemented using basic electronic devices such as resistors, capacitors, op
amps and MOS transistors. The electronic characteristics of the resistor, capacitor and op amp can be found in the literature [84] while a brief description of the operation of an MOS transistor is presented in Appendix A.

The overall objective of this review is to discuss and evaluate the performances of previously published computational circuits based on both theoretical and SPICE [85] (Simulation Program with Integrated Circuit Emphasis) simulation results. All the SPICE simulations presented in this chapter have been generated using the parameters of the MIETEC 2.4 µm CMOS process which are presented in Appendix B.

Finally the last section of this chapter presents an overview of the issue of weight storage.

2.1 Synapse designs

The synapse is by far the most important building block in the hardware implementation of ANNs. Using Sejnowski-Rosenberg's MLP as a design example, the number of synaptic circuits would represent 99.4% of the total number of cells. This illustration gives a clear indication to why researchers involved in the design of ANN simulators have continued, and will continue, to dedicate attention to the development of efficient synaptic cells. The function of a synapse is to produce an output signal which is the result of the product of the synaptic input signal and a weight. Both input and weight signals can be of either sign resulting in a four-quadrant multiplication.

2.1.1 Resistor-based synapse

The most straightforward synaptic design is based on Ohm's law [84]. Ohm's law states that the current flowing through a resistor is related to the product of the differential
potential across that resistor and its conductance, thus offering the multiplying feature required by the synaptic element. Fig. 2.1 shows a resistor and its equivalent synaptic representation. One of the terminals is connected to a virtual ground so that the differential potential across the resistor is $V_x$ which also corresponds to the input signal of the synapse. The current flowing through the resistance, illustrated in Fig. 2.1 as $I_u$, is usually selected as the output signal of the synapse since it simplifies the implementation of the summing function. Consequently the conductance $G$ determines the value of the weight as the resistor-based synapse function is expressed as

$$I_u = G.V_x$$  \hspace{1cm} (2.1)

The advantage of such a technique is that it allows the integration of large neural networks, since high resistor density is achievable using dedicated material such as amorphous silicon (approximately one resistor per 10 $\mu$m²) [20]. However one of the major drawbacks is that once a circuit is fabricated, the function of the network is not easily alterable since the system does not allow for reprogrammability. As a consequence the design has no learning capability. Furthermore, the mapping characteristics of a network may vary from one chip to another. These variations are mainly due to the effect of limited accuracy of integrated resistors. Moreover a resistor-based synapse does not allow for an inhibitory type of connection, unless the driving neuron provides for inverting and non-inverting outputs [20-21].

2.1.2 Transistor-based synapse

Learning in artificial neural systems requires a synaptic circuit that permits the weight to be externally adjustable. Such an increase in flexibility is achieved in VLSI technology by exploiting the ease with which the conductance or transconductance of MOS
transistors can be altered [47]. The techniques discussed in this section utilise MOS devices as either voltage-controllable resistors or electronically adjustable transconductors.

2.1.2.1 Single-Transistor Synapse

An MOS transistor operating in the linear mode of conduction (Appendix A) may be viewed as a voltage-controllable two-terminal resistor [47-49]. This may be illustrated by considering the drain current flowing through an N-channel MOS device

\[ I_D = K_N \cdot \left( (V_{GS} - V_{FB} - \phi_B)V_{DS} - \frac{1}{2}V_{DS}^2 \right) 
\begin{align*}
- \frac{2}{3} \gamma \left( (V_{ds} - V_{BS} + \phi_B)^{\frac{3}{2}} - (\phi_B - V_{BS})^{\frac{3}{2}} \right)
\end{align*}
\]

(2.2)

where \( V_{gs} \) is the gate-to-source potential, \( V_{ds} \) is the drain-to-source potential, \( V_{bs} \) is the bulk-to-source potential, \( V_{fb} \) is the flat-band voltage, \( \phi_B \) is the surface potential and \( \gamma \) is the body effect coefficient (Appendix A). The so-called transconductance factor is given as

\[ K_N = \mu_N C_{ox} \frac{W}{L} \]

(2.3)

where \( \mu_N \) is the electron mobility, \( C_{ox} \) is the thin oxide capacitance per unit area, and \( W \) and \( L \) are the channel width and channel length of the transistor respectively.

Utilising a binomial expansion technique, (2.2) can be expanded in Taylor's series resulting in the following expression

\[ I_D = K_N \cdot (V_{GS} - V_T) \cdot V_{DS} + h(V_{DS}) \]

(2.4)

where the first term is linearly dependent on \( V_{ds} \) and \( h(V_{DS}) \) represents all the non-linear terms of the drain current.
Considering the synaptic design shown in Fig. 2.2 and assuming that the input signal \( V_X \) is small enough so that the contribution of the non-linear terms compared to that of the linear term are negligible, then the output current may be expressed as

\[
I_U = K_N (V_W - V_T) V_X
\]  

(2.5)

When this is related to (1.1) it can be deduced that the value of the synaptic weight is given by \( W_{ij} = K_N (V_W - V_T) \) which is tuneable via the control potential \( V_w \). However to ensure that the synaptic circuit performs as desired, it is necessary that both source/bulk and drain/bulk junctions of the transistor be reversed-biased which requires that

\[
V_B \leq V_{X_{\text{min}}}
\]  

(2.6)

The family of simulated static characteristics of Fig 2.3 shows the output current versus the input signal over the range -1 V... 1 V for different values of \( V_w \), indicating the practicability of an MOS transistor as a variable-synaptic element. These simulations were performed with a bulk potential of -1 V, in order to fulfil the proviso of (2.6), and the channel width and channel length of the transistor set to 2.4 \( \mu \)m and 24 \( \mu \)m respectively. These results show that the linear characteristic of the resistor deteriorates significantly as the control signal decreases toward the threshold voltage. This non-linearity leads to substantial multiplication inaccuracies which are undesirable during the learning process. With reference to (2.4), the distortion is due to the higher order terms in \( V_{DS} \) contained in \( h(V_{DS}) \).

2.1.2.2 Dual-Transistor Synapse

The extent of the distortion can be appreciated by considering the second order approximation to the drain current of an NMOS transistor operating in the so-called linear mode of conduction
This second-order approximation is sufficient since the effect of the third and subsequent higher orders are relatively small. It may be noted that the square term contributes significantly to the drain current when the drain-source potential approaches the non-saturation conduction limit given by

\[ V_{DS} = V_{GS} - V_T \]  \hspace{1cm} (2.8)

Several linearisation techniques have been suggested [48-49]. Here a dual-transistor synapse is used as an illustrative example [48]. The scheme is shown in Fig. 2.4 and its kernel is essentially a pair of matched devices which are driven by a control signal \( V_w \) and are fed by a balanced differential input voltage \( v_x = V_x - (-V_x) \). Note that both output terminals share the same reference potential \( V_{Ref} \). In practice this virtual short-circuit is achieved using an op amp.

The current flowing through each transistor can be expressed as follows

\[ I_{M1} = K_N \left[ (V_w - V_{Ref} - V_T)(V_x - V_{Ref}) - \frac{1}{2}(V_x - V_{Ref}) \right] \]  \hspace{1cm} (2.9)

and

\[ I_{M2} = K_N \left[ (V_w - V_{Ref} - V_T)(-V_x - V_{Ref}) - \frac{1}{2}(-V_x - V_{Ref}) \right] \]  \hspace{1cm} (2.10)

Thus the difference output current of the synapse is given by

\[ i_U = I_{M1} - I_{M2} = K_N (V_w - V_T) v_x \]  \hspace{1cm} (2.11)

This expression shows that, at the expense of an additional device, the linearity is considerably improved. The simulation results shown in Fig. 2.5 further demonstrate the improvement.
Although such a structure enhances the performance of synaptic multiplication, it is still lacking the fundamental bipolar weight feature and has the additional inconvenience of requiring a balanced differential input signal.

2.1.2.3 Cross-coupled quad synapse

To achieve a synaptic connection which offers a bipolar weight characteristic implies the use of a four-quadrant multiplier. This feature is readily obtainable by combining a pair of two-quadrant multipliers. Such a design technique has been frequently used in analogue VLSI technology. If one considers the dual-transistor arrangement as a basic building block, then the circuit presented in Fig. 2.6 is a direct product of such a design approach. This scheme is commonly known as the cross-coupled quad [57-60] and consists of four matched devices $M_1-M_4$ driven by input voltages $V_x^+ = V_x + \frac{\nu_x}{2}$ and $V_x^- = V_x - \frac{\nu_x}{2}$, and control signals $V_w^+ = V_w + \frac{\nu_w}{2}$ and $V_w^- = V_w - \frac{\nu_w}{2}$. Where $V_x$ is the common-mode input level, $\nu_x$ is the dynamic input voltage, $V_w$ and $\nu_w$ are the quiescent and dynamic control potentials, respectively. It can be shown that, utilising (2.7), the difference output current of both dual-transistor arrangements $DT_1$ and $DT_2$ are given by

$$i_{U1} = I_{M1} - I_{M2} = K_N \left( V_w + \frac{\nu_w}{2} - V_x - V_T \right) \cdot \nu_x$$ (2.12)

and

$$i_{U2} = I_{M4} - I_{M3} = K_N \left( V_w - \frac{\nu_w}{2} - V_x - V_T \right) \cdot \nu_x$$ (2.13)

Thus the difference output current of the synapse is expressed as

$$i_U = i_{U1} - i_{U2} = (I_{M1} + I_{M3}) - (I_{M2} + I_{M4}) = K_N \cdot \nu_w \cdot \nu_x$$ (2.14)
Since the synaptic output is independent of the threshold voltage of the transistors, the cross-coupled scheme can be implemented using either enhancement or depletion devices of either n- or p-channel thus permitting its integration into a wide range of CMOS processes. It is also interesting to note that the cell is symmetrical, hence allowing for permutability between the input and output nodes. Lehmann and Bruun [60] have exploited this valuable characteristic to minimise hardware requirements when implementing the back-propagation learning algorithm in analogue ANNs. Finally, balanced input and weight signals are not required.

The family of simulated static characteristics of Fig 2.7 shows the output current versus the input signal over the range -1 V... 1 V for various differential weight values. These simulation results confirm that, for an aspect ratio of $W/L = 2.4\mu m / 24\mu m$ for all transistors and a bulk potential of -1 V, the cross-coupled arrangement offers the possibility of programming weight values of either signs. It may be seen that the transfer characteristics in the first and second quadrants are asymmetrical and that the error reaches approximately 15%. This is mainly due to higher order non-linearities in the drain currents (i.e. cubic and quadratic terms in $V_{ds}$). However the major drawback is the need for an op amp, in order to provide for the virtual output short-circuit, which can occupy substantial silicon area.

### 2.1.2.4 Differential pair synapse

As mentioned above the role of the op amp, in synaptic designs based on MOS transistors acting as voltage-controllable two-terminal resistors, is to maintain a pair of nodes at the same potential. This section demonstrates how this requirement can be avoided when utilising the MOS device as a transconductor operating either in the saturation or subthreshold mode of conduction.
At the basis of their synaptic designs several researchers have exploited the versatile transconductor-based differential pair multiplier shown in Fig. 2.8. This structure is also referred to as the long-tail pair and divides a biasing current $I_w$ between a pair of transistors $M_1$ and $M_2$ as a function of the difference between the input voltages $V^+_x = V_x + \frac{v_x}{2}$ and $V^-_x = V_x - \frac{v_x}{2}$. Where $V_x$ and $v_x$ are the common-mode and dynamic input levels, respectively. Note that independent of the regime of operation the tail and the drain currents are related by

$$I_w = I_{M1} + I_{M2} \quad (2.15)$$

### 2.1.2.4.1 A differential pair in saturation MOS

The voltage-to-current transfer characteristic of the long-tail pair based on saturated devices can be determined utilising the simplified square-law given in Appendix A by

$$I_D = \frac{K_n}{2} (V_{GS} - V_T)^2 \quad (2.16)$$

If the two transistors making up the pair have well matched features, their drain currents are expressed as

$$I_{M1} = \frac{K_n}{2} \left( V_x + \frac{v_x}{2} - V_S - V_T \right)^2 \quad (2.17)$$

and

$$I_{M2} = \frac{K_n}{2} \left( V_x - \frac{v_x}{2} - V_S - V_T \right)^2 \quad (2.18)$$

where $V_S$ represents the common-source node potential. Combining (2.15), (2.17) and (2.18), the difference output current is given by

$$i_u = I_{M1} - I_{M2} = \sqrt{K_n I_w \cdot v_x \cdot \sqrt{1 - \frac{K_n v_x^2}{4 I_w}}} \quad (2.19)$$
This expression identifies the inherently non-linear nature of the transfer characteristic.

This can be assessed as follows. Substituting (2.19) into (2.17) and (2.18) the drain currents of the devices can be rewritten as

\[ I_{M1} = \frac{K_N}{2} \left( \frac{v_x}{2} + \sqrt{\frac{I_w}{K_N} \cdot \left( 1 - \frac{K_N v_x^2}{4 I_w} \right)} \right)^2 \] (2.20)

and

\[ I_{M2} = \frac{K_N}{2} \left( \frac{v_x}{2} + \sqrt{\frac{I_w}{K_N} \cdot \left( 1 - \frac{K_N v_x^2}{4 I_w} \right)} \right)^2 \] (2.21)

Thus it may be noted that these expressions are valid as long as \( v_x \) do not exceed the operating range given by

\[ -\sqrt{2} \cdot \frac{I_w}{\sqrt{K_N}} < v_x < \sqrt{2} \cdot \frac{I_w}{\sqrt{K_N}} \] (2.22)

If so it can be seen that the difference output current saturates at \(-I_w\) and \(I_w\) as appropriate. However for input signal well within this range of operation, (2.19) may be approximated to a first order as

\[ i_U = \sqrt{K_N I_w} v_x \] (2.23)

It can be appreciated that the difference output current is independent of the drain-to-source potentials of \(M_1\) and \(M_2\), as long as these do not reach the minimum saturation limit \(V_{ds}\) given in (2.8), thus allowing for a resistive load to be directly connected to the synapse output [50-52], [69] (i.e. without the need of an impedance adaptor such as an op amp). However the saturation conditions expressed in (2.22) are dependent upon the weight control signal \(I_w\). This brings to the fore the fundamental problem of the
differential pair in saturation MOS which is that the smaller the weight value the narrower the input operating range. This is illustrated in the graph of simulated static characteristics depicted in Fig. 2.9.

2.1.2.4.2 A Differential pair in subthreshold MOS

With the assumption that the drain-to-source quiescent potential is greater than four times the thermal voltage \( V_T = kT/q \) (approximately 26 mV at room temperature), but low enough to disregard the Early effect, the drain current of an NMOS transistor biased in the subthreshold region is related to its gate-to-source potential (Appendix A) as

\[
I_D = I_X \exp \left[ \kappa \frac{V_{GS}}{V_T} \right] \tag{2.24}
\]

where \( \kappa \) is a measure of the effectiveness of the gate in controlling the channel current and \( I_X = I_{DO} W/L \), where \( I_{DO} \) is the characteristic current of the transistor. Thus the currents flowing through a differential pair of matched devices can be expressed as

\[
I_{M1} = I_X \exp \left[ \kappa \frac{V_X + \frac{V_X}{2}}{V_T} \right] \exp \left[ -\kappa \frac{V_S}{V_T} \right] \tag{2.25}
\]

and

\[
I_{M2} = I_X \exp \left[ \kappa \frac{V_X - \frac{V_X}{2}}{V_T} \right] \exp \left[ -\kappa \frac{V_S}{V_T} \right] \tag{2.26}
\]

Combining these two expressions and (2.15), it can readily be shown that the sum of the device currents at the common-source node is given by

\[
I_W = I_X \exp \left( \kappa \frac{V_X - V_S}{V_T} \right) \left[ \exp \left( \kappa \frac{V_X}{2V_T} \right) + \exp \left( -\kappa \frac{V_X}{2V_T} \right) \right] \tag{2.27}
\]
Hence the difference output current \([22], [79]\) can be formulated as

\[
i_u = I_w \cdot \frac{\left(\exp\left(\frac{\kappa \cdot V_x}{2V_t}\right)\right) - \exp\left(-\frac{\kappa \cdot V_x}{2V_t}\right)}{\left(\exp\left(\frac{\kappa \cdot V_x}{2V_t}\right) + \exp\left(-\frac{\kappa \cdot V_x}{2V_t}\right)\right) + \tanh\left(\frac{\kappa \cdot V_x}{2V_t}\right)} = I_w \cdot \text{tanh}\left(\frac{\kappa \cdot V_x}{2V_t}\right)
\]  

(2.28)

As in the saturation mode of conduction, the transfer characteristic has a non-linear response with respect to the differential input and saturates at \(|I_w|\) if \(|V_x|\) is greater than a few \(V_t\) (e.g. \(|V_x| \approx 6V_t\)). Nevertheless given that the tanh function can be expanded in series as

\[
\tanh(x) = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \frac{17}{315}x^7 + \ldots \text{ for } |x| < \frac{\pi}{2}
\]  

(2.29)

then for \(|V_x| \leq V_t / \kappa\), (2.28) can be approximated to a first order as

\[
i_u = \frac{\kappa \cdot I_w}{2V_t} \cdot V_x
\]  

(2.30)

Thus, biased in the weak inversion, the long-tail pair also operates as a linear transconductance multiplier for input voltage less than \(V_t / \kappa\). Although the linear range of operation is smaller than when the scheme is operated in the saturation region it is however independent of the biasing current, as can be seen in the family of simulated static characteristics shown in Fig. 2.10. It may also be deduced from (2.30) that, as with saturated devices, the scheme can be directly loaded \([22],[67],[79]\). However the fundamental advantage of subthreshold operation over the saturation mode of conduction is that power dissipation is diminished since current and biasing voltage levels are typically lower.
2.1.2.5 Multiple differential pair synapse

The differential pair is however a two-quadrant multiplier and can therefore only provide unipolar weight values. Different design techniques have been suggested to extend this feature to a four-quadrant multiplication.

The first approach is based on a long-tail pair loaded by a sign switching cell which is made up of four switching transistors $M_1-M_4$ controlled by a binary value $B_i$ [51], [79] as shown in Fig. 2.11. The difference output current of the complete structure is then given by

$$I_{out} = I_{out}^{+} - I_{out}^{-} = \begin{cases} I_{in}^{-} & \text{if } B_i = 1 \\ -I_{in}^{+} & \text{if } B_i = 0 \end{cases}$$

(2.31)

where $I_{in}$ is the difference output current of a long-tail pair biased either in the strong or weak inversion. The difficulty with this technique is that the weight is controlled by both an analogue and digital signal.

The second approach consists of coupling two differential pairs $DP_1$, $DP_2$ which share the same input voltages $V_{x1}$ and $V_{x2}$, and are biased by $I_w^{+} = I_w + \frac{i_w}{2}$ and $I_w^{-} = I_w - \frac{i_w}{2}$ respectively. Where $I_w$ and $i_w$ are the quiescent and dynamic components of the biasing currents, respectively. This scheme is also known as the modified Gilbert multiplier and is shown in Fig. 2.12. In the saturation regime [47], [52] the transconductance characteristic of this cell can be obtained utilising (2.19) and is expressed as

$$i_U = i_{DP1} - i_{DP2} = \frac{1}{\sqrt{K_N.I_w}} \cdot \left( \sqrt{1 + \frac{i_w}{2} - \frac{K_N.v_x^2}{4.I_w}} - \sqrt{1 - \frac{i_w}{2} - \frac{K_N.v_x^2}{4.I_w}} \right) \cdot v_x$$

(2.32)

which simplifies to
\[ i_U = \sqrt{K_NI_w} \left( \sqrt{1 + \frac{i_w}{2I_w}} - \sqrt{1 - \frac{i_w}{2I_w}} \right) \cdot v_x \]  

(2.33)

for \(|v_x|\) much smaller than \(\sqrt{(2I_w/K_N)}\). Where \(i_{dp1}\) and \(i_{dp2}\) represent the difference output current of the differential pair DP\(_1\) and DP\(_2\) respectively. While for devices operated in the subthreshold region [22], [67] the output current

\[ i_U = i_w \cdot \tanh \left( \frac{v_x}{2V_t} \right) \]  

(2.34)

is deduced substituting \(i_{dp1}\) and \(i_{dp2}\) by (2.28), where it has been assumed that all devices have identical features. The substantial advantage of subthreshold over saturation mode is that the synaptic weight is controlled in a linear fashion. The static behaviour of this multiplier operated in strong and weak inversion is depicted in Fig. 2.13 and 2.14 respectively. Note that, when biased in saturation mode (2.32), the linearity range of \(v_x\) can be widened at the expense of increasing the quiescent component of the biasing currents.

The alternative and by far the most commonly used synaptic design, shown in Fig. 2.15, exploits the MOS version of the Gilbert cell [70] which utilises an additional long-tail pair to provide for the biasing currents of the preceding scheme. The fixed biasing current \(I_g\), is then distributed between the differential pairs DP\(_1\) and DP\(_2\) via a third differential pair DP\(_3\) in a manner determined by the difference weight voltage \(v_w = V_{w1} - V_{w2}\). Operated in the saturation mode of conduction [53-55], the synapse output current can be derived by combining (2.20), (2.21) and (2.32) and is formulated as

\[ i_U = \frac{K_N}{2} \left( \sqrt{\frac{2I_B}{K_N} v_w^2 + \frac{v_w}{\sqrt{2}}} \right)^2 \cdot v_x \]  

(2.35)
which is non-linearly related to \( v_x \) and \( v_w \) in an interactive fashion. Nevertheless, for small input and weight signals \( (2.35) \) can be approximated as

\[
i_u = \frac{K_N}{\sqrt{2}} \cdot v_w \cdot v_x
\]  

(2.36)

Note that the linearity range of the grounded differential pair is fixed and defined by the value of \( I_b \) and \( K_N \), while the synaptic weight is only related to the transconductance factor of the transistors. Thus the operating range can be adjusted at will without affecting the connection strength. Combining \((2.34)\), and \((2.28)\), the difference output current of the Gilbert structure becomes

\[
i_u = I_B \cdot \tanh \left( \frac{v_x}{2V_t} \right) \cdot \tanh \left( \frac{v_w}{2V_t} \right)
\]  

(2.37)

when biased in the subthreshold region [22], [62-64], [68-69]. If the signal swing of both \( v_x \) and \( v_w \) is limited to \( |V_t| \), \( \tanh(x) \) approximates to \( x \) \((2.29)\) and the Gilbert cell then behaves as a linear transconductance multiplier since its transfer characteristic \((2.37)\) simplifies to

\[
i_u = \frac{I_B \cdot \kappa^2}{4V_t^2} \cdot v_w \cdot v_x
\]  

(2.38)

The static characteristics shown in Fig. 2.16 and 2.17 confirm that irrespective of the mode of conduction, the Gilbert cell generates a difference output current related to the product of two differential voltages which may be used to represent the input and weight signal of a synapse. The major difference is that in the saturation mode, Fig. 2.16, the linearity range of the inputs can be adjusted to fit a given design, while in the weak inversion, Fig. 2.17, it is fixed and substantially smaller.
2.1.2.5.1 Wide range Gilbert multiplier

From the above analysis and simulation results, it would appear that the Gilbert multiplier biased either in strong or weak inversion offers an ideal solution to the simulation of artificial synapses. Unfortunately the circuit is constrained by some biasing conditions which are mainly due to the fact that its design is based on a stack of long-tail pairs. One of the solutions to this problem is to isolate the top and bottom levels of the structure utilising a pair of current mirrors as shown in Fig. 2.18 [22], [53-54]. Although the wide range Gilbert multiplier virtually eliminates the biasing constraints of the Gilbert cell, it creates further inconveniences namely:

- The weight and input stages of the multiplier are integrated using opposite channel devices, thus requiring the need for a twin well process;

- The power dissipated by the modified scheme is twice that of the Gilbert circuit since the sum of currents flowing in both current mirrors is $2.1g$;

- Compared to the original scheme, the number of transistors is increased by 70%, therefore considerably reducing the number of synaptic connections that are implantable on a given area of silicon; and

- Since the number of transistors to be matched is greater, the wide range Gilbert cell will typically generate larger offset errors.

In the next chapter it will be shown that it is possible to eliminate these effects whilst maintaining a similar transfer characteristic by utilising an arrangement in which the bias current is distributed across two differential pairs by modulating the bulk potentials of the devices.
2.1.3 Hybrid transistor/capacitor-based synapse

The multiplying operation performed by the synaptic connection can also be simulated using an hybrid transistor/capacitor technique. One approach is based on the fact that a charge \( Q_u \) stored on a capacitor is related to the product of the difference potential \( V_x \) across that capacitor and its capacitive value \( C_w \) which is given by

\[
Q_u = C_w V_x \tag{2.39}
\]

Here the charge is chosen as the output of the synapse in order to facilitate the implementation of the summation function. This issue is discussed in the next section. It may be noted that this technique is similar to that used for the resistor-based synapse with the exception that the output signal is a quantity of electrical charge instead of a current. Although the relationship between \( Q_u \) and \( V_x \) is valid at any time, the synapse is simulated using a switched-capacitor structure, as shown in Fig. 2.19 (a), which exploits the charge conservation principle [86-88]. The transistor switches are controlled via a pair of two-phase clock signals, as depicted in Fig. 2.19 (c). When the switch associated with \( \phi_1 \) is closed, a charge packet, with magnitude proportional to \( V_x \), is stored on the capacitor which is then transferred to the output of the synapse during the clock phase \( \phi_2 \), see Fig. 2.19 (c) for more details. Since \( C_w \) is regarded as the weight, an array of weighted capacitors is required in order to emulate variable connection strengths. The additional disadvantage is that only positive weight values are simulated, unless a complementary opposite sign structure as shown in Fig. 2.19 (b) is adopted.

The alternative solution to the former problem is to view the switched-capacitor circuit of Fig. 2.19 (a) as a conductance of value

\[
G_c = C_w f \tag{2.40}
\]
where \( f \) is the switching frequency of the control signals which can therefore be used to modify the strength of a synaptic connection. However it has been reported [75-76] that it is more elegant to utilise the switching speed as the synaptic input when the post-neuronal states are represented as sequences of pulses modulated in time, thus limiting the number of on-chip Voltage Control Oscillators (VCO) to a minimum. The weight is then represented as the input voltage of the switched-capacitor scheme.

However all of these switched-capacitor techniques suffer from the clock feedthrough phenomenon [89] which occurs during the turning-off transient of an MOS switch.

### 2.2 Adder designs

Most of the synaptic schemes presented in the literature are developed such that their output signals are either currents or charges. This choice is primarily made for the straightforward reason that the summation function can be achieved simply by connecting the synapse outputs to a common bus bar. Thus such an implementation technique leads to a compact summer circuit since it does not involve the use of additional components.

#### 2.2.1 Current-based summer

Kirchhoff's current law stipulates that the sum of all the currents flowing into a node is equal to the sum of all the currents flowing out of that same node. This theorem is at the basis of the current summer circuit shown in Fig. 2.20. The currents \( I_{u1}, I_{u2}, \ldots, I_{uj}, \ldots, I_{UN} \) which enter the node are the output signals of the synaptic connections linked to an
analogue AN. Thus, applying Kirchhoff's principle, the output of the summer is given by

\[ I_s = \sum_{j=1}^{N} I_{uj} \] (2.41)

where \( N \) represents the number of connecting links associated to an AN.

### 2.2.2 Charge-based summer

A charge-based summer exploits the charge conservation principle that an electrical charge \( Q \) cannot be either created or destroyed and that, at any given time, the sum of all the charges stored in an electrical node is equal to zero. Applying these basic electrical concepts to a neuron circuit involving connected switched-capacitor synapses, the quantity of electrical charge flowing out of the summing node, similar to the current summer scheme depicted in Fig. 2.20, is then given by

\[ Q_s = \sum_{j=1}^{N} Q_{uj} \] (2.42)

where \( Q_{uj} \) is the output charge produced by the \( j^{th} \) switched-capacitor synapse, during the switching phase \( \phi_2 \) (see Fig. 2.19 (c) for details).

### 2.3 Activation function generator

The role of the activation function generator is to contain (squash) its weighted-sum input between two specified limits and supply an output signal which is compatible with the input levels required by the synapses it feeds. As a consequence its design is dependent upon the type of pre- and post-synaptic signals and the model of the squashing function which can be either a step, a ramp or a sigmoid function of either
unipolar or bipolar characteristics. Although applications exploiting the inherent properties of a fully connected feedforward neural network require far less activation function generators than connection links, it is however still desirable that the silicon area associated to each squashing building block, as determined by the number and size of electronic devices, be minimised in order to maximise implementation efficiency. This feature becomes a requirement when reconfigurability is involved, since an identical number of synaptic and thresholding building blocks are needed [53].

2.3.1 Squashing function based on a transimpedance amplifier

When the design of the interconnection link is based on the principle of either a fixed or variable conductor, the characteristic of the thresholding generator is that of a non-linear resistor, since the signal provided by the summing junction is a current and the signal required by the input of the synapse is a voltage. The non-linearity in the transfer function of the resistive element represents the squashing function. It may also be noted that such a current-to-voltage converter needs to provide low input and output impedances so that the multiplying characteristics of the synapse are unaltered whatever the number of connections it is either fed by or supplying to, thus allowing for a large fan-in and fan-out. All these features are displayed by a transimpedance amplifier scheme [20-21], [47-49], [54], [57-58]. Fig. 2.21 depicts the structure of a transimpedance amplifier, which is associated with the cross-coupled quad multiplier, consisting of an op amp and four MOS transistors operating as variable two-terminal resistors. When the op amp is operating in its linear range the transfer characteristic of such a scheme can be obtained utilising (2.7), since the matched transistors (M1-M4) are biased in their triode region, and is given by
\[ \nu_Y = \frac{i_S}{K_N(V_{C1} - V_{C2})} \]  

(2.43)

where \( i_s \) is the difference input current produced by the current summer. The gain of the current-to-voltage conversion can be adjusted via the difference control potential \( \nu_c = \nu_{C1} - \nu_{C2} \). The transfer characteristic displays a saturation behaviour, thus providing the thresholding required by an AN, whenever \( i_s \) attempts to generate an output signal greater than the supply voltage of the op amp. This is illustrated in the graph of simulated static characteristic depicted in Fig. 2.22.

Although a complete multi-input programmable analogue MOS vector multiplier offers a feasible solution to the implementation of multi-layer feedforward ANNs, it is not efficient since the silicon area associated with the analogue neuron can be of an order of magnitude larger than that of the synapse [58].

### 2.3.2 Diode-based current-to-voltage converter

When utilising long-tail pairs, biased either in the saturation or subthreshold mode of conduction, as the basis of a synaptic design the characteristic of the thresholding function generator is also that of a non-linear resistor for the same reasons as stated above. However, the differential pair possesses the following valuable features:

- High input impedance, because the difference signal is applied to the gates of the devices; and

- Low output conductance, since the drain-to-source voltages of the transistors have a minute influence in the multiplying characteristics;

thus, the current-to-voltage converter does not need to be buffered in order to allow for large fan-in/fan-out. As a consequence the non-linear load can be implemented using a simple CMOS arrangement of biased diodes as shown in Fig. 2.23 [53], [56], [64]. Note
that, for a fully differential scheme, the resistor consists of four matched transistors $M_1$-$M_4$ and two biasing voltage sources $E_1$, $E_2$. The transfer characteristic of such a scheme can be determined assuming that the current flowing through a diode connected transistor is exponentially related to its gate-to-source potential as in (2.24). Thus the output difference voltage in Appendix C is given by

$$v_Y = V_Y^+ - V_Y^- = \frac{2V_t}{\kappa}.\sinh^{-1}\left(\frac{i_s}{4.1_X}\right)$$

where $I_X$ is the quiescent bias current and is given by

$$I_X^+ = I_X.\exp\left[-\kappa.\frac{V_{dd} - (E_1 + E_2)}{2V_t}\right]$$

where $V_{dd}$ is the supply voltage.

Because the drain current of the device is an exponential function of the gate-to-source voltage, the I/V characteristic of the resistor is highly sensitive to the biasing arrangement. This may be illustrated as follows. Given that the sinh$^{-1}$ function can be expanded as

$$\sinh^{-1}(x) = x - \frac{1}{2.3}x^3 + \frac{1.3}{2.4.5}x^5 - \frac{1.3.5}{2.4.6.7}x^7 + \ldots \ldots \quad \text{for } |x| < 1$$

Then for $|i_s| \leq 2I_X$ the effective driving point resistance may be approximated to the first order as

$$r = \frac{v_Y}{i_s} \approx \frac{V_t}{2.\kappa.I_X} = \frac{V_t}{2.\kappa.I_X.\exp\left[-\kappa.\frac{V_{dd} - (E_1 + E_2)}{2V_t}\right]}$$

It can therefore be appreciated that the resistance is highly sensitive to variation in both the supply and bias sources. The simulation results depicted in Fig. 2.24 clearly highlights this undesirable effect. It can be seen that within the supply range 2.9 to 3.1 V
the nominal resistance of the scheme varies by at least 60%. It may also be added that
the non-linearity associated with the I/V characteristics is not necessarily the ideal
squashing function. The next chapter will demonstrate how these two issues have been
tackled.

2.3.3 Thresholding generator for pulse-stream

The role of the thresholding function circuit, when the synapse array exploits a
switched-capacitor design, is to convert, in a non-linear fashion, a sum of charges into a
stream of pulses whose frequency is modulated by the magnitude of that sum. This has
been achieved by converting the charge packets into a correlated analogue voltage,
utilising a leaky integrator whose design exploits a standard op amp integrated into a
resistive/capacitive negative-feedback loop, and then subsequently transforming that
voltage into a frequency modulated signal using a VCO [75-76]. Such an electronic
system is depicted in Fig. 2.25. Note that the feedback resistor $R_s$ can be implemented as
a switched-capacitor resistor driven by a global clock signal (common to all neurons),
thus allowing the gain of the generator to be externally controlled. The non-linearity is
introduced either by clamping the voltage across the feedback resistor or utilising a
VCO which embodies a sigmoid transfer characteristic [77]. However, within its linear
region the output voltage of the leaky integrator is defined by the following differential
equation

$$\frac{\partial V_Y}{\partial t} = \frac{V_Y}{R_s C_s} - \frac{1}{C_s} \sum_{j=1}^{N} Q_{uj} f_j$$

(2.48)

where $Q_{uj}$ is the amount of charge, per clock cycle, delivered by the $j^{th}$ synapse, $f_j$ is the
frequency of the clock (i.e. pre-synaptic input) and $R_s$ and $C_s$ are respectively the
feedback resistor and capacitor.
Due to the nature of the above expression, the system requires five times the time constant \( \tau_s = R_s C_s \) to settle to an output steady state activity given by

\[
V_Y(t) = -R_s \sum_{j=1}^{N} Q_{u_j} f_j
\]  

(2.49)

thus limiting the computational speed of an ANN. It should be emphasised that due to the transient behaviour of the synapse, a ripple signal, superimposed onto the output of the leaky integrator, is generated and has a magnitude given by

\[
V_{Y\text{ripple}} = V_w \frac{C_w}{C_s}
\]  

(2.50)

where \( V_w \) is the weight voltage of the switched-capacitor connection and \( C_w \) is its associated capacitor. This undesired variation has the effect of limiting the processing accuracy of the neuron. Both expressions (2.48) and (2.50) encapsulate the fundamental compromises between speed and accuracy of computation of pulse stream neural networks based on a switched-capacitor design.

### 2.4 Weight storage

As mentioned earlier, the characteristics of a neural network are determined by its topology and the weights associated with each connection. This section will only deal with issues related to weight. As mentioned earlier, the strength of each weight is estimated during the learning procedure. The abilities of the network to use these weights in a subsequent application is achieved by storing their values utilising a non-volatile memory circuit. The building block utilised to perform this task is usually referred to as either the weight storage cell or the long-term memory scheme. It may be noted that as one of these cells is associated with each synapse, it is desirable that its
silicon area be minimised allowing for a systolic simulation of ANNs. Since learning is an intensive iterative process, whereby the weight values can be altered many times [12], [18], [34], [91-92] (i.e. depending upon the type of algorithm), it is then essential that the speed at which the data is written in the memory cell be high in order to allow for rapid training. To implement this feature either for an on-chip or an off-chip system, it is also recommended that the memory cell offers a linear characteristic which may lead to a simpler and more compact weight storage controlling circuitry. The storage scheme should also offers a weight resolution of at least 8 bits in order to allow convergence during learning [93-94]. Utilising the above mentioned criteria, the following subsections will assess the viability of some storage circuits which have been developed in relation to the synaptic designs described earlier.

2.4.1 Digital weight storage

One straightforward weight storage design is based on a binary weighted array of devices. The elements can be either resistors [49] or capacitors [88] or current sources [50-51], [79] depending on the principle upon which the synapse is designed. Note that this technique has been inspired by the memory system used in a digital computer. A memory scheme based on weighted current sources is depicted in Fig. 2.26. For this example, transistors $M_1$ to $M_n$ behave as current sources with magnitudes determined by the biasing voltage $V_{bias}$. Transistors $S_1$ to $S_N$ acts as switches which are controlled by binary signals stored on non-volatile digital memory cells. The output current of the circuit is given by

\[ I_w = b_1 + 2b_2 + 4b_3 + \ldots + 2^{N-1}b_N = \sum_{i=1}^{N} 2^{i-1}b_i \]  

(2.51)
where \( b_i \) is either "0" or "1" according to the binary state of \( B_i \), \( I \) is the unit current source and \( N \) represents the number of bits. Note that the number of transistors needed to implement the digital-to-analogue converter circuitry alone (i.e. without including the digital memory cells) is \( 2^N + N -1 \). In the case of learning rules which require high weight resolution (i.e. at least 8 bits), this inherent feature mitigates against the small area condition required for a practical synaptic connection. Furthermore the size of those devices must also be large enough to obtain a monotonic conversion [82], even though mismatching effects can in practice be reduced by taking advantage of a common centroid layout [19]. To optimise the silicon area associated with digital storage cells, a design based on shift registers has been suggested [50-51]. However such schemes suffer from low learning speeds since the weight data are processed serially.

### 2.4.2 Capacitive weight storage

A basic capacitive sample/hold system can be used as a storage scheme [47], [52-54], [56], [58], [67-69], [75-76]. Its structure is shown in Fig. 2.27, and consists of a switching transistor \( M_s \) which is controlled by a switching signal \( \phi \) and a hold capacitance \( C_{\text{hold}} \). In order for the quantum of stored charges to remain fixed during the holding time the sensory circuit (i.e. synapse) must provide for a high impedance. In practice this can be achieved by connecting the capacitor to either the gate or bulk of an MOS device. Although most mixed analogue/digital processes offer the possibility of integrating poly-poly capacitors, the cell area may be significantly reduced by implementing \( C_{\text{hold}} \) as an NMOS device \( M_{\text{hold}} \) for which the bulk, drain and source terminals are grounded. This is justified by the fact that the capacitance is inversely proportional to the thickness of the dielectric which is that of the gate oxide. However
this approach is subject to one requirement which is that the difference potential across the device must, at all times, be greater than $V_r$ in order for it to be biased in strong inversion [74]. Provided that this condition is fulfilled, then the gate of the transistor acts as the top plate, the inversion layer as the bottom plate and the value of the capacitor may be approximated by

$$C_{\text{hold}} = W_{\text{hold}} \cdot L_{\text{hold}} \cdot \frac{\varepsilon_{\text{ox}}}{t_{\text{ox}}}$$ \hspace{1cm} (2.52)

where $W_{\text{hold}}$ and $L_{\text{hold}}$ are respectively the width and the length of $M_{\text{hold}}$, $t_{\text{ox}}$ is the thickness of the insulator and $\varepsilon_{\text{ox}}$ is its permittivity.

The basic sample/hold circuit suffers from two major drawbacks:

- During the holding time, some stored charge (information) is lost due to leakage occurring into the substrate via the reverse-biased source/drain-to-substrate diodes of the switching transistor, and also through subthreshold conduction.

- The stored charge is perturbed during the switch-off transient of the access transistor $M_a$.

These issues will be thoroughly discussed in chapter 4.

2.4.3 EEPROM technology for weight storage

An alternative approach to the implementation of a non-volatile, alterable, compact weight memory cell is to use the programmable threshold-voltage characteristic of either charge trapping or floating gate MOS transistors [71], [95-96]. This technology was originally developed for the design of Electrical Erasable Programmable Read Only Memory (EEPROM). These devices operates on the principle that charges trapped in
one or more extra layers of insulated gate situated between the control gate and the channel effectively produce a shift in the threshold voltage \( V_T \). Thus the current flowing through a Floating Gate MOS (FGMOS) device is determined by the amount of charge deposited on the floating-gate as well as the potential applied on the control gate, drain, source and substrate. Fig. 2.28 shows the cross-section of an FGMOS transistor (charge injector mechanism is not shown). Note that such a structure can be fabricated utilising a standard double-poly CMOS process wherein poly-1 and poly-2 layers are respectively utilised to act as the floating and controlling gate [96]. FGMOS transistors can be grouped in different categories depending upon the type of mechanism used to inject or remove charge from the floating gate. The most common methods [71] are:

- Avalanche injection; and

- Fowler-Nordheim tunnelling injection.

The process of charging and discharging a floating gate device is energy intensive. To obtain an avalanche breakdown phenomenon from a reverse biased n-p junction, a voltage in excess of +30 V is required, whereas tunnelling either electrons or holes from the channel to a floating gate through a thin oxide necessitates pulses of ±15 V of amplitude. Although tunnelling is compatible with today's low voltage VLSI technologies, the shift in \( V_T \), created by a train of constant pulses, is proportional to the logarithm of the number of pulses and the characteristics of the charging and discharging procedure are diverse. This strong non-linearity makes the weight value extremely difficult to program accurately. On the positive side however, since leakage through the oxide insulator is extremely low, the retention time of programmed devices can range from 1 to 10 years.
2.5 Summary

In this chapter, the most common analogue and hybrid designs used to simulate the primary functions of ANN have been discussed.

It has been shown that synaptic programmability can be obtained utilising transistors acting either as voltage controllable resistors or as adjustable transconductors biased either in the subthreshold or saturation mode of conduction. To provide for signal adaptation, the resistor-based synapses have, however, the disadvantage of requiring an op amp which is area hungry. The alternative solutions, which are based on combinations of differential pairs, remove this burden. Analysis and simulations revealed that the Gilbert multiplier and its derivative may be programmed to behave either as an inhibitory or as an excitatory synapse and that the saturation mode of conduction offers more control over signal range than subthreshold operation, which is an advantage acquired at the expense of an increase in power dissipation since current and biasing voltage levels are typically higher. However the Gilbert cell is liable to some biasing conditions which are costly, in power dissipation and silicon area, to overcome.

The switched-capacitor synapse as a variable conductor is compact; however it lacks computing resolution due to clock feedthrough phenomenon and its associated squashing generator bargains speed against accuracy of computation and is area hungry since it involves at least one op amp.

It has also been established that when utilising synaptic schemes based on either analogue or hybrid designs the summation function is simple to implement since no electronic devices are required.
Weight storage schemes based on designs such as digital memory, capacitive element and EEPROM technologies respectively suffer from area inefficiency, limited retention time and non-linear programmability.
Figure 2.1: Resistor-based synapse. (a) electronic representation, (b) equivalent synaptic model.
Figure 2.2: An MOS transistor synapse.

Figure 2.3: Static characteristics of an MOS transistor synapse. $I_U$ versus $V_X$. 
Figure 2.4: Dual-transistor synapse.

Figure 2.5: Static characteristics of a dual-transistor synapse. $i_u$ versus $v_x$ for $(W/L)_{m_1} = 2.4\, \mu m / 24 \, \mu m$ and a reference potential of 0 V.
Figure 2.6: Cross-coupled quad synapse.

\[ V_W^+ = V_W + \frac{V_W}{2} \]
\[ V_W^- = V_W - \frac{V_W}{2} \]
\[ V_X^+ = V_X + \frac{V_X}{2} \]
\[ V_X^- = V_X - \frac{V_X}{2} \]

\[ i_U = (I_{M1} + I_{M3}) - (I_{M2} + I_{M4}) \]

Figure 2.7: Static characteristics of the cross-coupled quad. \( i_U \) versus \( v_X \) for \( (W/L)_{M1,2,3,4} = 2.4\mu m / 24\mu m \), a reference potential of 0 V, a quiescent input signal of 0 V and common-mode control signal of 3 V.
Figure 2.8: Differential-pair multiplier.
Figure 2.9: Static characteristics of the differential-pair in saturation MOS. $i_U$ against $v_x$ for $(W/L)_{M1,2} = 2.4\mu m/6\mu m$ and a common-mode input signal of 3V.

Figure 2.10: Static characteristics of the differential-pair in subthreshold MOS. $i_U$ against $v_x$ for $(W/L)_{M1,2} = 24\mu m / 2.4\mu m$ and a common-mode input signal of 1V.
Figure 2.11: Sign switching cell.

Figure 2.12: Modified Gilbert cell.
Figure 2.13: Static characteristics of the modified Gilbert multiplier in saturation MOS. $i_U$ against $v_X$ for $(W/L)_{M1,2,3,4} = 2.4 \mu m / 6 \mu m$, a quiescent biasing current of 5 $\mu A$ and a common-mode input voltage of 3 V.

Figure 2.14: Static characteristics of the modified Gilbert multiplier in subthreshold MOS. $i_U$ against $v_X$ for $(W/L)_{M1,2,3,4} = 24 \mu m / 2.4 \mu m$, a quiescent biasing current of 100 nA and a common-mode input voltage of 1 V.
Figure 2.15: MOS version of the Gilbert multiplier.
Figure 2.16: Static characteristics of the Gilbert multiplier in saturation MOS. $i_U$ against $v_x$ for $(W/L)_{M1,2,3,4,5,6} = 2.4 \mu m / 6 \mu m$, a biasing current of 10 $\mu$A, a common-mode input voltage of 5 V and a quiescent weight potential of 2 V.

Figure 2.17: Static characteristics of the Gilbert multiplier in subthreshold MOS. $i_U$ against $v_x$ for $(W/L)_{M1,2,3,4,5,6} = 24 \mu m / 2.4 \mu m$, a biasing current of 200 nA, a common-mode input voltage of 3 V and a quiescent weight potential of 1 V.
Figure 2.18: Wide range Gilbert multiplier.
Figure 2.19: Switched-capacitor synapse. (a) excitatory synaptic connection, (b) inhibitory synapse and (c) transient characteristics.
Figure 2.20: Current summer.
Figure 2.21: Squashing function based on a transimpedance amplifier.
Figure 2.22: Static characteristic of the transimpedance amplifier. 
\(v_y\) versus \(i_s\) for \((W/L)_{M1,2,3}=2.4\mu m/6\mu m\) and \(V_{CI} - V_{CO} = 0.2\ V\).
Figure 2.23: I/V converter implemented utilising a CMOS arrangement of biased diodes.
Figure 2.24: Static characteristics of the diode-based I/V converter.

$v_Y$ against $i_S$ for $(W/L)_{m1,2,3,4} = 12 \mu m / 2.4 \mu m$ and $E_1 = E_2 = 0.89 \text{ V}$. 
Figure 2.25: Thresholding generator for pulse-stream.
Figure 2.26: Digital memory based on weighted current sources.
Figure 2.27: Sample/hold memory cell.

Figure 2.28: Floating-gate MOS transistor.
Chapter 3

Subthreshold design for feedforward neural networks

The aim of this chapter is to identify the problems associated with the design of existing building blocks, namely multipliers, current-to-voltage converters and weight storage schemes, in the context of a larger system such as a feedforward ANN when MOS devices are operated in the subthreshold mode of conduction and to present alternative solutions. A design method which exploits the exponential transfer characteristic of MOS transconductors is presented, based on the fact that the activation function of a neuron, which is incorporated into a whole analogue system, can be distributed over the next layer of synapses without affecting the overall behaviour of the network [64]. Two different cells have been specifically designed, to overcome some of the disadvantages associated with subthreshold biasing and to emulate the functions required by the above mentioned design method, namely a four-quadrant multiplier [64] and a load [99]. The structure of the former circuit is based on a cross-coupled quad design in which differential multiplication is obtained by driving the bulk (also known as the back gate)
terminals of MOS transconductors. The latter scheme displays a transfer characteristic similar to that of the CMOS configuration of biased diodes presented in the previous chapter, with the difference that the sensitivity to power supply variations has been substantially improved utilising an arrangement incorporating feedback. The performances of these building blocks are assessed based analytically and via simulation. The discussions will include their relative advantages and disadvantages. The final objective of this chapter is to examine the potentials, in terms of speed and ease of programmability, of a neural network based on the proposed technique.

3.1 Statement of problems

It has been demonstrated in the previous chapter how four-quadrant multipliers, exploiting the transconductance characteristic of MOS transistors, such as the Gilbert cell and its derivatives, offer a viable solution to the emulation of synaptic connections. It has also been shown that when the devices are operated in the subthreshold mode of conduction power dissipation is diminished not only because current levels are typically lower than transistors biased into strong inversion, but also because the biasing voltages can be reduced. However in the context of a whole system, combination of this primitive building block with that of a weight storage cell based on a sample-and-hold design and a non-linear load poses some challenges. This section concentrates on the problems associated to the non-linear dependence on both of the inputs of the Gilbert cell.

For a typical $\kappa$ value of 0.84, the linear range of the inputs of the Gilbert multiplier, implemented utilising NMOS devices, are limited to $\pm V_t / \kappa = \pm 30 \text{ mV}$. Such a
confined linear input swing imposes limits on both the weight resolution and the dynamic range of the thresholding generator as shown below.

The weight resolution is determined by both the accuracy of the weight storage scheme $V_s$ and the linearity range of the synaptic weight input $V_r$ and is given by

$$W_r \leq \log_2 \left( \frac{V_r}{V_s} \right) \quad \text{(bits)}$$

(3.1)

Even if utilising compensation techniques such as the dummy device [90], to minimise error due to the clock feedthrough phenomenon, one may expect the relative accuracy of a sample-and-hold storage scheme to be 2 mV at best. Thus, it can be seen from (3.1) that the resolution of the weights is limited to 4 bits which is well below the precision necessary for the successful training procedure of a feedforward ANN [93-94].

As mentioned in the previous chapter, in order to achieve compatibility, the activation function generator must supply a signal which is squashed within the input range of the synapses it is feeding. With signal levels as low as those mentioned above the system would therefore be highly susceptible to noise.

A possible solution to these problems would be to linearise the inputs of the multiplier utilising two pre-processing cells which would be characterised by a tanh$^{-1}$ type distortion. This may be achieved using diode connected transistors as current-to-voltage converters [70]. However this would have the inconvenience of considerably increasing the silicon size, as determined by the number of transistors, of the synaptic cell.

It is shown in the next section how the undesired saturation associated with the input of a modified Gilbert multiplier (i.e. tanh function) could be exploited in the design of a feedforward neural network.
3.2 Design method

Let us consider the local section of a fully connected feedforward neural network as depicted in Fig. 3.1. It has been shown in the introduction that the neuron transfer function can be expressed in the form

\[ x_{i}^{m+1} = f(S_{i}^{m}) = f\left( \sum_{j} W_{ij}^{m} x_{j}^{m} + b \right) \]  

(3.2)

where \( f(.) \) is the sigmoid activation function, \( x_{i}^{m+1} \) is the output of the \( i^{th} \) neuron in the layer \( m \) and \( x_{j}^{m} \) represents the input of the \( j^{th} \) connection link in the layer \( m \).

The solution that is presented here makes use of the non-linearity associated with the input of a transconductance multiplier to perform both the activation function and synaptic multiplication. The proposed design technique is based on the observation that the activation function of the \( i^{th} \) neuron in layer \( m \), depicted in Fig. 3.1, can be distributed over each of the forward layer synapses without modifying the overall behaviour of the neural network [64] as shown in Fig 3.2. It may be noted that the transfer function of that neuron is unchanged and is given by (3.2). It can consequently be seen that for the hidden and output layers, the thresholding and synaptic weighting functions could be combined to form what will be referred to as a Thresholding-Synapse (TS). The output and input of a TS block are related as

\[ P_{ki}^{m+1} = W_{ki}^{m+1} f(S_{i}^{m}) \]  

(3.3)

where \( P_{ki}^{m+1} \) and \( S_{i}^{m} \) are the output and input of the TS block respectively; \( i \) and \( k \) define the driving neuron in the layer \( m \) and the driven node in the layer \( m+1 \) respectively. The following subsection indicates that it is possible to implement (3.3) using a simple
non-linear transconductance multiplier consequently eliminating the need for a non-linear load.

3.2.1 Possible circuit implementation

An analogue circuit implementation of the TS cell as required by the neural architecture shown in Fig. 3.2 can be realised utilising the transconductance multiplier depicted in Fig. 3.3. The core of the scheme is a modified Gilbert cell. Note that the top two current sources denoted $I_{\text{ref}}$ have the effect of removing the quiescent components of the output current of the multiplier. Assuming that all transistors have matched characteristics, the difference output current of the suggested circuit (2.34) can be expressed in the form

$$I_{pk} = (I_w - I_{\text{ref}}) \cdot \tanh \left( \frac{V_{\text{si}}}{2V_t} \cdot \kappa \right)$$

When this expression is related to (3.3) it can be seen that the value of the synaptic weight of the TS is given by $W_i = I_w - I_{\text{ref}}$ which is adjustable via the biasing current source $I_w$ and the activation function is a tanh function. Although such a circuit offers an ideal solution to the simulation of a TS block so far as:

- It provides for the most commonly used activation function in the simulation of feedforward ANNs which is the sigmoid function;

- The weight is linearly controllable over a bipolar wide range; and

- It is compact since it integrates a minimum number of elements (4 active transistors and 4 current sources of which one of the latter is required to be alterable).

some problems arise when one needs to consider a suitable analogue memory.
3.2.2 Current memory circuit

Utilising the above transconductance multiplier to simulate the TS cell the weight signal is represented as a current. The memory must behave as a linearly controllable current source. As mentioned in the previous chapter, such a storage circuit can be constructed by digitally controlling the sum of weighted current sources which are generated via transistors operating as transconductors. However, this digital-to-analogue conversion technique is not suitable for weight resolution value greater than 3 bits for reasons of poor efficiency in silicon area. An alternative solution would be to make use of a basic voltage-to-current transfer cell, whereby a voltage stored on a capacitor is transduced to a current via a transistor [97]. The circuit diagram of such a scheme is shown in Fig. 3.4 (a). Here the two switch-transistors are driven by a two-phase clock signal $\phi_1$ whose transient states are illustrated in Fig. 3.4 (b). When both switches are closed, the input current $I_{mem}$ charges the holding capacitor $C_s$ to a specified voltage $V_s$ such that at the end of this sampling period, $T_s$, the current flowing through the storage transconductor $M_s$ is the memorised current. During the holding phase, $T_h$, the switches are turned off and $M_s$ operates as a current source which is loaded by a diode connected transistor $M_i$. Note that during this period the voltage stored on the capacitor is isolated and is therefore memorised. Note also that $V_s$ is used to generate the weighted current required by the synaptic circuit via an additional transconductor $M_s'$ whose characteristics match those of $M_s$. However several problems are associated with this type of current memory.

First, during the switching off transient, part of the charge stored in the inversion layer of the switch-transistor is injected $q_{ij}$ on $C_s$ reducing the gate voltage of the storage transconductor by
Furthermore, during the sampling mode, the drain-to-source voltage of M₃, which operates as a diode biased in the subthreshold region, is approximately $V_T$. In the holding mode this voltage changes to $V_{dd}$ minus the voltage drop across the load which is also approximately $V_T$. Hence, when the transfer circuit is switched from tracking to holding this variation is fed back to $C_s$ via a coupling capacitance $C_c$. $C_c$ is represented by the dotted capacitor in Fig. 3.4 (a). This effect increases $V_e$ by roughly the following amount

$$\Delta V_{st} \approx (V_{dd} - 2V_T) \frac{C_c}{C_s + C_c}$$

(3.6)

It may be noted that in the weak inversion region, the value of $C_c$ is approximately that of the drain-to-gate overlap capacitance [74].

Finally, during the holding phase the charges accumulated on $C_s$ are discharged at a rate determined by both the reverse-biased current of the source/drain-to-substrate junction, represented by a dotted diode in Fig. 3.4 (a), and the subthreshold conduction of the feedback transistor. This leakage decreases the stored voltage and therefore the memorised current.

To reduce $\Delta V_{st}$ (3.5) Guggenbühl et al. [97] have suggested the combination of two solutions which are:

- Reducing the amount of injected charge utilising a dummy-switch compensation technique; and
- Increasing the value of the holding capacitor by means of a Miller enhancement method.
Whereas $\Delta V_\omega$ (3.6) is attenuated by replacing the storage transconductor by a regulated cascode device whose feedback capacitance is lowered by about 100 times. However these compensating techniques are only applicable when the current memory cell is operated in strong inversion.

To bring the current error caused by clock feedthrough and parasitic capacitive coupling within acceptable limits, (i.e. $< 0.1\%$) Pain and Fossum [98] have proposed an alternative scheme which makes use of a feedback arrangement. The cell, shown in Fig. 3.5 (a) in its NMOS version, has been developed for a low-power application and is therefore suitable for weak inversion current levels. Note that the circuit is controlled by a pair of two-phase clock signals $\phi_1$ and $\phi_2$ and their respective inverses $\bar{\phi}_1$ and $\bar{\phi}_2$. The transient states of $\phi_1$ and $\phi_2$ are depicted in Fig. 3.5 (b). The sampling operation takes place when the switches associated to $\phi_1$ and $\phi_2$ are closed with the dynamics of this phase defined by the non-linear differential equation

$$I_{\text{mem}} = I_{C_s} + I_{M_5} = C_s \frac{\partial V_s}{\partial t} + I_x \exp\left(\frac{k \cdot V_x}{V_i}\right)$$  \hspace{1cm} (3.7)

where $V_s$ is the gate-to-source potential of the storage transconductor. The solution to this differential equation is

$$I_{M_5}(t) = \frac{I_{\text{mem}}}{1 - \varphi \exp\left(\frac{k \cdot I_{\text{mem}}}{C_s V_i} t\right)}$$  \hspace{1cm} (3.8)

where the constant $\varphi$ is dependent on the magnitude of $I_{M_5}$ at the origin.

The current-error $i_e$, resulting from the change of state in $\phi_1$, is then partially fed back to the holding capacitor, via a pair of current mirrors $M_1$-$M_2$ and $M_3$-$M_4$, creating a negative change in the stored voltage. This compensation phenomenon takes place until
the error decreases to a level which cannot be detected by the feedback loop. As a consequence, the current flowing through $M_s$ is approximately $I_{\text{mem}}$ to the precision of a residual current. This correction process is also described by a non-linear differential equation

$$i_e = -\frac{C_s}{\eta} \frac{\partial V_s}{\partial t} = I_{M_s} - I_{\text{mem}} = I_X \exp\left(\frac{\kappa}{V_t} V_s\right) - I_{\text{mem}}$$

(3.9)

where $\eta$ is the ratio of the current mirrors. The solution to this differential equation is given as

$$i_e(t) = \frac{I_{\text{mem}}}{1 - \frac{I_e}{I_{\text{mem}}} \exp\left(\frac{\eta \kappa I_{\text{mem}}}{C_s V_t} t\right)} - I_{\text{mem}}$$

(3.10)

where $I_e = i_e(t = 0)$. Note that for a stable operation of the feedback structure the ratio of the current mirrors should be smaller than 1. Furthermore, the correction scheme relies on a change in $V_s$, induced during the transition from sampling to error compensation mode, to be positive. This requires the coupling capacitance $C_c$ to be greater than the gate capacitance of $\phi_i$ switch.

(3.8) and (3.10) indicate that the time required to charge the gate capacitor and subsequently compensate for errors due to switching transient such that the difference between the current flowing in $M_s$ and $I_{\text{mem}}$ is within one Least Significant Bit (LSB) precision, is approximately

$$T_{sc} = T_s + T_e = 5. \frac{(1 + \eta) C_s V_t}{\eta \kappa I_{\text{mem}}}$$

(3.11)

which is inversely proportional to the memorising current. Thus for currents as low as 10 nA, and a typical storage capacitance of 1 pF with a mirroring factor of 0.2, this period can extend to up to 90 $\mu$s. Note that during the holding mode, $\phi_1 = \phi_2 = 0$, 100
transistor M, acts as the load and its mirrored current is short-circuited to the ground via \( \bar{\phi}_2 \) switch.

To compensate for the loss of information due to charge leakage, the stored current value may be periodically regenerated [71] via a Digital-to-Analogue Converter (DAC) circuit. A sequential refreshing technique is desirable to minimise the number of DACs. Therefore the amount of synaptic weight that can be refreshed by a single DAC is determined by both the holding and the sampling/compensation periods

\[
N_s = \frac{T_h}{T_{sc}}
\]

where \( T_h \) is the interval between update phases such that the precision of the weight is maintained at 1 LSB. A numerical approach suggests that for a maximum current of 510 nA with 1 LSB corresponding to 2 nA and assuming that the leakage current is approximately 1 pA, the duration of the holding phase would be about 120 \( \mu \text{s} \) since the accuracy of the gate voltage between two quantized levels is of the order of 120 \( \mu \text{V} \) for weight values situated in the upper range. These results would imply that the maximum number of synaptic weights that could be refreshed by a single DAC would only be one, thus indicating the impracticability of a current memory scheme, and therefore that of the modified Gilbert cell as a means to simulate a TS cell. Note that if the devices were operated in the strong inversion region the ratio of synapses to DACs would be increased by 3 orders of magnitude.
3.3 The conceptual building blocks of the Neural Network

This section shows that it is possible to obtain a transconductance relationship similar to that of the modified Gilbert multiplier in an arrangement in which the bias current is distributed across two differential pairs by modulating the bulk potentials of the devices, thus eliminating the need for a current memory design. It will also be shown that this circuit combined with that of a compensated sample/hold using a dummy switch scheme offers the possibility of simulating 8 bits weight resolution. Finally, it will be demonstrated that the sensitivity to biasing variations of a load arrangement based on diode-connected transistor has been substantially improved.

3.3.1 A four-quadrant multiplier

The four-quadrant multiplier [64], as shown in Fig. 3.6, has been developed to efficiently implement the TS block and consists of a current source $I_p$ and two differential pairs $P_1$ and $P_2$ of matched transistors operating in the subthreshold mode of conduction. Differential multiplication is obtained by applying one differential input $v_x = V^*_x - V^x$ between the gate terminals of $M_1$ ($M_a$) and $M_2$ ($M_b$) whilst the second signal $v_w = V^*_w - V^w$ appears between the bulk terminals of $M_1$ ($M_2$) and $M_3$ ($M_4$). It may be added that the differential pairs must be integrated on separate wells.

With the assumptions that the drain-source potential $V_{ds} \geq 4V_t$ but remains low enough to disregard Early effect, the drain current of a PMOS transistor operating in the weak inversion is given in Appendix A by

$$I_D = -I_X \exp\left[-\kappa \frac{V_{gs}}{V_t}\right] \exp\left[-(1-\kappa) \frac{V_{bs}}{V_t}\right]$$

(3.13)
where all the parameters are defined in Appendix A. Typical parameters for a minimum-size device (2.4 μm x 2.4 μm) fabricated in a standard analogue 2.4 μm p-substrate are for a p-type transistor, \( I_x = 26.6 \times 10^{-18} \text{ A} \) and \( \kappa = 0.7 \). It should also be emphasised that the normal operation of a PMOS transistor requires that the source/bulk and drain/bulk junction be reverse-biased, i.e. \( V_{bs} \geq 0 \) and \( V_{bd} \geq 0 \).

The transfer characteristics can be determined as follow. Utilising (3.13), it can be shown that the sum of the devices currents at the common source node is expressed as

\[
I_B = -(I_{M1} + I_{M2} + I_{M3} + I_{M4}) = I_x \exp \left( \frac{V_s}{V_i} \right) \left[ \exp \left( -\kappa \frac{V^*_s}{V_i} \right) + \exp \left( -\kappa \frac{V^*_x}{V_i} \right) \right] \\
\left[ \exp \left( -(1-\kappa) \frac{V^*_w}{V_i} \right) + \exp \left( -(1-\kappa) \frac{V^-_w}{V_i} \right) \right]
\]

(3.14)

while the difference output current is given by

\[
i_o = (I_{M2} + I_{M4}) - (I_{M1} + I_{M3}) = I_x \exp \left( \frac{V_s}{V_i} \right) \left[ \exp \left( -\kappa \frac{V^*_x}{V_i} \right) - \exp \left( -\kappa \frac{V^-_x}{V_i} \right) \right] \\
\left[ \exp \left( -(1-\kappa) \frac{V^*_w}{V_i} \right) - \exp \left( -(1-\kappa) \frac{V^-_w}{V_i} \right) \right]
\]

(3.15)

Thus combining (3.14) and (3.15) it can readily be shown that

\[
i_o = I_B \tanh \left[ \kappa \frac{V^*_x}{2V_i} \right] \tanh \left[ (1-\kappa) \frac{V^-_w}{2V_i} \right]
\]

(3.16)

This transconductance relationship is similar to the one obtained by the MOS version of the Gilbert cell (2.37) with the exception that the coefficient of the factor, \( v_w / 2V_i \), is 1 - \( \kappa \) instead of \( \kappa \). This variation has the desirable effect of extending the linearity range of differential input \( v_w \) from \( V_i / \kappa \approx 37 \text{ mV} \) to \( V_i / (1-\kappa) \approx 90 \text{ mV} \). It will be shown
later how this extended linearity has been exploited. Note that the gain of the multiplier is adjustable through $I_B$.

Since, the core of the multiplier uses only four devices, mismatching and its effect on input offset may be more readily manageable. This will be covered in the next chapter. It may also be noted that having only one stack of transistors will permit a decrease in supply voltage, thus reducing the power dissipation.

### 3.3.1.1 Limits of operation

The operating range of the $v_w$ differential input is, however, limited by the fact that the source/bulk junctions of transistors $M_1$-$M_4$ should not be forward-biased. This condition leads to the following requirement

$$V_w + v_w \geq V_s + v_s \quad (3.17)$$

where $V_w$ is the quiescent bulk common-mode input level, $V_s$ and $v_s$ are the quiescent and dynamic source potentials, respectively. The quiescent and dynamic source voltage can be obtained by rearranging (3.14) as follows

$$V_s = V_t \left[ \ln \left( \frac{I_B}{4.1X} \right) + \kappa \frac{V_X}{V_t} + (1 - \kappa) \frac{V_w}{V_t} \right] \quad (3.18)$$

and

$$v_s = V_t \left[ \ln \left( \text{sech} \left[ \kappa \frac{V_X}{V_t} \right] \right) + \ln \left( \text{sech} \left[ (1 - \kappa) \frac{V_w}{V_t} \right] \right) \right] \quad (3.19)$$

where $V_X$ is the quiescent gate potential. Note that whatever the sign or amplitude of the dynamic potential applied to either the gate or bulk terminals, the dynamic signal at the common-source node will always be negative. Hence combining (3.17), (3.18) and
disregarding the influence of the dynamic signal \( v_x \) the maximum input swing of the bulk input can be expressed as

\[
|v_{W\text{max}}| = V_i \left[ \ln \left( \frac{I_b}{4.1X} \right) + \kappa \frac{(V_X - V_W)}{V_i} \right]
\]  
(3.20)

Therefore, to fully exploit the extended linearity range of the bulk dynamic input, \( I_b, V_X, V_W \) and the aspect ratio of the transistors \( W/L \) needs to be carefully chosen so that the signal swing remains within the limits defined by the above expression.

### 3.3.1.2 Simulation results

The transconductance multiplier performance has been optimised via SPICE simulations based on a 3 V power supply using Level 3 model parameters for the MIETEC 2.4 |im CMOS process. With bias currents set at 250 nA, extensive simulations have shown that a maximum bulk dynamic and linearity range of ± 250 mV can be obtained for an aspect ratio of \( W/L = 10/1 \) for the transistors where the common mode voltage of the gate and bulk terminals are set to 1.5 V and 2.55 V respectively. The family of simulated static characteristics shown in Fig. 3.7 (a) and (b) suggest that over the linear range of \( v_w \) (3.16) may be approximated as

\[
i_o = g . v_w . \tanh \left( \kappa \frac{v_X}{2. V_t} \right)
\]  
(3.21)

where \( g = I_o (1 - \kappa) / 2. V_t \) which is similar to the transfer function of the modified Gilbert cell. Thus this transconductance multiplier may be employed to implement the function described by the TS building block.

It has however been observed in Fig. 3.7 (b) that the linearity range of the bulk difference input is greater than that which was predicted by the analytical results. This variation is mainly due to the fact that the coefficient which measures the effectiveness
of the gate in controlling the channel current is a function of the bulk-to-source potential of the device [72], [61], and that its value tends to 1 as the level of $V_{bs}$ gets larger. This effect is however desirable. Comparing (3.21) and (3.3) it may be noted that the signal $v_w$ could be used to control the value of the weight. The fact that its linearity range is of the order of $\pm 250 \text{ mV}$, (3.1) indicates that the system would allow for an 8 bits weight resolution as long as the precision of the weight storage scheme is within 2 mV. Note that the effectiveness of the gate at controlling the barrier energy is inversely proportional to the square-root of the device substrate doping [74]. Thus a weak doping would tend to increase $\kappa$ towards unity and further extend the linearity range of the $v_w$ input of the multiplier. This influence may be observed by simulating an NMOS version of the circuit.

Further simulations based on a range of values for the key static variables $I_b$, $I_x$, $V_x$ and $V_w$ have been carried out and suggest that the dynamic range estimate given in (3.20) is reasonable.

### 3.3.2 The load

The load converts the summed (neural) current to a potential as required to drive the inputs of the following layer. Given that the thresholding function is incorporated within the synapse, to fulfil the conditions of implementation suggested in (3.3), it might be assumed that a linear current-to-voltage converter would be needed. However, it will be shown in the next section that moderate levels of non-linearity in the load resistor characteristics would have little effect on the overall performance of the MLP, and consequently a design based on a minimum number transistor structure can be used.
It has been shown in the previous chapter that the current-to-voltage conversion can be achieved using a CMOS arrangement of biased diodes. It has also been noted that because the drain current in a subthreshold mode device is an exponential function of the gate-source voltage, the I/V characteristic of such a load resistor would be highly sensitive to biasing arrangements.

Note that the sensitivity could be substantially enhanced by keeping the difference potential $V_{ds} - (E_1 + E_2)$ to a constant level, see (2.44) for details. This may be achieved by rearranging the position of the bias sources $E_1$ and $E_2$, as indicated in the circuit diagram of Fig. 3.8. It is important to note that unlike in the original scheme, the core diode group $M_1$-$M_4$ and the associated biasing elements can be implemented using either PMOS or NMOS devices, thus allowing for a more compact integration. The biasing potentials are generated by two diode connected transistors $M_5$ and $M_6$, and are controlled by a feedback arrangement comprising transistors $M_7$-$M_9$ together with a current source $I_2$. Since the potential across the diodes $M_8$ and $M_9$ is nominally fixed by $I_2$, variations in the supply ($V_{dd}$) are distributed across the gate-source junctions of the shunt transistor $M_7$ and the biasing transistor $M_5$. With the proviso that these transistors have matching characteristics and that their quiescent currents are similar, it follows that approximately half of the supply voltage variation will appear across the device $M_5$. Given the symmetry of the scheme it also follows that an equal variation will appear across the diode connected transistor $M_5$. With the biasing potential across the core diode group (i.e. $V_{ds} - E_1 - E_2$) buffered from supply variations the load resistance sensitivity is correspondingly reduced.

This phenomenon can be accounted for by assuming that the drain current of a PMOS transistor is related to its gate-source potential as
It is shown in Appendix D that, based on this transistor model, the transfer characteristic of the scheme is given

\[ v_o = \frac{2V_i}{\kappa} \cdot \text{sinh}^{-1} \left( \frac{i_o}{4I_z \cdot \sqrt{1 + \frac{I_c}{I_{M7}}} } \right) \]  

(3.23)

where \( I_c \) is the current biasing the core diode group and \( I_{M7} \) is the drain current of the shunt device \( M_7 \). It can be seen that since \( I_c \) is nominally independent it follows that the sensitivity of the resistance value is determined by the static ratio \( I_c / I_{M7} \). Any given level of insensitivity is achievable by ensuring that the shunt current is sufficiently larger than the core group current. Although the sensitivity is improved at the expense of an increase in the static current consumption, when considering a battery-powered system it is however a justifiable expense since supply voltage cannot in practice be regarded as a constant.

3.3.2.1 Simulation results

The family of simulated static characteristics in Fig. 3.9, is based on a bias current \( I_z \) of 3.125 nA, an aspect ratio of \( W/L = 5/1 \) for all transistors and a 3 V power supply. They show the level of sensitivity to a power supply variation of ± 100 mV. For such biasing conditions the sensitivity is enhanced by 40 times while the static current consumption is increased by a factor of 30. However, it can be appreciated that under these conditions of operation the transfer characteristic of the load can be approximated as

\[ v_o = \frac{2V_i}{\kappa} \cdot \text{sinh}^{-1} \left( \frac{i_o}{4I_z} \right) \]  

(3.24)
Note that even in this mode of operation the power dissipated by the load is of the same order as the multiplier.

Extended simulations have also indicated that increasing the aspect ratio $W/L$ of all devices has the effect of decreasing the static current ratio $I_c / I_m$ and consequently improve the sensitivity of the nominal resistance to power supply variation. It may be pointed out, as a matter of interest, that to render the load completely insensitive to supply variation it would be necessary to substitute both biasing diodes $M_5$ and $M_6$ by two ideal current-controlled voltage sources. This could be achieved by buffering the biasing potential $E_1$ and $E_2$ utilising two op amps operating as followers. However, in the context of neural network design, this solution would not be area efficient.

3.4 Circuit Implementation

Following the results presented in the above section it can be seen that an analogue implementation of the neural network architecture depicted in Fig. 3.2 can be realised, as shown in Fig. 3.10, using a Transconductance-Thresholding-Synapse (TTS) [64] to emulate the function of the TS block. The TTS building cell is composed of a four-quadrant multiplier and two grounded current sources whose values are half of the biasing current source. The role of the latter devices is to eliminate the quiescent component of the output signal of the multiplier so that the difference output current of the TTS contains only a dynamic term. This choice has been made in order to avoid high current levels at the summing junction which would require larger wiring busses and create undesirable parasitic coupling capacitance. It can readily be demonstrated (3.21) that in the linear range of $V_{wl} (|V_{wl} \leq 2V_i / (1 - \kappa))$ the difference output current of a TTS can be expressed in the form
\[ I_{pk} = g \cdot V_{wi} \cdot \tanh \left( \kappa \cdot \frac{V_{si}}{2 \cdot V_i} \right) \]  

Note the graphic illustration of a TTS cell is circle embodying a thresholding and a multiplying sign. The summation function can simply be achieved by connecting the outputs of the TTS to a common bus bar. Finally the current-to-voltage conversion can be realised using the load arrangement described earlier whose transfer characteristics (3.24) can be approximated as

\[ V_{si} = \frac{2 \cdot V_i}{\kappa} \cdot \sinh^{-1} \left( \frac{I_{si}}{4 \cdot I_z} \right) \]  

For reasons of convenience, the symbols of both the summation and load circuits have been brought together, as shown in Fig. 3.10.

### 3.4.1 Combined load and transconductance-thresholding-synapse

The transfer characteristic of an analogue TS block can be obtained by combining (3.25) and (3.26) and its normalised input and output variables are related as

\[ P_k = \frac{g}{4 \cdot I_z} \cdot V_{wi} \cdot \tanh \left[ \sinh^{-1}(S_i) \right] \]  

where \( P_k = I_{pk} / 4 \cdot I_z \) and \( S_i = I_{si} / 4 \cdot I_z \).

When this expression is compared to (3.3) it can be seen that the value of the synaptic weight is given by

\[ W_i = V_{wi} \cdot \frac{I_a \cdot (1 - \kappa)}{8 \cdot I_z \cdot V_i} \]  

which is adjustable in a bipolar fashion via the difference potential \( V_{wi} \), and the activation function is \( \tanh[\sinh^{-1}(.)] \). The fact that the transfer function is not an ideal sigmoid is due to the non-linearity associated to the current-to-voltage converter.
However, the $\tanh[\sinh^{-1}(.)]$ function is similar to the most commonly used form of activation function which is the hyperbolic tangent, as shown in Fig. 3.11. As a consequence (3.27) may be approximated as

$$P_k = \frac{B}{4I_z}V_{wi}\tanh[S_i]$$  \hspace{1cm} (3.29)

Following this analysis it can be appreciated that the non-linearity associated with the load circuit has a minor influence on the transfer characteristic of a feedforward neural system. It will be shown in the next subsection that the presence of the distortion actually enhances the processing speed of the ANN.

Note that to provide a better immunity against noise all signals associated with the TTS and load are differential. Although this improvement is at the expense of lowering transistor integration density, this trade-off is however tenable when one considers the low current and voltage levels associated with weak inversion operation.

Comparing (3.29) and (3.3) it can be deduced that an analogue implementation of a fully connected feedforward neural network can be achieved utilising TTS and load circuits as is illustrated in Fig. 3.12. It can be noticed that since the thresholding function is combined with the input of the synapse:

* An activation function (i.e. $\tanh(.)$) is associated with each node of the input layer of the network; and

* The output of each neuron situated in the output layer does not contain a thresholding element.

A solution to both of these problems will be offered in the next chapter.
3.4.2 Speed of the network

The processing speed of the analogue system discussed above is mainly determined by the amount of current that the TTS circuit can source/sink at its output and by the resistive and capacitive loads it is driving. It can be expected that operating speed will be lower than for a system whose devices are biased in the strong inversion since subthreshold current levels are much lower. However, this is partially compensated for by the fact that, in weak inversion in contrast to strong inversion, the layer of charge throughout the length of the channel is negligible, thus the gate-to-source capacitances are relatively small [74]. The transient characteristic of the analogue structure shown in Fig. 3.12 is described by the simplified small-signal equivalent circuit shown in Fig. 3.13. The current I is controlled by the output current of the TTS, R represents the resistive load, and \( C_L \) the total capacitance at the TTS output and is given by

\[
C_L = N \cdot C_i + C_r \tag{3.30}
\]

where \( C_i \) is the average input capacitance of a TTS, \( N \) is the number of TTS inputs connected at the output of the driving TTS-load, and \( C_r \) is the input capacitance of the current-to-voltage converter which is an order of magnitude smaller than \( C_i \). Thus for large networks (i.e. \( N > 10 \)), \( C_L = N \cdot C_i \) may be used as an approximate relation.

If \( I_{\text{max}} \) is the maximum current available at the output of a TTS, and \( C_L \) is assumed to be linear, then the rate of change of the output voltage is defined by the non-linear differential equation

\[
I_{\text{max}} = C_L \frac{\partial V}{\partial t} + 4.12 \cdot \sinh \left( \frac{V}{2 \cdot V_t} \right) \tag{3.31}
\]

One of the solutions to this differential equation is given by
\[ v(t) = \frac{4V_t}{k}\tanh^{-1}\left\{ \frac{I_{\text{max}}}{I_Z} \tanh\left[ \frac{\kappa\sqrt{I_{\text{max}}^2 + I_Z^2}}{4V_tC_L}t \right] \right\} \]  

(3.32)

Evaluating this expression with the software package: Mathcad, shows that for a maximum output current \( I_{\text{max}} = I_B/2 = 125 \text{ nA} \) (single ended) and a load biasing current of 3.125 nA, the settling time of (3.32), to within 1% of error, is about 1.5 \( \mu \text{s} \) for a capacitive load of 0.25 pF (i.e. \( N \approx 5 \)). As depicted in Fig. 3.14, a SPICE transient simulation also indicates a similar processing speed. Note that this rate of change corresponds to that of a single layer and is mainly dependent upon the number of TTS connected at the output of the driving TTS-load arrangement. An approximation of the processing speed of an MLP would be determined by the sum of the transient time for each layer which is consequently proportional to the size of the network. The above results may therefore be considered as exemplars. However the fact that the resistive element has a non-linear characteristic substantially improves the time response of the neural network as is demonstrated below.

To obtain an ANN system that would accurately possess the characteristic given in (3.29), the transfer function of the resistor would need to be

\[ V_{Si} = \frac{V_t}{2kI_Z}I_{Si} \]  

(3.33)

which is the effective driving-point resistance of the non-linear load at the origin (3.26). Alternatively the computation rate would be determined by the solution of the following differential equation

\[ I_{\text{max}} = C_L \frac{\partial v}{\partial t} + \nu \frac{2I_Z \kappa}{V_t} \]  

(3.34)
which yields, as shown in Fig. 3.14, a relatively longer settling time of 8 μs. This result may be explained, in physical terms by the fact that for the structure shown in Fig 3.12 the effective value of the resistive element decreases as the current flowing through it increases, thus one would expect the charging/discharging time to be reduced.

3.4.3 Weight range

It can be noted in (3.28) that, since the signal range of the weight control input is limited to ± 250 mV, the maximum and minimum weight values (i.e. weight range) are determined by the ratio $I_g / I_z$. Thus for a synapse tail current of 250 nA and a load biasing current of 3.125 nA the weight range is set to ± 10. Note that this range can easily be altered to suit any given neural network problems since $I_g$ and $I_z$ can be externally programmed.

It is also interesting to note that the power dissipation can be traded off against processing speed (3.32) or vice versa without effectively changing the weight range.

3.5 Summary

The limited input signal range for a subthreshold mode transconductance multiplier has been extended by distributing some thresholding operations for feedforward neural networks over to the inputs of the synapses.

The conceptual circuit of a TTS cell based on a new four-quadrant MOS analogue multiplier has been presented. The circuit is area efficient and less likely to be vulnerable to mismatching effects since a minimum number of devices are utilised. It has a low power dissipation because a minimum stack of devices permits a reduction in the supply voltage. It offers a wide range linear differential input and the non-linear tanh
function associated to the other differential input has been used to achieve the activation function required by the TS. The conditions of operation of the multiplier have also been analytically derived and verified via a series of SPICE simulations.

An alternative means to resolve the problems associated with load arrangements based on diode connected transistors has also been presented. Sensitivity to power supply is however improved at the expense of power dissipation which remains low enough not to be considered a burden. The cell is also area efficient since a single type of MOS transistor is utilised.

Analytical and simulation studies have also shown that the non-linearity associated with the load does not interfere with the characteristics of the network and considerably enhances the processing speed.

It has also been established that the weight range of the synapses can readily be adjusted and power dissipation traded off against processing speed irrespectively.
Figure 3.1: Local section of a fully connected feedforward neural network.
Figure 3.2: Feedforward neural network with the activation function distributed over the following synapses.
Figure 3.3: Transconductance multiplier based on the modified Gilbert cell
Figure 3.4: Current memory cell. (a) circuit diagram, (b) transient characteristics.
Figure 3.5: NMOS version of Pain and Fossum current memory cell. (a) circuit diagram, (b) transient characteristics.
Figure 3.6: Circuit diagram of the four-quadrant multiplier.

\[ V_c = V_{c1} + (V_{c2}) \]

\[ 2 + 4 \]

\[ M_1 + M_3 \]

\[ V_{dd} \]

\[ I_B \]

\[ V_X^+ \]

\[ V_X^- \]

\[ I_{M1} \]

\[ I_{M2} \]

\[ I_{M3} \]

\[ I_{M4} \]

\[ i_0 = (I_{M2} + I_{M4}) - (I_{M1} + I_{M3}) \]
Figure 3.7: Static characteristics of the four-quadrant analogue multiplier.
(a) $i_o$ against $v_x$ and (b) $i_o$ against $v_w$ for $(W/L)_{M1,2,3,4} = 24\mu m / 2.4\mu m$, a common-mode gate potential of 1.5 V and a quiescent bulk signal of 2.55 V.
Figure 3.8: Current-to-voltage converter comprising a feedback mechanism.

Figure 3.9: Static characteristics of the current-to-voltage converter for various supply voltage, $v_o$ against $i_o$ for $(W/L)_{M1-9} = 12\mu m / 2.4\mu m$ and $I_z = 3.125$ nA.
Figure 3.10: Analogue implementation of a section of a feedforward neural network
Figure 3.11: Sigmoid activation function.
Figure 3.12: Analogue implementation of a fully connected feedforward neural network.
Figure 3.13: Simplified small-signal equivalent circuit of a TTS and load.

Figure 3.14: Transient characteristics of the combined TTS and load circuits.

- Simulation results
- Analytical results
- Linear load
Chapter 4

Design and implementation of a neural network chip

An analogue technique for implementing a fully connected feedforward neural network and the conceptual circuits of the basic building blocks along with their performances have been discussed in the previous chapter. The objective of this chapter is to present the design, in a detailed fashion, of all the elements that have been developed for the implementation of a whole neural network chip. The complete structures of the two previously discussed schemes which make up the body of the network are presented along with an extra two cells that have been conceived to overcome the problems associated with the effect of combining the thresholding and synaptic functions. The issue of weight storage is also covered in this chapter. This will include an overall presentation of the system and a thorough discussion, based on analytical results, of its performance and that of the cells it is incorporating. The overall behaviour of a basic 2 layer analogue feedforward neural network such as the XOR which incorporates 9 synapses, 3 neurons and a weight storage scheme has been assessed via simulation.
studies. A part of this chapter will also be dedicated to introducing the whole design of neural network chip that comprises a 10:6:3 MLP and a 2:2:1 feedforward network. The final section of this chapter will be devoted to a discussion on the layout of the entire system and the physical features of the neural network chip will also be outlined.

During the development of the chip, the choice of the design techniques, and consequently the circuit structures, has been made keeping in mind five different issues namely, the size of the most frequently used cells as determined by the number and dimensions of the transistors involved, the power dissipation, the speed of operation, the accuracy of computation and the ease of integration.

4.1 The complete circuits of the neural network

It has been indicated in the first chapter that an MLP is composed of multiple arrays of identical modules namely, multipliers, summers and thresholding generators. It results from such a type of framework that any sort of MLP topology can be created by simply modifying the size of these groups of array. However since the combination of the basic neural units has been modified, as explained in the previous chapter, it may be shown how this valuable feature has been retained at the circuit level.

4.1.1 The transconductance-thresholding-synapse

As mentioned earlier, the TTS cell includes three current sources of which one is biasing the core of the four-quadrant multiplier and the other two have the function of eliminating the quiescent component of the difference output current. It is essential that the current values of the latter two sources are matched and are a fraction of the former source by a factor of half in order to avoid a quiescent current build up at the summing
junction which could drive the load arrangement into saturation. It is therefore desirable that these current sources are controlled at the circuit level rather than at the chip level to avoid large mismatching between devices resulting from long distance parameter variations [82]. This would be even more apparent as the devices are biased in the weak inversion. Thus in order to render the circuit modular and to prevent substantial current offsets, all the transistorised current sources are driven by a single bias potential $V_b$, as depicted in Fig. 4.1. The output DC cancellation is performed by the two matched NMOS devices $M_4$ and $M_{19}$, while a third transistor $M_8$, of identical characteristics, is generating the biasing current via a PMOS current mirror consisting of equal size elements $M_5$, $M_6$, and $M_7$. The dimensions of the transistors $(W/L)$ are given in $\mu$m. A nominal gain of 2 is achieved for the current mirror by utilising two multiplier transistors ($M_5$-$M_7$) of identical dimensions rather than a single device whose width would be double that of the reference element ($M_4$). This choice was made in order to minimise error in the current ratio due to geometrical mismatching engendered by the etching effect of the polysilicon gate of the transistors [74]. Furthermore, it has been shown in the previous chapter that the common source node potential changes when a signal is applied to either input of the multiplier. A similar phenomenon also takes place between the output nodes of the TTS as it is directly feeding a load circuit. Since both of these variations appear between the drain and source terminals of the transistors forming the current sources, long channel devices have been used as a means to diminish the effect of channel length modulation. Note the compromise between size and accuracy of computation.

The effect of transistor mismatch on the difference output current of the complete TTS scheme was assessed via a series of Monte Carlo SPICE simulations [19]. Pelgrom
et al. [82] have suggested that the dominant source of mismatch between two identical MOS transistors biased in the weak inversion is the variation in the zero bias threshold voltage $V_{TO}$ and that its variance is given by

$$\sigma^2(V_{TO}) = \frac{A_{VTO}^2}{W.L} + S_{VTO}^2 D^2$$  \hspace{1cm} (4.1)$$

where $A_{VTO}$ is the area proportionality constant of the threshold voltage and $S_{VTO}$ describes the variation of $V_{TO}$ with respect to the distance $D$ between the components. Both of these parameters are process dependent and can be empirically determined. However Pelgrom et al. experimental results suggest that for a 2.5 μm, 50 nm gate-oxide CMOS process, which is a wafer fabrication technology similar to that of MIETEC, the value of these parameters are:

- For an NMOS: $A_{VTO} = 30 \text{ mV} \mu \text{m}$, $S_{VTO} = 4 \mu \text{V/} \mu \text{m}$; and
- For a PMOS: $A_{VTO} = 35 \text{ mV} \mu \text{m}$, $S_{VTO} = 4 \mu \text{V/} \mu \text{m}$.

It may be noted that, for small devices situated at a short distance apart, the distance dependent component of the variance is negligible with respect to the area component.

According to this data, a series of 100 Monte Carlo simulations of the whole TTS circuit, when both of its inputs are short-circuited and with a biasing current of 250 nA, indicate a standard deviation of the difference output current (i.e. $\sigma(I_{PK})$) of 13 nA. Note that $\sigma(I_{PK})$ is approximately 10% of the full scale output current. This offset current can be compensated for, during training, by modifying the weight of the biasing connection of the neuron. This error correction technique is in practice adequate since it is most unlikely that the sum of the offset currents of a batch of TTSs will exceed the maximum output current of the biasing synapse.
As shown in Fig. 4.2, the advantage of utilising a single bias potential \( (V_b) \) is that it allows for the gain of a large group of TTSs to be straightforwardly controlled. \( V_b \) can be generated utilising a diode connected transistor biased by an external current source. However, due to the fact that the TTS cell \( I \) and \( N \) can physically be far from each other their biasing current can substantially be mismatched. This difference can be as high as 15\% for distances of the order of 1 mm and manifests itself as a variation of the gain, which can also be compensated for during training (see (3.28)).

4.1.2 The load

The biasing technique described above can also be applied to the load. However, since the value of the current source \( I_z \) is substantially lower than that of the TTS, the effect of process parameter variations over long distance is potentially more devastating. According to the transfer characteristic of the \( I/V \) converter given in (3.26), it can be appreciated that a change of 1 nA in \( I_z \) leads to a 20\% variation in the effective driving point of the resistance. To prevent this undesired phenomenon from occurring, the structure of the feedback mechanism has been altered to allow a biasing current level similar to that of the TTS to be used while retaining identical resistive characteristics.

As mentioned in the previous chapter, the role of \( I_z \) is to generate a constant potential \( E_o \) across two diode connected transistors. Note that the biasing requirements could be satisfied by substantially increasing the width of these diodes. However this would be achieved to the detriment of silicon size. The alternative solution, as indicated in the circuit diagram of Fig. 4.3, is however area efficient since a single diode \( (M_g) \) is used to induce \( E_o \). Furthermore the structure of the feedback mechanism, which includes the shunt and diode devices together with the biasing source, has been rearranged so that \( I_z \)
can be implemented utilising an NMOS device \((M_n)\). That is because, as suggested by Pelgrom \textit{et al.}, matching is poorer for p-channel transistors than for n-channel ones [82]. For the same reason as stated earlier, a long channel biasing element has also been employed to limit the effect of variations in the supply voltage on the current \(I_Z\). The choice of the other components' aspect ratios were influenced by the necessity for suitable power dissipation, sensitivity to supply voltage variations and silicon area.

Using (4.1) as a means to model the standard deviation of the threshold voltage of each transistor, a series of Monte Carlo SPICE simulations of the whole load have shown that, for an un-driven input and a biasing current set to 100 nA, a differential input voltage offset of about 10 mV can in practice be expected. A similar batch of simulations also indicated that the resistance of the load can vary from its nominal value by as much as 13% if situated as far as 1 mm from the biasing generator.

4.1.3 The pre-processing cell

As mentioned in the previous chapter, the adoption of a TTS approach results in the introduction of a tanh function non-linearity at the input layer connection. This requires that the input data be pre-processed [64] to compensate for the influence of this non-linear function. Where the neural network is incorporated in a programmable loop, the pre-processing can easily be accommodated within the host computer software.

However, if the neural hardware is physically interfaced to the real world analogue sensors, the non-ideality demands a physical solution. Although this can be achieved by designing an extra synaptic cell to perform the required operation the preferred approach makes use of a Horizontal Resistance (HRes) [22], [100] arrangement which displays a \(\tanh^{-1}(.)\) pre-distortion function. The structure of the resistor and its symbol are shown in Fig. 4.4. It consists of a core of four diode-connected transistors \((M_1-M_4)\) and two
current sources implemented by \((M^5-M^6)\) and \((M^{10}-M^{13})\). Note that the values of these biasing currents are identical and are twice that of the TTS ones. This choice was primarily made for the following two reasons:

- To make sure that the synaptic weights associated to either the input or hidden layers have a similar range of operation.

- To ensure input/output compatibility for the network design, and allow for the possibility of interconnected neural network chips and the implementation of recurrent networks.

The additional advantage is that both HRes and TTS cells can share the same biasing generator as long as their biasing transistors have identical aspect ratios.

Following the Kirchhoff's current law, it can readily be shown that

\[
i_i = \frac{(I_1 - I_2) - (I_3 - I_4)}{2}
\]  

(4.2)

For a core of matched transistors operating in the subthreshold mode of conduction (2.28) it can also be demonstrated that the difference output currents of the top and bottom differential pairs are respectively given by

\[
I_1 - I_2 = -2.1B \cdot \frac{\exp\left(-\kappa\frac{V^+_{Si}}{V_i}\right) - \exp\left(-\kappa\frac{V^-_{Si}}{V_i}\right)}{\exp\left(-\kappa\frac{V^+_{Si}}{V_i}\right) + \exp\left(-\kappa\frac{V^-_{Si}}{V_i}\right)} = 2.1B \cdot \tanh\left(\kappa\frac{V_{Si}}{2V_i}\right)
\]  

(4.3)

and

\[
I_3 - I_4 = -2.1B \cdot \frac{\exp\left(\kappa\frac{V^+_{Si}}{V_i}\right) - \exp\left(\kappa\frac{V^-_{Si}}{V_i}\right)}{\exp\left(\kappa\frac{V^+_{Si}}{V_i}\right) + \exp\left(\kappa\frac{V^-_{Si}}{V_i}\right)} = -2.1B \cdot \tanh\left(\kappa\frac{V_{Si}}{2V_i}\right)
\]  

(4.4)
Thus combining (4.2), (4.3) and (4.4) it follows that the difference input voltage can be written as

\[ V_{si} = \frac{2}{\kappa} \cdot \tanh^{-1} \left( \frac{I_i}{2 I_B} \right) \]  

(4.5)

where \( I_i \) is the differential input current and \( V_{si} = V_{si}^+ - V_{si}^- \). Accordingly, the normalised function of the input layer synapse is obtained by combining (3.25) and (4.5) and can be expressed as

\[ p_K = \frac{V_{wi} \cdot I_B \cdot (1 - \kappa)}{8 I_Z \cdot V_t} \cdot I_i \]  

(4.6)

where \( I_i = I_i / 2 I_B \). This expression confirms the linearisation of the input layer synaptic function with the weight given as: \( W_i = V_{wi} \cdot I_B \cdot (1 - \kappa) / 8 I_Z \cdot V_t \). The other benefit of using a horizontal resistance is that a differential current is needed to drive the input connection and consequently limits external noise pickup. Note that (4.5) is valid for differential input currents smaller than \( |2 I_B| \). The simulated static transfer characteristic of the HRes are presented in Fig. 4.5 and confirm the \( \tanh^{-1} \) distortion. Extra simulation, whose results are depicted in Fig. 4.6, have also been conducted to confirm that when a TTS is fed by an HRes its characteristics are linearised.

A series of Monte Carlo SPICE simulations of the HRes circuit suggested that for a null differential input current one can expect a 4 mV output offset due to mismatching effects. Observe that this standard deviation is noticeably smaller than that of the previously discussed scheme. This result may be explained by the fact that in the latter circuit the biasing sources have been implemented using a multiple transistor design [19], [65].
4.1.4 The output layer activation cell

Having combined the thresholding and synaptic functions, it has previously been mentioned that the neurons of the output layer are deprived of their squashing functions, see Fig. 3.11 for more details. To conform to MLP design, these activation functions can readily be implemented utilising extra TTS circuits for which the weight values are set to one. These specific TTSs will be referred to as Output Neurons (ONs). Implementing the output layer neuron function using a TTS results in a differential current output which also reduces external noise pickup. Furthermore the structure of the TTS has been slightly modified to introduce an internal current amplifier thereby allowing an external resistive/capacitive load to be driven directly. The scheme and its symbolic representation are shown in Fig. 4.7 and consists of a basic TTS cell and additional current mirrors. The gain of the amplifier has been limited to eight, mainly for reasons of power consumption and silicon area. Over the linear range of $V_{\text{w}}$, the differential output current of the ON is determined by the transfer characteristic of the TTS (3.25) times the gain of the current amplifier which may be approximated as

$$I_o = 8gV_{\text{w}} \tanh \left( \kappa \frac{V_{\text{Si}}}{2V_t} \right)$$  \hspace{1cm} (4.7)

Since a multiple transistor technique has been employed to minimise mismatching in the current amplifier circuit, one can therefore expect the standard deviation of the difference output current offset of the ON to be roughly that of the TTS times the gain, which is an approximation that has been confirmed by a series of Monte Carlo simulations.
4.1.5 The core of a feedforward neural network

As shown in Fig. 4.8 the core of a fully connected feedforward ANN can be efficiently implemented in analogue low power VLSI technology using:

- TTS circuits to simulate synaptic and sigmoid thresholding functions;
- Common bus bars to achieve the summation operations;
- Non-linear resistive elements to convert the summed currents to potentials as required by the following layer of TTSs.
- HRes to restore the linearity of the input nodes; and
- ONs to emulate the activation function of the output layer neurons.

It may be noted that, since both the inputs and outputs of the network are compatible, the design technique allows for recurrent neural networks, such as the Hopfield network, to also be implemented.

4.2 Weight storage scheme

So far, it has been theoretically shown that, based on the principle of distributing the thresholding function over the next layer of synapses, the basic processing functions of feedforward and possibly recurrent neural networks can readily be simulated, in a systolic manner, utilising TTS, load, HRes and ON analogue circuits. However, before proceeding to the design of a prototype neural network chip, it is necessary to consider the issue of analogue weight storage.
The design strategy for the weight storage scheme is based on the need to maintain the weight quiescent potential at a fixed level so as to insure maximum swing capability (3.20).

4.2.1 Capacitive storage

It may be noted that in the proposed neural network design the weight values are controlled in a bipolar fashion by the differential potentials $V_{wi}$ at the bulk terminals of the transistors forming the core of the TTS cells (3.28). As mentioned in chapter 2, ideally the analogue memory cell in a neural network design would be characterised by nonvolatility, a high resolution (at least 8 bits), easy access, high update speed (needed in the learning phase) and minimum use of die area. Unfortunately, the analogue storage device that satisfies all of these conditions has yet to be developed.

To surmount this issue, the proposed solution was to thoroughly investigate the characteristics of a sample/hold circuit, in order to fully optimise its performances.

As mentioned in chapter 2, the basic capacitive sample/hold system shown in Fig. 4.9 which consists of a PMOS switching transistor $M_S$ and an holding capacitance $C_{\text{hold}}$ suffers from both charge injection and charge leakage. These effects can severely limit the resolution of the weight signal when corresponding sensing systems are biased in the weak inversion region (3.1). Note that the holding capacitor can be implemented using an NMOS transistor since the quiescent component of the weight signal will always be much greater than $V_T$.

4.2.1.1 Charge injection

Throughout the sampling period the switching transistor conducts, thus a finite amount of hole carriers are stored in its channel region to form the inversion layer. During the
turning-off transient, when the gate voltage of $M_s$ rises from ground toward $V_{dd}$, these charges evacuate the channel through the source and the drain connections. Thus the packet of charge injected onto the holding capacitor induces an error voltage. Although the inversion layer disappears as the source-to-gate voltage of the switching device drops below $-V_T$, the offset voltage continues to rise, since charges are still fed through via the gate-source overlap capacitance, until the clock signal reaches a steady state (i.e. $V_{dd}$).

Based on the lumped model suggested by Sheu et al. [89], the magnitude of the error voltage induced on the holding capacitor in Appendix E can be approximated as

$$
V_e = \sqrt{\frac{\pi U C_{\text{hold}}}{2 K_p}} \left( \frac{C_{ov} + \frac{1}{2} C_G}{C_{\text{hold}}} \right) \text{erf} \left( \sqrt{\frac{K_p}{2 U C_{\text{hold}}}} \cdot (V_w + V_T) \right) \cdot \left( V_{dd} - [V_w + V_T] \right) + \frac{C_{ov}}{C_{\text{hold}}} (V_{dd} - [V_w + V_T])
$$

(4.9)

where $U$ is the rising-rate of the clock (assumed to be constant), $K_p$ is the transconductance parameter of $M_s$, $C_{ov}$ is the gate overlap capacitance, $C_G$ is the gate capacitance, $V_w$ is the input signal voltage and $V_T$ is the threshold voltage of the switching device which is $V_T = V_{TO} - \gamma \left( \sqrt{V_{dd}} - V_B \right) - \sqrt{\phi_B}$.

For a minimum feature size switching transistor, an holding capacitor of 1 pF and a switching-off period of 5 nS, Fig 4.10 shows the level of clock feedthrough voltage predicted by (4.9) and a SPICE simulation. These results indicate that over the weight signal range one can expect a maximum switch-induced error voltage of the order of 5 mV, thus limiting the resolution of the weight to 6 bits.

In order to improve the accuracy of such a storage scheme, a standard dummy compensation technique has been considered [90]; see Fig. 4.11, which consists of an
additional half-width short-circuited transistor \( M \), driven by an antiphase clock \((\phi)\). Following [89] the effect of clock feedthrough in Appendix F is given by

\[
\nu_e = \sqrt{\frac{\pi \cdot U \cdot C_{\text{hold}}}{2 \cdot K_p \cdot 2 \cdot C_{\text{hold}}}} \cdot \exp \left[ \frac{K_p}{U \cdot C_{\text{hold}}} \left( V_{dd}^2 + 2 \cdot V_{TW}^2 - 3 \cdot V_{dd} \cdot V_{TW} \right) \right] \cdot \left[ \text{erf} \left( \sqrt{\frac{K_p}{2 \cdot U \cdot C_{\text{hold}}} \cdot V_{TW}} \right) - \text{erf} \left( \sqrt{\frac{K_p}{2 \cdot U \cdot C_{\text{hold}}} \cdot (2 \cdot V_{TW} - V_{dd})} \right) \right] - \frac{C_{ox}}{2 \cdot C_{\text{hold}}} \cdot (V_{dd} - V_{TW})
\]  

(4.10)

where \( V_{TW} = V_T + V_w \).

Given that the \( \text{erf}(\cdot) \) function may be approximated as

\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \cdot (1 - \frac{x^2}{3})
\]

for \( x >> 1 \)  

(4.11)

Then for a fast switching-off transient, i.e. \( U >> K_p \cdot V_{TW} / 2 \cdot C_{\text{hold}} \), (4.10) may be simplified to

\[
\nu_e = -\frac{C_{ox}}{2 \cdot C_{\text{hold}}} \cdot (V_{dd} - V_{TW}) \cdot \left( 1 - \exp \left[ \frac{K_p}{U \cdot C_{\text{hold}}} \cdot (V_{dd} - V_{TW}) \cdot (V_{dd} - 2 \cdot V_{TW}) \right] \right)
\]  

(4.12)

It can be appreciated that, within this condition of operation, for an input signal set to \((V_{dd} / 2) - V_T\) the error voltage induced by the turning off of the switching transistor is completely compensated for by the dummy device. However, for a 4.8\( \mu \)m / 2.4\( \mu \)m switching device, a switching-off period of 5 ns and an holding capacitor of 1 pF, as shown in Fig. 4.12, (4.12) and a SPICE simulation show that the absolute error in the full range of stored voltage is reduced to a level where the required 8 bits weight resolution is readily achievable.
4.2.1.1 Charge leakage

As mentioned in the previous chapter, the problem of leakage may be addressed by continuously refreshing the weight information using on-chip circuitry that converts value stored in digital memory into an analogue signal. However, the number of parasitic leakage paths for the dummy compensated sample/hold circuit is more than double that of the simple scheme. Assuming a total leakage current of 5 pA, the voltage stored onto the capacitor would need to be restored once every 400 μs if an error of no more than 2 mV is to be allowed. In order to reduce this refreshing frequency and consequently power dissipation, a twin-capacitor structure has been adopted, although such a choice does impact on die size.

Assuming that both storage arrangements have matched characteristics (i.e. similar leakage current), the weight signal is influenced and therefore needs to be updated when one of the capacitors has reached complete discharge. Under worse case conditions, that is when \( V^* \) is set to 2.8 V, a refreshing cycle of 40 ms could then be tolerated.

4.2.2 The refresh mechanism

The differential weight potential is generated using the on-chip circuit shown in Fig. 4.13. The structure consists of a DAC, two silicon resistors and a pair of buffer amplifiers. Row and column decoders direct the differential voltage to the refreshed TTS. In our case, the digital weights are stored in an external Random Access Memory (RAM), however, this component could easily be integrated within the chip. The maximum refreshing speed is mainly dependent on the RAM access time, the conversion time and the time required to charge the storage capacitor. In the switch-on mode, \( M_s \) is biased in triode mode region and the charging time constant is
\[ \tau = R_s C_{\text{hold}} = \frac{1}{K_p (V_{dd} - V_{TW})} \cdot C_{\text{hold}} \]  

(4.13)

where \( R_s \) represents the channel resistance of the path transistor. For \( V_w \) set to 2.8 V, the maximum time constant is approximately 22 ns, therefore the holding capacitor is fully refreshed to within 1% of its desired value in 75 ns, (i.e. 5\( \times \tau \)).

### 4.2.2.1 Digital-to-analogue converter

For reasons of compactness and speed, the DAC uses a binary-weighted current source technique, see Fig. 4.14. The basic current multiplier consists of 70 identical transistors \( M_1 \) to \( M_{70} \) with common gate and source nodes. Note that the LSB is implemented using four transistors connected in series whereas the Most Significant Bit (MSB) consists of thirty-two devices combined in parallel. This design approach was considered since it limits the number of transistors while the influence of the body effect on the weight of the LSB still remains negligible. The currents are controlled by PMOS current switches.

The clock feedthrough associated with the bit switch, is compensated by the addition of a dummy transistor in the output current line. To limit the influence of the output voltage swing, the output conductance of the transistors forming the current sources of the four most significant digits (\( b_3-b_0 \)) have been decreased utilising a feedback amplifier as in the simple regulated cascode circuit [101]. The output current may be expressed as

\[ l_{DAC} = \frac{l_f}{4} \sum_{i=1}^{8} 2^{i-1} \cdot b_i \]  

(4.14)

where \( b_i \) is either "0" or "1" according to the binary state of \( B_i \) and \( l_f \) is the biasing current (set to 80 nA). A further bank of transistors, not shown in Fig. 4.14, generates a reference current, \( l_{ref} \), identical to the most significant binary-weighted source.
The linearity error associated with random mismatches in the conversion elements is critically important. Hence, the circuit yield is a function of the matching accuracy of the current sources. Following Lakshmikumar et al. [81], the yield of the proposed structure may be expressed (Appendix G) as

\[
G = \prod_{j=1}^{254} \text{erf}\left(\frac{63 + \frac{1}{2} + \frac{1}{4}}{\sqrt{2} \cdot (255-j) \cdot \sqrt{2} \cdot \frac{\sigma_{I_F}}{I_F}}\right)
\]

(4.15)

where \(\sigma_{I_F} / I_F\) represents the standard deviation for a unit current source in percent. As can be seen in Fig. 4.15, a yield of 100 percent can therefore be achieved for a matching standard deviation in the unit current source of about 1.5 percent. However, Pelgrom et al. [82] indicate that such standard deviation can only be obtained for devices operating in the saturation mode of conduction. Within this condition of operation and assuming the variance of the transconductance factor \((\sigma^2(K_n))\) to be negligible with respect to the variance of the threshold voltage, the standard deviation in the current source may be approximated by

\[
\sigma_{I_F} = \frac{2 \cdot A \cdot V_{TO}}{\sqrt{W \cdot L \cdot (V_{gs} - V_T)^2}}
\]

(4.16)

where \(W\), \(L\) and \(V_{gs}\) are receptively the channel width, length and gate-to-source potential of the transistor used as a current source. Given that for a saturated n-type MOS device \((V_{gs} - V_T)^2 = 2 \cdot I_d \cdot L / W \cdot K_n\), the length of the transistor for a given standard deviation may be expressed as

\[
L = 2 \cdot A \cdot V_{TO} \cdot \sqrt{\frac{K_n}{2 \cdot I_d \cdot \sigma_{I_F} / I_F}}
\]

(4.17)
It may be pointed out that in the saturation mode of conduction, the standard deviation of the drain current is inversely proportional to the drawn length ($L$) of the device. In our process $K_N$ is nominally $51.7 \times 10^{-6} \, \text{A/V}^2$. Consequently, to obtain a DAC with an integral non-linearity of $\pm 1$ LSB with a 100% yield requires a length greater or equal to 70 $\mu$m. The smallest possible size transistor ($W/L = 3.6 \mu\text{m} / 70 \mu\text{m}$) was used to build the current multiplier and a common centroid layout adopted to further reduce mismatching effects.

### 4.2.2.2 Operational amplifier

The DAC outputs are converted to differential potentials via two silicon resistors. To ensure high speed operation and sufficient current driving capability, two op amps configured as buffers interface the DAC-resistor structure and the weight storage mechanism. With the bias supply-voltage equal to 3 V and the common-mode weight inputs set at approximately 2.55 V the op amps are required to operate at near top rail input and to provide a wide unity-gain bandwidth to allow for fast settling time on low load capacitor (1 pF), and dissipate as little as possible power, within a small die area. A structure offering most of these features is shown as Fig. 4.16, and is based on a design described by Hogervorst et al., [102]. This two-stage structure comprises a folded cascode input stage and a class-AB output stage. Since common mode input bottom rail operation is not needed in our system, the P-channel differential input pair and the $g_m$ control circuitry have been omitted leading to a more compact and simple structure. The initial choice of device aspect ratios was determined by power dissipation, slew-rate, silicon area and input offset considerations. Final drawn values were developed using a SPICE simulator.
The simulated open loop frequency response of the op amp loaded by a 10 kΩ resistor and 10 pF capacitor is shown in Fig. 4.17. The result indicates that a unity-gain bandwidth of approximately 3.5 MHz and a unity-gain phase margin of 63° are achieved for a biasing current of 2.5 μA. Extra simulations have also confirmed that, configured as a unity gain buffer, the amplifier offers adequate performances even when operating in near top-rail common-mode input signal. Note that a multiple transistor technique was used in order to limit the effects of process parameter variations and consequently minimise the offset voltage of the input stage. A batch of Monte Carlo simulations has suggested that a 5 mV input offset can be expected. However this error level can be reduced if one considers using a common-centroid layout structure.

4.2.2.3 Decoders

Since the chip includes 98 TTSs and 5 ONs, analogue weight information is necessarily refreshed and updated in a sequential manner. The clock signal (ϕ) for the access switch in each TTS (or ON) is generated by addressing the appropriate row and column. To reduce the number of external pin connections to a minimum, the signals are generated using a 3-to-8 row decoder and 4-to-14 column decoder based on standard commercial designs. However, to synchronise the clock signal with the analogue weight information, an extra control command line has been added to the decoders.

4.3 Simulation of a neural network

The capacities of the proposed designs (i.e. TTS, load, ON, HRes, capacitive storage and weight refreshing scheme) have been assessed via a whole neural network simulation in SPICE. To limit the size of the complete system, as determined by the
number of transistors, and consequently simulation time a simple 2:2:1 MLP performing an XOR function was selected. The structure of the network is shown in Fig. 4.18 (a) and comprises 3 input nodes (including bias), 2 hidden processing units and an output neuron and is emulated, in Fig. 4.18 (b) using 3 HRes, 9 TTSs, 3 I/V converters, an output neuron and a complete weight refreshing scheme. For reasons of simplicity, the holding capacitors are not represented in the diagram.

An extra TTS has been added to each neuron to bias the weighted sum. The inputs to these bias synapses were set to unity via an HRes. Note that since this input level is common to all biasing TTSs a single bias input has therefore been used. This biasing technique will also be applied to the design of a neural network chip since:

- It limits the number of biasing inputs to one and therefore reduces the amount of HRes cells;

- Power dissipation is consequently diminished, while processing speed is unaffected since the bias is kept to a constant level; and

- The number of connection pads is also minimised.

To avoid having to train the analogue network, the sequences of digital data applied at the inputs of the DAC were determined utilising a set of weight values, given by a trained neurocomputer, which were then converted using (3.28) and (4.14). Fig. 4.19 shows the performances of the simulated XOR network. The first 20 µs of the simulation was dedicated to charge each weight storage capacitor pair to its appropriate difference analogue voltage using time multiplexing. The four distinct input patterns in Fig. 4.19 (a) were subsequently fed to the network in a sequential manner. It may be noted that as expected, the propagation time of the 2 layer network is approximately
4 µs since the maximum weight value was set to 10, i.e., the bias current of the TTSs was 250 nA and $I_{z} = 100$ nA.

Extra simulations have also shown that the complete refreshing scheme (i.e., DAC, op amp and sample/hold cells) offers a suitable level of functionality and that one can expect a maximum refreshing speed of 1 µs. This result also suggests that one of this structure could be used to refresh or update, in a time multiplexed manner, the weight values of 40 000 TTSs, since a 40 ms holding time is expected.

It may be appreciated that a physical silicon implementation would not necessarily exhibit the performances predicted either by simulation or theoretical results for several reasons including modelling imperfections and process/device tolerances. Therefore the above mentioned results should be regarded as guidelines.

### 4.4 The neural network chip

A prototype chip was fabricated using the Eurochip 2.4 µm double metal, double poly p-well type process. Since information characterising the behaviour of subthreshold mode devices was unavailable, much attention has been directed to the design of individual test structures. The primary aim of this prototype design was to demonstrate that the proposed technique along with the suggested circuits can be trained to perform practical pattern recognition tasks. The neural network chip consists of:

- Two TTSs, an ON and two I/V converters as test structures;

- An XOR network comprising 3 source nodes (including bias), 2 hidden processing neurons and 1 output unit;

- An 10:6:3 MLP;
• A row and column decoder for weight addressing;

• A DAC, two silicon resistors and two differential amplifiers for weight refreshing; and

• Two biasing generators which provide the bias potentials required by the TTS, HRes, ON and load circuits.

Note that no HRes circuit was integrated for test purpose, that is because its transfer characteristic will readily be available from any inputs of the two networks.

4.4.1 Exclusive OR network

The XOR function has been included for test purposes since its simple but highly non-linear characteristics serves as a benchmark problem. As mentioned earlier, the network consists of 3 input nodes (including bias), 2 hidden processing units and an output neuron and is implemented using 3 HRes, 9 TTSs, 3 I/V converters and an ON. The outputs of the 2 hidden processing units were made available for measurement purposes via two pairs of CMOS switches. These switches were utilised so that the capacitive loads associated with the bonding pads could be isolated from the network and consequently processing speed would be optimised.

4.4.2 QRS complex detector

To further demonstrate the capability of the proposed approach, a 10:6:3 perceptron, trained to perform the seemingly complex function of detecting the QRS complex of foetal electrocardiogram (ECG) has been included. This type of application has already received much attention [80]. It may also be noted that this exercise is similar to the cardioverter defibrillators [79].
Although Ifeachor et al. [80] results suggest that a 20:6:1 feedforward network performs marginally better than a 10:6:1 MLP in detecting QRS complexes, the latter scheme was however integrated in order to comply with wafer fabrication process limitations (i.e. limited amount of bonding pads). Nevertheless an extra two output neurons were added so that the quality of the detected QRS complexes could be classified in three different categories (i.e. good, average and poor). The complete structure of the QRS complex detector is shown in Fig. 4.20 and was simulated using 87 TTSs, 11 HRes, 9 I/V converters and 3 ONs.

4.5 Layout techniques

To exploit the modular structure of fully connected feedforward neural networks, a full custom layout approach [19] was adopted for generating the masks of the basic building blocks described earlier in this chapter. In order to minimise the complexity of inter-cell connections and consequently enhance device density, signals generally shared by adjacent circuits were distributed via banks of horizontal and vertical metal wires. This layout strategy, as shown in Fig 4.21, allows for the implementation of any size matrix of unit elements. At the circuit level, to achieve acceptable matching between critical devices (i.e. differential-pair transistors and mirror transistors) several considerations were taken into account:

- The devices were integrated as close as possible to each other and arranged with the same orientation; and

- A common-centroid-symmetry layout design was employed, in which devices are connected around a central point.
Since the chip includes both analogue and digital circuits, to attenuate the noise interaction attention was paid to the:

- Use of separate analogue and digital supply connections;
- Maximum physical separation of the analogue and digital elements where possible, see floor plan of the chip in Fig. 4.22;
- Use of guard rings.

Since a large capacitance is usually associated with the bonding-pad diodes of an electrostatic discharge protection cell, all the analogue inputs and outputs of the chip were left unprotected from static electric charges so as to maximise the processing speed of the networks. Note that some protections are however offered by the reverse-biased source/bulk and drain/bulk junctions of the transistors incorporated in the HRess and ONs.

The chip includes 100 pins and the circuits occupy an active area of 2.79mm × 2.26mm (containing approximately 10 000 transistors and 214 capacitors). A photomicrograph of the prototype chip is shown in Fig. 4.23 and its principal features are given in Table 4.1.

4.6 Summary

The complete structures of the building blocks making-up the core of feedforward ANNs have been presented and their vulnerability to process parameter variations evaluated via series of Monte Carlo SPICE simulations. Results indicated that the expected imperfections in the basic processing cells may be compensated for during training.
The non-linear input and un-thresholded output problems, associated with distributing the squashing function over to the next layer of synapses, have respectively been overcome by introducing an HRes and an ON which also facilitate compatibility between the inputs and outputs of the network and decrease vulnerability to external noise pickup.

A dummy-compensated sample-and-hold circuit has been suggested as a means to implement the storage scheme since it readily allows for an 8 bits weight resolution. To control the weight voltages an on-chip update/refresh mechanism consisting of a DAC, two silicon resistors, a pair of op amps configured as buffers and a row and column decoder has been developed.

The behaviour of the suggested ANN design has been assessed via a series of SPICE simulations of the XOR network. Results revealed a reasonable level of performance. Finally a prototype neural network chip has been developed in order to experimentally evaluate the performance of the proposed feedforward ANN design technique.
Figure 4.1: The complete circuit of a transconductance-thresholding-synapse.

Figure 4.2: Array of TTS circuits.
Figure 4.3: The complete circuit of the load.
Figure 4.4: Horizontal resistor.
Figure 4.5: Static transfer characteristic of the horizontal resistance.

Figure 4.6: Static transfer characteristics of the transconductance-thresholding-synapse fed by an horizontal resistance.
Figure 4.7: Circuit diagram of the output neuron.
Figure 4.8: The complete analogue implementation of a fully connected feedforward neural network.
Figure 4.9: Basic PMOS-based sample-and-hold circuit.

Figure 4.10: Level of clock feedthrough voltage of a single 2.4/2.4 μm PMOS switch. Switching speed $U = 3 \text{ V} / 5 \text{ ns}$, holding capacitor $C_{\text{hold}} = 1 \text{ pF}$.

- Calculated
- Simulated
Figure 4.11: PMOS-based dummy-compensated sample-and-hold circuit.

Figure 4.12: Level of clock feedthrough voltage of a PMOS-based dummy-compensated sample-and-hold circuit. Switching speed $U = 3 \text{ V} / 5 \text{ ns}$, holding capacitor $C_{\text{hold}} = 1 \text{ pF}$, $M_s = 4.8 / 2.4 \mu\text{m}$.
- Calculated
- Simulated
Figure 4.13: The complete refreshing mechanism.
Figure 4.14: Circuit diagram of the digital-to-analogue converter.
Figure 4.15: DAC yield versus current-source mismatch.
Figure 4.16: Circuit diagram of the differential amplifier.
Figure 4.17: Simulation results of the open-loop frequency response of the differential amplifier.
Figure 4.18: XOR neural network.
Figure 4.19: Simulated transient characteristics of the XOR network.
(a) Logic table of the XOR gate.
Figure 4.20: Structure of the QRS complex detector.
Figure 4.21: Layout of a matrix of TTSs.
Figure 4.22: Floor plan of the neural network chip.
Figure 4.23: Photomicrograph of the prototype chip.
<table>
<thead>
<tr>
<th>Fabrication process</th>
<th>2.4-μm CMOS double-poly double-metal</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTS unit size</td>
<td>174μm x 147μm</td>
</tr>
<tr>
<td>I/V converter unit size</td>
<td>147μm x 102μm</td>
</tr>
<tr>
<td>Output-neuron unit size</td>
<td>202μm x 176μm</td>
</tr>
<tr>
<td>Horizontal-resistance unit size</td>
<td>174μm x 87μm</td>
</tr>
<tr>
<td>DAC size</td>
<td>405μm x 382μm</td>
</tr>
<tr>
<td>OpAmp size</td>
<td>444μm x 209μm</td>
</tr>
<tr>
<td>Silicon resistances</td>
<td>201μm x 145μm</td>
</tr>
<tr>
<td>Row decoder size</td>
<td>468μm x 277μm</td>
</tr>
<tr>
<td>Column decoder size</td>
<td>944μm x 296μm</td>
</tr>
<tr>
<td>Number of TTSs</td>
<td>98</td>
</tr>
<tr>
<td>Number of I/V converters</td>
<td>14</td>
</tr>
<tr>
<td>Number of output-neuron</td>
<td>5</td>
</tr>
<tr>
<td>Number of Horizontal-resistance</td>
<td>14</td>
</tr>
<tr>
<td>Weight memory device</td>
<td>capacitor</td>
</tr>
<tr>
<td>Active die size</td>
<td>2.79mm x 2.26mm</td>
</tr>
<tr>
<td>Supply voltage</td>
<td>3V</td>
</tr>
<tr>
<td>Package</td>
<td>100-pin PGA</td>
</tr>
</tbody>
</table>

Table 4.1: The chip features.
Chapter 5

Performance of the prototype neural network chip

The potentials of the various neural network modules have been analytically estimated in the previous chapter and simulation studies have also indicated similar levels of practicality. In this chapter, the performances of these elements are experimentally assessed and compared with those mentioned above. To this end special test set-ups have been developed which allow for measurements to at least 1% accuracy. A thorough review of the potentials of various learning algorithms suitable for analogue neural structure in-loop is presented. The objective of this analysis is to examine the basic concept of these algorithms in order to highlight their performance level in terms of learning quality, flexibility, convergence speed, and hardware cost and consequently select the most suitable for the given neural network structure. A section of this chapter will also be dedicated to introducing the experimental set-up that has been developed to monitor the training and testing of the analogue neural networks. The performances of the XOR and QRS complex detector networks are subsequently presented and compared with those predicted by the analysis and SPICE simulations. Finally, these
results will be used to judge the level of usefulness of the proposed implementation technique in simulating fully connected feedforward neural networks.

5.1 Experimental characteristics of the basic building modules

An assessment of the performances of the analogue neural network modules, i.e. TTS, I/V converter, ON, HRes, weight storage scheme and refreshing mechanism are presented in this section. Ten samples of the prototype chips were made available by MIETEC for test purposes, thus allowing for an evaluation of the effects of local and global process parameter variations on the test structures.

5.1.1 Transconductance-thresholding-synapse

For measurement purposes, the differential output currents of the individual TTSs were converted to differential potentials via the use of a pair of matched transimpedance amplifiers consisting of a low noise JFET input op amp and a feedback passive resistor of 1 MΩ. As shown in Fig. 5.1, these outputs were subsequently sensed using a difference amplifier whose gain was set to 5. The differential input potential at the gate terminals was applied externally using a function generator, whereas the weight differential voltage was generated using the on-chip DAC-op amp arrangement and directed to the bulk inputs of the TTS under test by the weight address decoders.

Experiments were carried out over a wide range of biasing currents and confirmed that the gain factor of the differential output current is proportional to $I_b$. The measured static transfer curves relating the output voltage $V_{TS}$ to the input voltages $V_{si}$ and $V_w$ are shown in Fig. 5.2 (a) and (b), respectively and correspond to a bias current of 250 nA.
These transfer characteristics were obtained using the X-Y mode measurement facility of a digital oscilloscope and show excellent correlation with the simulation results of Fig. 3.7. The measurements also confirm that over the linear range of the bulk differential input potential, the differential output current can be approximated using (3.25). Although the bias current was set to 250 nA, the gain of the TTS is about 50% higher than anticipated. This increase may principally be due to the significant distance (approximately 2 mm) between the biasing generator and the TTS under test. Distance of this order will exacerbate parameter variations associated to unavailable non-uniform doping.

Extensive measurements on all available TTSs have also been conducted to establish the magnitude of output offset current when both inputs are short-circuited. To ensure accuracy, the offset contribution of the test set-up was evaluated and accounted for during the experiments. The results shown in Table 5.1 indicate a standard deviation of approximately 12 nA which closely agrees with the prediction.

5.1.2 Current-to-voltage converter

Since the TTS biasing current is set to 250 nA, to obtain a weight range of ± 10 the non-linear load must display a driving point resistance of approximately 15 MΩ. To measure the transfer characteristic of such a high-value resistor the experimental configuration shown in Fig. 5.3 was used in order to minimise external noise pickup and measurement distortion. The floating current source was implemented using an arrangement of high-precision passive resistors and amplifiers and controlled by a triangular shape voltage signal of 0.4 V_{pp}. The scheme exhibited a 1 μA/V transconductance. The generated differential potential was subsequently detected using low noise JFET input op amps operated as voltage followers driving a difference amplifier.
whose gain was also set to 5. Since the test set-up loaded the resistive element with a substantially large capacitor (i.e. at least 5 pF) measurements were conducted at 10 Hz so as to avoid frequency related attenuation.

Results suggested, as shown in Table 5.2, that the nominal resistance (i.e. the slope resistance measured at the origin) is inversely proportional to the bias current $I_2$ over the range 15 nA ... 60 nA, thus facilitating the control of the weight range. It has also been observed that within this range of biasing current the nominal resistance is three times smaller than expected. This variance may be caused by either batch-to-batch or wafer-to-wafer rather than local or global die process parameter variations, since the scale of error was approximately the same for all 10 prototype chips. Fig. 5.4 depicts a family of measured I-V static characteristics obtained for a bias current of 30 nA and supply voltages ranging from 2.9 V to 3.1 V in equal 100 mV increments. These measurements validate that the transfer characteristic is relatively insensitive to power supply variations and displays a sinh⁻¹ non-linearity.

Extra tests have also indicated that, as shown in Table 5.3, the load arrangement is less vulnerable to local process parameter variations than anticipated. The standard deviation of the output offset voltage is of the order of 5 mV. These experiments were conducted with the floating current source removed from the above mentioned test set-up.

5.1.3 Horizontal resistance

The experimental arrangement, shown in Fig. 5.5, consisted of an inverting amplifier where the device under test (i.e. HRes) together with a passive resistor defined the closed loop gain. Given that the expected value of the horizontal resistance, within its linear range, is in the order of 210 kΩ, a passive resistor of 200 kΩ was employed to
obtain an absolute gain of approximately one. To fix the quiescent common-mode input level to \( V_{dd}/2 \) the non-inverting input of the op amp was set to 1.5 V. With a triangular shape signal, having an amplitude of 0.2 \( V_{pp} \) and a common-mode component of 1.5 V, applied at the input of the structure, the current to voltage static characteristic of the horizontal resistance shown in Fig. 5.6 was obtained for a bias current \( I_b \) of 250 nA. With the gain of the inverting amplifier chosen to be one, the settings for the digital oscilloscope were 50 mV/div in X mode and 50 mV/div in Y mode which was also inverted. The results also indicate close agreement with the theoretical prediction (4.5), i.e. the HRes exhibits the \( \tanh^{-1} \) distortion and a nominal resistance of approximately 250 kΩ. It has also been noted that, as expected, the resistance value is a function of the controlled bias current generator; the higher the bias current, the lower is the resistance.

The performance of a TTS fed by an HRes was experimentally assessed utilising the set-up presented in Fig. 5.1 wherein the function generator was substituted by the arrangement previously described. The experimental evidence, whose results are depicted in Fig. 5.7, confirms the predicted linearisation of the input of the TTS for differential input current within ±2×\( I_b \).

The output voltage offset measurements were conducted on the HRes associated to the A input of the XOR network of each prototype chip. The results are given in Table 5.4 and exhibit a standard deviation of 4 mV which closely agrees with the value predicted by the Monte Carlo study.

**5.1.4 Output neuron**

Measurements were conducted utilising an experimental set-up similar to that used for characterising the TTS, with the exception that the value of the feedback passive resistances of the transimpedance amplifiers were halved so as to contain the output
signal well within the supply voltages of the op amps since the differential output current is magnified. With the biasing current set to 250 nA, Fig. 5.8 (a) and (b) present the static measurements which relate the differential output current to both differential input voltages $V_{si}$ and $V_w$ respectively. These results, when compared with those of the TTS, confirm that the gain of the internal current buffer of the ON is about 8. This outcome was expected because both structures are juxtaposed on the die, thus their biasing currents must be similar. One can therefore conclude with some certainty that the 50% gain error of the TTS, as mentioned earlier, and that of the ON is due to global process parameter variations.

As predicted by the series of Monte Carlo simulations, the standard deviation of the output current offset of the individual ONs, given in Table 5.5, is 97 nA which is approximately that of the TTS's times the gain of the internal current amplifier.

To determine the driving capacity of the ON a transient experiment was conducted. The step response of the ON exhibited in Fig. 5.9 was acquired with the structure supplied by a square wave signal of 0.4 $V_{pp}$ embedded in a 1.5 V quiescent common-mode component and loaded by its associated detecting scheme. This measurement indicates, for a maximum positive weight signal, a propagation delay of 1.5 µs and a transition time of 1.4 µs both of which are mainly due to the limitations of the op amps (i.e. slew rate) utilised within the sensing arrangement.

### 5.1.5 Weight storage scheme

As foreshadowed, the holding time is a critical parameter for any weight storage scheme. The weight persistence was determined by applying a 200 mV differential potential at the test TTS input and loading its weight storage capacitors with a maximum positive value. To maximise subthreshold conduction during the charge
leakage process, the input of the access switches was subsequently set to the lowest possible weight control voltage. The differential output current detection was achieved utilising the sensing scheme depicted in Fig. 5.1 in order to accurately correlate the weight voltage decay to the static transfer characteristics formerly presented in Fig. 5.2; Fig. 5.10 shows the output signal variation of the TTS in the absence of refreshing. Two distinct periods may be identified. During the initial 60 seconds charges stored on both capacitors are leaked at a similar rate. Nonetheless the differential output current is slightly corrupted for the following two reasons:

- Mismatch in the leakage paths degrades the stored differential potential; and

- The gain of the TTS is influenced by the change in the quiescent common-mode weight signal. The higher the common-mode, the lower is the gain.

In the subsequent phase, the capacitor precharged with the most positive weight signal $C_{\text{hold}}$ has reached complete discharge (i.e. $V_{w}^{+} = V_{dd}$). However charges still flow between the complementary capacitive element $C_{\text{hold}}$ and the substrate, thus reducing the magnitude of the differential weight voltage and consequently that of the output signal at a much quicker rate. The maximum holding time, at room temperature, corresponds to the time taken for the output to drop by $1/127$. Measurements indicate that a refresh cycle time of 5 seconds would be acceptable. Results also suggest a mismatch in the leakage paths of approximately 0.4 fA and a leakage current per path of about 5 fA rather than 5 pA as anticipated. In the event of a design based on a single capacitive storage device this latter result implies that the weight would need to be restored at least three times per second.

Charge injection also has an important influence on the weight storage mechanism. As noted earlier, clock feedthrough has the effect of modifying the charge stored onto
the holding capacitors and occurs during the switch-off transient of the access transistors, thereby inducing an error in either the differential output of a TTS or an ON. The level of switched induced error voltage was experimentally evaluated utilising the test set-up presented in Fig 5.1. A differential potential of 100 mV was applied at the input of the individual TTS while a null weight was loaded onto its holding capacitors. Fig. 5.11 shows the variations in the TTS output signal when the weight is continuously refreshed at a rate of 80 times a second. This measurement clearly illustrates that the quantity of stored charge is corrupted due to clock feedthrough and capacitive coupling between the decoder signal lines and the holding capacitors. To estimate the contribution of each effect additional experiments have been conducted. The results revealed that when the row and column signal lines, crossing over $C_{\text{hold}}$, are switched independently a disturbance of 30 mV and -10 mV respectively is caused, while charge injection induces a relatively larger change of 120 mV. When this is related to the TTS static transfer characteristics, it implies that the error caused by clock feedthrough occurring at the holding capacitors level is 30 mV which appears to be much larger than expected. A similar amplitude has been detected over the whole weight range, thus suggesting that the source of this interference is related to additional digital signal wires running above the top plate of $C_{\text{hold}}$ rather than charge injection. This hypothesis seems plausible as charge injection should be further compensated since a twin-capacitor memory structure is utilised.

5.1.6 Refreshing mechanism

Linearity and monotonicity in the DAC element of the weight-update circuitry are essential to the learning process. Monotonicity is assured if the difference between a bit current and the sum of the lower-order bit currents is positive and the differential
linearity depends on the magnitude of that difference. Table 5.6 shows the measured values of the weighted current sources for each of the 10 prototype DACs. For a biasing current $I_d$ of 80 nA, these measurements confirm the monotonicity in the static transfer characteristics and indicate a differential linearity of ± 1 LSB with a circuit yield of 100%. It may also be noted that when the absolute value of the DAC input changes from 127 to 128, the DACs exhibit an accuracy of ± ½ LSB. This higher than expected level of precision may be due to the fact that the bit current transistors were integrated around a common centroid point.

The dynamic characteristic of the weight refreshing system is also an important feature. The maximum settling time is the interval required to fully charge the holding capacitors when the digital stored weight changes from one extreme to the other. This refreshing speed is a determining factor of the following two fundamental limits of our analogue ANN design:

- The maximum number of TTSs that could be refreshed by a single asynchronously time-share refreshing scheme; and

- The learning time.

As depicted in Fig. 5.12, a maximum refreshing time of 3 μs was obtained for an op amp biasing current $I_{op}$ of 2.5 μA. This result indicates that for a holding time of 5 seconds a maximum number of $1.5 \times 10^6$ synaptic-weights could be refreshed/updated using the suggested design. Since this is greater than the number of synapses that can readily be integrated on a large silicon die, (i.e. 10mm $\times$ 10mm) refresh rate considerations appear not to be restrictive.
5.2 Training algorithm for analogue feedforward neural networks

As mentioned earlier the input/output mapping function achieved by a feedforward neural network is dependent upon the weight associated to each synaptic connection. To determine a weight point in the weight space that may fit the desired function, the neural network is stimulated utilising an input pattern within a set of examples of input-output pairs. The difference between the actual output and the desired target is then used to modify the free parameters of the network, which are usually initially set to small random values, in order to diminish the mapping error. This error measure for a pattern \( p \) is defined as the sum of the square errors

\[
E_p = \sum_{k=1}^{n} (T_{pk} - O_{pk})^2
\]

(5.1)

where \( n \) is the number of output units. \( T_{pk} \) and \( O_{pk} \) represent the target and actual outputs, respectively, for pattern \( p \) at the output \( k \). When this procedure is repeated for each of the training samples, one epoch is complete. This method is commonly referred to as stepwise supervised learning [18] and takes place until the averaged sum of the error measure from each input/output vector in the training set has reached an acceptable limit. The mean-squared mapping error or cost function is expressed as

\[
E_{Av} = \frac{1}{N} \sum_{p=1}^{N} \sum_{k=1}^{n} (T_{pk} - O_{pk})^2
\]

(5.2)

where \( N \) is the number of training patterns. At any given time in the proceeding, the magnitude and direction by which the connection strengths of the network are altered depends on the value of the error measure and the nature of the update rule (also known
as the algorithm) that is used to minimise it. However, since the education process is conducted on a pattern-to-pattern basis, the size of the former is usually kept small in order to avoid the loss of previously stored information. As a consequence to make the feedforward neural network emulate the teacher many training epochs are usually required, thus rendering the learning procedure intensively iterative and computational. Generally, to make the path through the weight space more stochastic, thus allowing a wider exploration of the error surface, the patterns are presented in a random fashion from one epoch to another [12].

An alternative approach to stepwise learning is batch mode. In batch mode learning the complete set of training vectors are consecutively presented to the network in order to estimate the average squared error. The weights are updated once, at the end of an epoch rather than after each training pattern, utilising (5.2). This technique may counteract the forgetting effect of stepwise learning, especially if a large number of irregular patterns are to be taught. However this is acquired at the expense of an increase in the likelihood of trapping the error surface in a local minimum since the influence of a random update sequence is not available.

A neural network is said to be trained when the error surface has reached a global minimum, i.e. the weight vector of the MLP approximates, within some degree of accuracy, the required input/output mapping function. When this state of operation is attained one can dispose of the teacher and independently use the MLP to perform the task it was trained for. However it is important to note that a neural network subject to a training procedure may never converge toward a global solution. This outcome does not necessarily indicate the non-existence of the unique set of weights needed to produce the desired mapping characteristics. However in the eventuality of failure one may consider altering either the initial starting conditions (i.e. the initial values of the free
parameters), the learning rule, the topology of the network or eventually revising the number and diversity of the training patterns [12].

The fundamental characteristic of a trained neural network is that it has the ability to generalise well [18]. That is to say that if all the training patterns were sampled throughout the mapping space, the network should be able to provide a satisfactory response to an apparently unknown input pattern by interpolation.

Several supervised learning algorithms for feedforward neural networks have been developed for neurocomputer simulators [12], [18], [34], [91-92]. However, a limited number of those are adaptable to analogue hardware implementations [51], [93-94] for various reasons that are presented later. The EBP also commonly referred to as the back-propagation is by far the most commonly used algorithm to train MLP in software on conventional digital computers. It was first described by Paul Werbos in his Ph.D. thesis in 1974 and was popularised more than a decade later in 1986 by Rumelhart and his colleagues [91]. The standard EBP learning rule [56], [60] and derivatives [55], [59] have been successfully implemented in analogue VLSI hardware in order to speed up the learning process and to render the neural network system completely autonomous. However it will be shown in the next section that the update technique is also computationally demanding. Thus, for a given silicon area, the hardware cost of integrating the EBP on-chip can severely degrade the density of primary neuronal circuits.

In the following section our experience with back-propagation and weight-perturbation techniques is discussed.
5.2.1 Back-propagation

The back-propagation algorithm is based on a first-order approximation of the steepest descent method which uses the Jacobian gradient to determine a suitable direction of movement of the weights. According to the steepest descent method, the correction applied to a weight is proportional to the negative of the gradient of the error surface with respect to that particular weight [12], [18] and is given by

\[
\Delta W_{ij}^a = -\frac{\eta}{2} \frac{\partial E_p}{\partial W_{ij}^a}
\]

(5.3)

where \(\eta\) is referred to as the learning rate. The factor \(\frac{1}{2}\) was introduced in order to simplify subsequent analytical derivations. For a feedforward neural network \(O_{pk}\) in (5.1) may be expressed as follow

\[
O_{pk} = f\left(\sum_k b_k^R + W_{kl}^R f\left(\sum_m b_m^* + W_{lm}^* f\left(\ldots f\left(\sum_n b_n^Q + W_{hn}^Q x_n\right)\ldots\right)\right)\right)
\]

(5.4)

where \(W_{kl}^R\) represent the synaptic weight that is attached to the \(k^{th}\) neuron of the output layer and is fed by the \(l^{th}\) neuron located in the antecedent layer. \(W_{hn}^Q\) is the weight vector associated to the input layer and \(x_n\) is the \(n^{th}\) input node of the network. Combining (5.1), (5.3) and (5.4) the update rule of the back-propagation algorithm is given [12], [18] by

\[
\Delta W_{ij}^a = \eta Y_{j}^{a-1} \delta_i^a
\]

(5.5)

where \(Y_{j}^{a-1}\) is the output signal of \(j^{th}\) neuron situated in the \(a-1^{th}\) layer and \(\delta_i^a\) is referred to as a local error signal and is defined as follow

\[
\delta_i^a = \begin{cases} 
f'\left(S_k^R\right) \left[T_{pk} - O_{pk}\right] & \text{if } a \text{ represents the output layer} \\
 f'\left(S_i^a\right) \sum_v \left[\delta_v^{a+1} \cdot W_{vi}^a\right] & \text{if not}
\end{cases}
\]

(5.6)
where $S^i_\alpha$ is the sum signal of the $i^{th}$ neuron located in the $\alpha^{th}$ layer and $f'(\cdot)$ is the first derivative of the thresholding function. To be able to compute this first derivative the EBP learning rule requires that the activation function of the neurons be continuous. This is the reason why the sigmoid function is by far the most commonly used activation function in the design of feedforward neural networks. It may be observed that the local error signal of a given hidden unit is a function of all of those associated to each neuron located in the subsequent layer. Hence, a complete learning iteration consists of feeding the input of the network with a training pattern and propagating the signal forward toward the output nodes. This first step is usually referred to as the forward path. The local error signal vector of the output layer is then estimated and propagated backward toward the input nodes in order to evaluate those of the hidden layers. The weights are then adjusted as follow

$$W^\alpha_{ij}(\text{new}) = W^\alpha_{ij}(\text{old}) + \eta_j Y^\alpha_{j} \delta^\alpha_i$$  \hfill (5.7)$$

Fig. 5.13 depicts the flow of information of the back-propagation algorithm for a 3:3:2 MLP. It may be noticed that as the output get closer to the target (5.6), the value of $\delta^\alpha_i$ reduces, leading to a slower progression towards the optimum. For an analogue hardware simulator such has the one described above, implementing this process with a limited weight resolution of 8 bits may present some problems. Several researchers [93-94] have stated that successful MLP gradient descent learning requires weight precision of between 8 and 13 bits. The actual number of bits depends on the problem to be solved and on the choice of $\eta$. Speed improvements and better convergence can be obtained by modifying the algorithm so as to estimate a near-optimum value of the learning parameter [12] and/or incorporating a momentum term [18]. However it can be appreciated that for each synaptic connection located in a hidden layer of the forward
path an extra three multipliers are required to update its weights. These additional computational costs make the standard back-propagation algorithm a questionable candidate for on-chip training.

To successfully evaluate the weight correction signal $\Delta W^a_{ij}$, the learning rule requires an accurate knowledge of the network connection strengths, the output state and internal activity level of each neuron and the sigmoid function derivative. For an in-loop training architecture (see section 5.3 for details) extra input/output pads would be required to provide the necessary information from the forward path to the teacher. This additional hardware cost also causes the EBP learning rule to be incompatible with chip-in-loop training. Note that for the analogue simulator described earlier the output states of the hidden neurons are not available since the thresholding and synaptic functions have been combined.

### 5.2.3 Weight perturbation

The Weight Perturbation (WP) learning algorithm was first presented by Jabri and Flower [92], and is also based on an approximation of the gradient descent method. The weight correction signal is estimated using a finite difference of the error/weight gradient which is a first order approximation of (5.3) and is given by

$$
\Delta W^a_{ij} = -\eta \cdot \frac{\partial E}{\partial W^a_{ij}} \approx -\eta \cdot \frac{E(W^a_{ij} + \text{pert}) - E(W^a_{ij})}{(W^a_{ij} + \text{pert}) - W^a_{ij}} = -\eta \cdot \frac{\Delta E(W^a_{ij})}{\text{pert}}
$$

where $\text{pert}$ is the weight perturbation signal. Note that the factor $\frac{1}{2}$ is not introduced since the gradient is approximated rather than analytically evaluated. However it is required that, for the approximation to remain satisfactory, both the perturbation and error variation are kept small in order to optimise the probability of convergence and
the level of generalisation. These conditions may cause the WP learning algorithm to converge at a slower rate than the EBP learning rule.

A complete training procedure involves feeding an input pattern to the network and evaluating the error signal (5.1). The connection strengths are then perturbed in turn and the corresponding changes in the network error are measured. Finally the weights are updated using

$$W_{ij}^{(\text{new})} = W_{ij}^{(\text{old})} - \eta \frac{\Delta E_{ij}}{\text{pert}}$$

(5.9)

Fig. 5.14 depicts the signal flow of the weight perturbation algorithm for a 3:3:2 feedforward neural network. Although the WP method is similar to the backpropagation technique, it is more suitable to analogue hardware implemented systems, because the weight update is simply a factor of the change in the network error. Note that knowledge about the non-linearity associated to the neuron and its first derivative is also no longer required. These advantages are partially offset in that for each iteration $Z+1$ forward paths are required, where $Z$ is the number of weights. Obviously, this may generate greater concern for a computer simulator than for an analogue VLSI system, since in the former the forward-path computation time may be greater than in the latter. For on-chip implementation of the WP algorithm $Z$ memory elements are required to temporally store the error variation associated to each weight. This hardware cost is clearly less complex than that of the EBP learning algorithm. It is also worth noting that for an in-loop training system, access to hidden neuron activity is not needed.

Comparing the EBP and WP algorithms, one can trade-off training speed for complexity of implementation. In our case the weight perturbation scheme appears to offer the best approach for a chip in the loop architecture.
5.2.3.1 Computer experiment

The performance of the WP learning algorithm has been assessed via extensive computer simulations for the task of detecting QRS complexes in a foetal ECG signal. The training data set included 306 input patterns of good, and poor quality QRS complexes and noise signals. Fig 5.15 shows the ECG signals from which the training patterns where obtained. A training pattern consisted of 10 samples which were extracted at even intervals from a window of 30 data points. The network, whose structure was depicted in Fig. 4.20, was trained using various perturbation and learning rate values, ranging from 0.02 to 0.32 and 0.05 to 0.2, respectively. A stepwise mode of training was adopted whereby the patterns were presented in an incremental fashion. In all cases, identical initial weight values and stopping criteria were imposed. The simulations showed that the quality of the generalisation was best when the learning rate and the perturbation values were set to 0.05 and 0.08, respectively. Fig. 5.16 shows the learning curve of the network. The training was terminated when the root mean square error of the network decreased below the set limit of 2%, and required approximately 135 epochs. During these experiments it was noted that increasing the learning rate had the effect of reducing the number of epochs required to reach the stopping criterion. However, it also had the effect of lowering the generalisation ability of the network. It may also be noted that these (computer simulated) learning experiments were achieved using high precision arithmetic and therefore "ideal" weight resolution and dynamic range.
5.3 Test set-up

Although, the neural network chip includes a weight refreshing mechanism and weight storage capacitors, the architecture is not completely autonomous in the sense that it does not incorporate an on-chip learning mechanism and memory elements to store the digital value of the weights. To the former issue few solutions are available. First, the neural networks can be trained off-chip. This technique involves evaluating a set of weights that will match the required mapping function utilising a digital neurocomputer simulator and downloading those connection strengths onto the chip. Unless the simulator takes into account the non-idealities of the analogue hardware which are offsets and gain errors, the analogue neural systems may not display the desired mapping function. Furthermore the possibility of characterising each processing unit of a large analogue structure is not always made available since it impairs on hardware cost (i.e. extra connection pads). An alternative solution consists of inserting the neural network in a loop whereby the learning procedure is monitored by an external host computer. This technique is commonly referred to as chip-in-loop training. Since the forward path is provided by the analogue ANN rather than a digital ideal system, the learning mechanism is therefore able to compensate for the distortions. The latter method emerges as the most suited to our present design.

To facilitate communication between the host computer and the analogue neural networks an interface board has been purpose designed. Fig. 5.17 shows the block diagram of the complete test set-up. The test bench includes a high speed digital weight memory circuit which incorporates a $128 \times 8$ bits static RAM to store the digitised value of the synaptic connections. The external storage scheme was designed so that the weights could be refreshed either using data stored in the static RAM or by the host.
computer. In the former case, i.e. when the networks are trained, the weights were sequentially updated at a rate of 80 times a second. This is achieved using a 7 bits serial counter driven by a 10 kHz clock signal. To avoid synchronisation problem during a weight refresh cycle, the RAM is enabled before the on-chip decoders and then disabled after. Twelve 8-bit high-speed DACs are utilised to convert the input pattern data, which are supplied by the host computer, into differential currents. The inputs of the DACs are multiplexed and latched since the host computer provides a limited number of digital input/output data channels. This design burden has the unfortunate consequence of limiting the speed of inter-communication and therefore increasing the training time. The differential output currents of the networks are converted into single-ended voltages utilising sensing schemes similar to that exploited to characterise the ON. These voltages are then transformed, also in a multiplexed manner, into digital words of 12 bits using a single flash Analogue-to-Digital Converter (ADC).

The whole test scheme can be operated either in training or generalisation mode. In the former state, since the retention time of the capacitive storage schemes is 5 seconds, a complete WP training cycle can be executed without having to refresh the weights. Hence analogue noise generated either by the effect of charge injection or capacitive coupling is disregarded. The analogue system can therefore fully enjoy the 8 bits weight resolution during training. When a network is trained its final connection strengths are stored in a file for possible later use. In the generalisation mode, the weights are downloaded onto the RAM and test patterns are fed to the system by the host computer.
5.4 Performance of the analogue neural networks

5.4.1 Exclusive OR network

The XOR network was trained using the weight perturbation algorithm, described above. To avoid saturation of the system during the early phase of the training, the synaptic weights were initially set to zero.

Training was performed using a stepwise approach, where one of the four patterns was presented at the inputs and all weights were updated prior to a pattern change. It may also be added that the patterns were presented to the network in an incremental fashion rather than in a random order. The neural network trained optimally when parameters $\eta$ and $\text{pert}$ were set to 0.05 and 0.2, respectively. Tests showed that if the value of the learning rate was greater than that suggested above, the system tended to settle to a local minimum rather than converging to a global solution. On the other hand, reducing $\eta$ below 0.05 had the effect of increasing the number of epochs required for the network to converge. In an ideal system, the accuracy of the weight perturbation algorithm improves as the perturbation is reduced. In practice however, thermal noise imposes lower limits on the minimum perturbation size and it is necessary to trade-off gradient accuracy for signal-to-noise ratio. Our tests established that a perturbation of 0.2 offers a good compromise.

A batch mode approach was then used to train the network whereby the weights were adjusted following a complete cycle through the training patterns. Results indicated that over its stepwise counterpart, batch mode training allowed larger values of the learning parameter before the system encountered any convergence difficulties. This approach resulted in a substantial improvement in the learning speed. According to
the experiments, as long as $\eta$ was smaller than 0.5 the WP training algorithm offered satisfactory generalisation levels.

Fig. 5.18 shows the learning curve of the XOR network for both stepwise and batch mode training. The training procedures were terminated when changes in the Root Mean Square (RMS) error were so minor as to be undetectable by the learning system. Notwithstanding substantial noise constraints, the network was able to reach a 3% RMS error. Fig. 5.19 illustrates the performances of the XOR network when trained using stepwise mode. As mentioned earlier, for training purposes, the differential output current was converted into a single-ended voltage using a circuit arrangement similar to the one suggested for the ON tests. With the XOR network driving such a load, as shown in Fig. 5.20, the propagation delay was typically 7 $\mu$s of which 1.5 $\mu$s is due to the limitations of the op amps utilised within the sensing arrangement. Thus the propagation time of the forward path is approximately 5.5 $\mu$s. This experimental propagation delay is slightly larger than that predicted by the simulation. This increase may principally be due to the fact that in the simulation study the wiring and coupling capacitances were not taken into account.

When the weight point, which satisfied the mapping function of the XOR network located in the first prototype chip, was downloaded onto the other 9 prototype structures, measurements indicated that in all cases, performances did not reach an acceptable level of operation. This was clearly due to the effect of random circuit offsets. However, results improved to a level comparable to the one shown in Fig. 5.19 after readjustment of the weights and clearly indicate that errors due to offsets and non-linearity can be accommodated by the learning process.
5.4.2 QRS complex detector

The 10:6:3 network was trained to detect QRS inputs using the same learning mode and set of input/output patterns as those employed for the digital equivalent network. The training program employed a weight perturbation procedure, similar to that used during computer simulation, with additional functions dedicated to communication between the host and the network inputs/outputs and the weight storage mechanism. It may also be mentioned that the weight perturbation procedure provided updated digital weights which had been rounded up to their nearest feasible value in order to virtually increase the weight resolution by $\frac{1}{2}$ bit. It was initially found that with all the synaptic weights set to a null value, the outputs of the QRS complex detector network were fully saturated. The solution adopted involved training the network to eliminate the saturated nodes. The weight point was subsequently adopted as an initial training set and the QRS complex detector successfully trained using 0.05 and 0.2 as learning rate and perturbation, respectively. Fig. 5.21 shows the learning curve of the network. It may be noted that the RMS error could not be reduced below 7%; a value comparatively greater than that of its analogous software simulator. This result is unsurprising given the limited weight resolution and the analogue thermal noise. Although the forward propagation time for the analogue network is substantially lower than that of its digital counterpart, the overall training time was, however, longer. The increased training time was mainly due to slow communication between the host and the test bench.

The effectiveness of both the digital and analogue simulators in detecting the QRS was assessed using test samples. The foetal ECG data was obtained from the Plymouth Perinatal Research Group. The ECG signals were digitised at 500 samples/s to a resolution of 8 bits. To minimise the influence of external noise source, such as mains,
during signal acquisition, it has been necessary to pre-process the raw ECGs. The pre-processing was achieved using an optimal linear phase Finite Impulse Response (FIR) band-pass digital filter, whose frequency response is shown in Fig. 5.22. The filter attenuates the mains frequency and successive harmonics components by approximately 40 dB. In the case of cardiac arrhythmia classification [79], it may be noted that pre-processing might not be needed since the system would be implanted near the information source, i.e. heart. The ECG was then normalised to lie in the range ±1 and serially fed to the inputs of the networks using time delays. Since the raw ECG was sampled 500 times a second, one in three samples were supplied to the systems so that a complete QRS complex fitted the window formed by the time delays. The generalisation level of the networks was determined using the following measure of performance

\[
P_m = \frac{(\text{Total } N^o\text{ of QRS complexes} - \text{N}^o\text{ of misses} - \text{N}^o\text{ of false detections})}{\text{Total } N^o\text{ of QRS complexes}}
\]  

(5.10)

This performance measure attains a value of 100% only if all the QRS complexes in an ECG are correctly detected, i.e. no misses and no false detections. The number of misses or false detections were determined by visually comparing the outputs of the neural networks and the filtered ECG. A QRS was assumed to be present if the output of the neural network exceeded a threshold level set to half of the maximum output neuron swing, i.e. zero.

The ability of the network to detect previously unseen sets of QRS complexes was tested on two foetal ECG taken from different foetuses. The first ECG signal contained 800 heart beats and the quality of the data could be described as medium, whilst the second was composed of 465 QRS complexes of poor quality. Table 5.7 shows the performances of both digital and analogue neural networks. The results suggest that a
digital simulator performs marginally better than an ANN designed using MOS devices operating in subthreshold mode on the problem of detecting QRS complexes from foetal ECG. The reasons are that the analogue system offers a limited weight resolution and thermal noise imposed a restriction on the minimum value of the perturbation parameter. Fig. 5.23 shows an example of QRS detection for a medium quality raw ECG signal. Note the constant delay between the raw and filtered ECG, which is characterised by the length of the FIR filter (i.e. 200).

5.5 Summary

The experimental characteristics of each building block used in the design of fully connected feedforward neural networks have been presented in this chapter. The measurements indicated that for all arrangements except the capacitive weight storage scheme, the performances were in close agreement with the theoretical and simulation predictions. The concept of the learning process has also been dealt with in this chapter. Although the back propagation learning algorithm is well suited to train MLP which is simulated by utilising digital neurocomputer, its application to analogue simulators incurs considerable hardware cost. While the WP learning rule allows a trade-off of learning speed against hardware complexity, it offers a more adequate solution to training utilising a chip-in-loop architecture. Computer simulation studies revealed that, based on the task of detecting QRS complexes in a foetal ECG signal, the WP reached excellent levels of convergence. To implement the WP learning algorithm a specially designed chip-in-loop system has been developed. Finally, both XOR and QRS detector analogue neural networks were successfully trained. These results suggest that in spite of offsets, gain errors and analogue noise feedforward ANNs may be efficiently
simulated exploiting the exponential relation of current and voltage offered by MOS transistors biased in their weak inversion.
Figure 5.1: Experimental set-up for measuring the transfer characteristics of the TTS
Figure 5.2: Measured static transfer curves of the TTS. (a) $V_{TS}$ against $V_{SP}$, (b) $V_{TS}$ against $V_{W}$. 
Figure 5.3: Test configuration used for measuring the characteristics of the load.

Figure 5.4: Experimental static characteristics of the I/V converter. $V_R$ against $V_C$ for $I_C / V_C = 1\mu A/V$. 
Figure 5.5: Experimental set-up for measuring the characteristics of the horizontal resistance.

Figure 5.6: Experimental static characteristic of the horizontal resistor. 
$V_H$ against $V_C$, $I_{Si} = V_C / 2 \cdot 10^3$. 
Figure 5.7: Experimental static characteristics of the transconductance-thresholding-synapse when fed by the horizontal resistance. $V_{TS}$ against $V_C$. 
Figure 5.8: Measured static transfer characteristics of the output neuron. (a) $V_O$ against $V_{Si}$ and (b) $V_O$ against $V_W$. 
Figure 5.9: Transient characteristics of the output neuron when loaded by its associated sensing scheme.
Figure 5.10: Variation of the TTS differential output current during the charge leakage process.
Figure 5.11: Effect of clock feedthrough and capacitive coupling onto the output of the TTS during weight refreshing.
Figure 5.12: Transient characteristic of the refreshing mechanism.
Figure 5.13: Signal flow of the error back-propagation algorithm.
Figure 5.14: Signal flow of the weight perturbation algorithm.
Figure 5.15: ECG training patterns.

Figure 5.16: Learning curve of the computer simulated QRS complex detector. Learning rate: $\eta = 0.05$ and perturbation: $pert = 0.08$. 
Figure 5.17: Experimental set-up for training and testing the analogue ANNs.
Figure 5.18: Learning curve of the analogue XOR neural network using the weight perturbation algorithm.
Figure 5.19: Performances of the XOR network when trained using stepwise mode.

Figure 5.20: Transient characteristics of the XOR network.
Figure 5.21: Learning curve of the analogue QRS detector network using the weight perturbation algorithm. Learning rate: $\eta = 0.05$ and perturbation: $pert = 0.2$. 

![Graph showing the learning curve of the analogue QRS detector network. The x-axis represents the number of epochs, and the y-axis represents the RMS error (%). The graph shows a decreasing trend in RMS error as the number of epochs increases.]
Figure 5.22: Frequency response of the linear phase FIR filter.

Figure 5.23: Example of QRS detection for medium quality ECG.
<table>
<thead>
<tr>
<th>Prototype Chip</th>
<th>Offset Current (nA)</th>
<th>Transconductance Thresholding Synapse 1</th>
<th>Transconductance Thresholding Synapse 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Damage during test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>29</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: Output current offset of the individual TTSs.
### Table 5.2: Nominal resistance of the I/V converter for various bias currents.

<table>
<thead>
<tr>
<th>Biasing current (nA)</th>
<th>15</th>
<th>30</th>
<th>60</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal resistance (MΩ)</td>
<td>30</td>
<td>15</td>
<td>8</td>
<td>2.5</td>
</tr>
</tbody>
</table>

### Table 5.3: Output voltage offset of the individual I/V converters.

<table>
<thead>
<tr>
<th>Prototype chip</th>
<th>Offset voltage (mV)</th>
<th>Current-to-Voltage converter 1</th>
<th>Current-to-Voltage converter 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>11.8</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.9</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7.4</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>12.4</td>
<td>14.4</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.6</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.4</td>
<td>7.2</td>
<td></td>
</tr>
</tbody>
</table>

216
<table>
<thead>
<tr>
<th>Prototype Chip</th>
<th>Offset voltage (mV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13.2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>15.8</td>
</tr>
<tr>
<td>4</td>
<td>7.2</td>
</tr>
<tr>
<td>5</td>
<td>11.4</td>
</tr>
<tr>
<td>6</td>
<td>7.9</td>
</tr>
<tr>
<td>7</td>
<td>13.7</td>
</tr>
<tr>
<td>8</td>
<td>15.2</td>
</tr>
<tr>
<td>9</td>
<td>14.1</td>
</tr>
<tr>
<td>10</td>
<td>6.4</td>
</tr>
</tbody>
</table>

Table 5.4: Output voltage offset of the individual HRes.

<table>
<thead>
<tr>
<th>Prototype Chip</th>
<th>Offset Current (nA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>347</td>
</tr>
<tr>
<td>3</td>
<td>69</td>
</tr>
<tr>
<td>4</td>
<td>165</td>
</tr>
<tr>
<td>5</td>
<td>121</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>63</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 5.5: Output current offset of the individual ONs.
### Table 5.6: DACs bit current sources.

<table>
<thead>
<tr>
<th>Prototype Chip</th>
<th>Bit Current (nA)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B_0$</td>
<td>$B_1$</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>41</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>41</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>41</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>41</td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td>21</td>
<td>41</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>41</td>
</tr>
</tbody>
</table>

### Table 5.7: Performances of the analogue and digitally simulated QRS complex detectors.

<table>
<thead>
<tr>
<th>Network Type</th>
<th>Number of QRS complexes</th>
<th>Number of missed beats</th>
<th>Number of false detections</th>
<th>Performances of NN (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analogue simulator</td>
<td>800</td>
<td>1</td>
<td>20</td>
<td>97.4</td>
</tr>
<tr>
<td>Digital simulator</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analogue simulator</td>
<td>465</td>
<td>20</td>
<td>2</td>
<td>95.2</td>
</tr>
<tr>
<td>Digital simulator</td>
<td></td>
<td>8</td>
<td>2</td>
<td>97.8</td>
</tr>
</tbody>
</table>

Table 5.7: Performances of the analogue and digitally simulated QRS complex detectors.
Chapter 6

Conclusions

6.1 Discussion of results

A succinct description of the evolution of artificial neural systems and their categorisation has been presented. After a detailed explanation of the concept and terminology of AN models and ANN topologies, it has been established that artificial neuronal functions may be simulated either as software on a digital computer or in hardware using either digital, analogue or hybrid VLSI technologies. It emerged that although software-based simulators offer flexibility in terms of computing resolution and topology alteration they lack processing speed and fault tolerance which are characteristics required by most real time applications. To acquire these features a systolic design approach is required and appears to be best achieved when using either analogue or hybrid VLSI techniques since they allow for compact processing units.

The adequacy of various previously reported analogue and hybrid circuits to perform as either synapses or neurons was explored. Studies revealed that an MOS transistor operated as a voltage controllable transconductor biased in the weak inversion is a resourceful processing unit with which to build primitive ANN functions since it allows
for compact structures which dissipate power levels compatible with implantable neuronal systems. However, because of the exponential transfer characteristic, the computational schemes either display narrow linear range or are highly sensitive to biasing conditions which limit the computing resolution to a level below that required for the training of feedforward ANNs. A survey also revealed that existing memory schemes such as the capacitor, the floating-gate transistor and the binary weighted array of devices, suffer from substantial drawbacks.

To fully enjoy the transfer characteristic of multipliers operating in the weak inversion, distribution of the thresholding operation of the hidden neurons of a fully connected feedforward neural network over to the inputs of the synapses they are feeding to form TS blocks, has been suggested. Although the modified Gilbert scheme can fulfil the role of a TS, the approach was compromised because of the lack of an adequate current memory cell. To overcome this problem a TTS based on a novel four-quadrant multiplier was introduced. The scheme is compact since it utilises a minimum number of devices. Experimental results correlated well with the analytical and simulation predictions and revealed that it can be supplied with voltages as low as 3 V, has a linear differential input as wide as ±250 mV while the other input displays the required tanh function and that process parameter variations have a limited impact.

It has also been shown that, although the required characteristic of the neuron is that of a linear load, moderate levels of non-linearity in the load resistor characteristic have a negligible effect on overall performance of MLPs with the exception that it considerably enhances the processing speed. To acquire this feature, a novel diode-based load arrangement has also been developed. Analytical and simulation predictions as well as experimental measurements indicated that immunity to power supply variations of
diode-based resistive elements can be substantially increased at the tolerable expense of power dissipation.

It has also been established that combining the novel TTS and load schemes to form the core of a feedforward ANN, the weight range of the synapses can readily be adjusted and power dissipation traded off against processing speed independently while the input and output nodes of the system are thresholded and un-thresholded respectively. An output neuron was developed as a means of squashing the output signals and horizontal resistances were used to linearise the inputs. Furthermore, since the input and output of the network were made to be compatible the proposed design could also be used to simulate recurrent ANNs.

Simulation and analytical predictions suggested that to obtain the required minimum weight resolution of 8 bits, the effect of clock feedthrough on a basic PMOS-based sample-and-hold circuit as a storage element had to be compensated for by utilising a dummy transistor. However this approach has the side effect of decreasing the retention time of the memory since the number of leakage paths is increased. A twin capacitor structure was then proposed to alleviate this problem and correspondingly attenuate refreshing speed. However experimental results revealed that such a storage scheme is vulnerable to either charge injection or capacitive coupling phenomenon.

To assess experimentally the viability of the proposed design, a neural network chip including the XOR network as a benchmark problem and a QRS complex detector based on a 10:6:3 MLP with an on-chip weight update/refresh scheme was designed and fabricated. A special purpose interface board was also built in order to train the networks utilising a chip-in-loop architecture. It has been suggested that since the retention time of the capacitive storage schemes is of the order of 5 seconds, during
training refreshment of the connection strengths of the networks is best achieved between epochs in order to avoid switched induced interference which attenuate the weight resolution below the required 8 bits. This approach facilitated successful training of the networks utilising the weight perturbation learning rule. In spite of noise constraints, the analogue networks were able to reach generalisation levels similar to those obtained from a digital simulator. Experimental results have also indicated that distortion due to device mismatching does not interfere in the learning performance of the networks and that the processing speed of the XOR marginally corresponds to that predicted by the simulations.

To conclude, the results of this study revealed that the transfer characteristic of an MOS transistor biased in the weak inversion can be fully exploited (i.e. gate and bulk terminals are driven) to design compact and low power neuronal functions which alleviate issues such as limited linearity range, restricted weight resolution, processing speed and mismatching effect.

6.2 Recommendations for future work

A review of the work reveals four possible refinements which could be investigated in further research.

6.2.1 Weight storage

As mentioned in chapter 5, the effects of either charge injection or capacitive coupling onto the weight signals are much larger than anticipated. Whilst the capacitive coupling can (in hindsight) be significantly decreased by re-routing analogue and digital signal lines the evaluation and reduction in switch-charge injection may be the focus of future
activities. A solution to these issues may lie in the use of FGMOS transistors. Although, as specified in chapter 2, this technology suffers from the fact that the threshold voltage $V_T$ varies linearly with the logarithm of the number of pulses and that the write and erase operations are asymmetrical; a diode-connected FGMOS transistor acting as a variable resistor biased by a constant current source, as shown in Fig. 6.1, may be used to replaced the storage capacitors and thus substantially improve area and power efficiency.

6.2.2 On-chip learning

An on-chip learning mechanism may be added to overcome the two aforementioned problems of FGMOS devices and also provide for a self adaptable neuronal system that could be used as an implantable functional electrical stimulator for people suffering from paralysed limbs, such as those with spinal cord injuries, and consequently facilitate interaction between the network and the changing condition of the patients due to fatigue.

6.2.3 Learning algorithm

In the previous chapter, it has been mentioned that although the weight perturbation learning algorithm is suitable to analogue on-chip implementation, this is partly negated by the fact that learning is slower than the error back-propagation. It would be of interest to investigate optimisation methods based on simple search techniques such as the feedforward method introduced by Petridis and Paraschidis [103] since, firstly, their extensive simulation results suggested that their method usually converges faster than the back-propagation and, secondly, it facilitates analogue hardware implementation even more than the weight perturbation would.
6.2.4 Sensitivity to temperature

As stipulated in the introduction, the transconductance characteristic of a weakly inverted MOS device is extremely sensitive to temperature drifts. Although for implantable neuronal systems the human body offers a stable thermal environment, it would also be of interest to evaluate the sensitivity of the suggested system to temperature variations.

6.3 Summary

In summary, this thesis presents four novel circuits, namely a four-quadrant multiplier, a current-to-voltage converter, an output neuron and an on-chip weight update/refresh mechanism, designed to simulate low energy feedforward neural networks suitable for battery-powered implantable neuronal systems. These have been analytically studied, simulated using the CAD tool SPICE, fabricated in a low-cost 2.4 μm double poly, double metal CMOS process and their operations experimentally confirmed. A 2:2:1 feedforward network and a 10:6:3 MLP were designed and successfully trained, to solve the XOR benchmark problem and to detect QRS complexes respectively, using the weight perturbation learning algorithm.
Figure 6.1: Transconductance-thresholding-synapse incorporating an FGMOS-based memory device.
Bibliography


implementation of a neural network memory with several hundreds of neurons',


1989.

[23] MAHER, M. A. C.; DEWEERTH, S. P.; MAHOWALD, M. A. and MEAD, C.
A.: 'Implementing Neural architectures using analog VLSI circuits', IEEE


representation by an analog silicon model of the auditory periphery', IEEE


[27] LAN, N.; FENG, H. and CRAGO, P. E.: 'Neural network generation of muscle
stimulation patterns for control of arm movements', IEEE Transactions on

[28] GRAUPE, D. and KORDYLEWSKI, H.: 'Artificial neural network control of
FES in paraplegics for patient responsive ambulation', IEEE Transactions on


235


Appendix A

The MOS transistor

A brief analytical description of the operation of Complementary Metal Oxide Semiconductors (CMOS) is presented in this Appendix. The reader may consult references [22], [61], [72-74] for more detailed discussions. The basic structure of a four terminal n-channel MOS (NMOS) transistor is shown in Fig. A.1 (a) and its symbolic representation in Fig. A.1 (b). The terminals are known as the Gate (G), the Drain (D), the Source (S) and the Bulk (B) which is also referred to as the substrate or back-gate. The current that flows between the drain and source nodes of the device depends upon the density of channel charge carriers which is controlled by the level of gate and source potentials. This current is commonly referred to as the drain current $I_D$.

The NMOS transistor models presented in the subsequent sections are also applicable to the p-channel MOS (PMOS) device by multiplying all currents and voltages by -1.

A.1 Strong inversion

For gate-to-source voltage ($V_{GS}$) greater than the threshold voltage ($V_T$) the concentration of electron carriers in the channel is high. In this mode of conduction, the channel also
known as the inversion layer is said to be strongly inverted and the drain current is mainly due to drift [74] and may be expressed as

\[
I_D = K_N \cdot \{(V_{GS} - V_{FB} - \phi_B)\sqrt{V_{DS}} - \frac{1}{2} V_{DS}^2 \}
\]

\[
- \frac{2}{3} \gamma \left[ (V_{DS} - V_{BS} + \phi_B)^{\frac{3}{2}} - (\phi_B - V_{BS})^{\frac{3}{2}} \right] \] (A.1)

where the parameters are defined in Table A.1. For analytical development this expression may appear too complicated and therefore needs to be simplified. This may be achieved utilising a binomial expansion technique. The Taylor’s series of \((1 + x)^n\) is given by

\[
(1 + x)^n = 1 + nx + \frac{n(n - 1)}{2!}x^2 + \frac{n(n - 1)(n - 2)}{3!}x^3 + \ldots \quad \text{for } x^2 < 1 \] (A.2)

If one considers the first \(3/2\) term in (A.1),

\[
(\phi_B - V_{BS} + V_{DS})^{\frac{3}{2}} = (\xi + V_{DS})^{\frac{3}{2}} = \xi^{\frac{3}{2}} \left( 1 + \frac{V_{DS}}{\xi} \right)^{\frac{3}{2}} \] (A.3)

where

\[
\xi = \phi_B - V_{BS} \] (A.4)

Utilising (A.2) to expand (A.3) gives

\[
(\phi_B - V_{BS} + V_{DS})^{\frac{3}{2}} = (\phi_B - V_{BS})^{\frac{3}{2}} + \frac{3}{2} V_{DS} \sqrt{\phi_B - V_{BS}} + \frac{3}{8} \frac{V_{DS}^2}{\sqrt{\phi_B - V_{BS}}} - \frac{1}{16} \frac{V_{DS}^3}{(\phi_B - V_{BS})^{\frac{3}{2}}} + \frac{3}{128} \frac{V_{DS}^4}{(\phi_B - V_{BS})^{\frac{5}{2}}} - \ldots \] (A.5)

for \(V_{DS} < (\phi_B - V_{BS})\). Thus substituting (A.5) into (A.1), the drain current may be approximated to
\[
I_D = K_N \cdot \left( V_{GS} - V_{FB} - \phi_B - \gamma \sqrt{\phi_B - V_{BS}} \right) \cdot V_{DS} - \frac{1}{2} \left( 1 + \frac{\gamma}{2 \cdot \sqrt{\phi_B - V_{BS}}} \right) V_{DS}^2
\]

\[
+ \frac{\gamma V_{DS}^3}{24.(\phi_B - V_{BS})} - \frac{\gamma V_{DS}^4}{64.(\phi_B - V_{BS})^2} \}
\]

(A.6)

It can be seen that the coefficient associated to the cubic and subsequently higher terms in $V_{ds}$ are comparatively much smaller than that of the square term. Hence the drain current may be further approximated to

\[
I_D = K_N \cdot \left[ (V_{GS} - V_T) \cdot V_{DS} - \frac{\delta V_{DS}^2}{2} \right]
\]

(A.7)

where

\[
V_T = V_{TO} + \gamma \left( \sqrt{\phi_B - V_{BS}} - \sqrt{\phi_B} \right)
\]

(A.8)

is the threshold voltage and

\[
\delta = 1 + \frac{\gamma}{2.\sqrt{\phi_B - V_{BS}}}
\]

(A.9)

$V_{TO}$ is the zero bias threshold voltage given by

\[
V_{TO} = V_{FB} + \phi_B + \gamma \sqrt{\phi_B}
\]

(A.10)

Note that when both the substrate is lightly doped and the gate oxide is thin, $\gamma$ is small and the drain current (A.7) simplifies to the well known form

\[
I_D = K_N \cdot \left[ (V_{GS} - V_T) \cdot V_{DS} - \frac{1}{2} V_{DS}^2 \right]
\]

(A.11)

Operated in strong inversion (also called above threshold) the NMOS transistor therefore displays an ohmic characteristic as long as the drain-to-source voltage do not reach the saturation level which occurs when $d I_D / d V_{DS} = 0$ and is given by
\[ V_{DS} = V_{GS} - V_T \]  \hspace{1cm} (A.12)

If so, substituting (A.12) into (A.11), an NMOS device biased in its saturation region of strong inversion is modelled by

\[ I_D = \frac{K_N}{2} (V_{GS} - V_T)^2 \]  \hspace{1cm} (A.13)

**A.2 Weak inversion**

When the density of mobile charge carriers is lower than depletion charge, the channel is weakly inverted. This condition of operation arises for gate-to-source biasing voltages lower than the threshold voltage. This is referred to as the subthreshold mode of conduction or weak inversion. The flow of electrons is caused by diffusion [22], [61], [72-74] which materialises as

\[ I_D = \frac{W}{L} \mu_N C_{ox} V_t^2 \gamma \frac{\sqrt{2} \sqrt{\phi_B - 5.5 V_t - V_{BS}}} {\exp\left(-2 \frac{\phi_B}{V_t}\right) \exp\left(\frac{\kappa}{V_t} \frac{V_{GS}}{V_t}\right)} \exp\left(1 - \frac{\kappa}{V_t} \frac{V_{BS}}{V_t}\right) \left[ 1 - \exp\left(\frac{V_{DS}}{V_t}\right) + \frac{V_{DS}}{V_O} \right] \]  \hspace{1cm} (A.14)

where \( \kappa \) and \( V_O \) are defined in table A.1. This expression maybe rewritten in a simpler form as

\[ I_D = \frac{W}{L} I_{DO} \exp\left(\kappa \frac{V_{GS}}{V_t}\right) \exp\left(1 - \kappa \frac{V_{BS}}{V_t}\right) \left[ 1 - \exp\left(\frac{V_{DS}}{V_t}\right) + \frac{V_{DS}}{V_O} \right] \]  \hspace{1cm} (A.15)

where \( I_{DO} \) is the characteristic current and is given by

\[ I_{DO} = \mu_N C_{ox} V_t^2 \gamma \frac{\sqrt{2} \sqrt{\phi_B - 5.5 V_t - V_{BS}}} {\exp\left(-2 \frac{\phi_B}{V_t}\right)} \]  \hspace{1cm} (A.16)

One can see that for drain-to-source quiescent potential greater than 4.\( V_t \) but remaining low enough to disregard the Early effect, (A.15) reduces to
\[ I_D = \frac{W}{L} I_{DO} \exp\left( \kappa \frac{V_{GS}}{V_t} \right) \exp\left( [1-\kappa] \frac{V_{BS}}{V_t} \right) \]  \hfill (A.17)

When the substrate potential with reference to that of the source is null, the above expression further simplifies to the well known form

\[ I_D = \frac{W}{L} I_{DO} \exp\left( \kappa \frac{V_{GS}}{V_t} \right) \]  \hfill (A.18)
Figure A.1: An n-channel MOS transistor. (a) structure, (b) symbol.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_N = \mu_N C_{ox} \frac{W}{L}$</td>
<td>$\mu A/V^2$</td>
<td>Transconductance parameter</td>
</tr>
<tr>
<td>$\mu_N$</td>
<td>cm$^2$/Vs</td>
<td>Electron mobility</td>
</tr>
<tr>
<td>$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$</td>
<td>F/m$^2$</td>
<td>Gate oxide capacitance per unit area</td>
</tr>
<tr>
<td>$\varepsilon_{ox} = k_{ox} \varepsilon_0$</td>
<td>F/m</td>
<td>Permittivity of silicon dioxide</td>
</tr>
<tr>
<td>$\varepsilon_0 = 8.854 \times 10^{-12}$</td>
<td>F/m</td>
<td>Permittivity of free space</td>
</tr>
<tr>
<td>$k_{ox} = 3.9$</td>
<td></td>
<td>Dielectric constant of the insulator</td>
</tr>
<tr>
<td>$t_{ox}$</td>
<td>m</td>
<td>Thin gate oxide thickness</td>
</tr>
<tr>
<td>$W$</td>
<td>m</td>
<td>Drawn channel width</td>
</tr>
<tr>
<td>$L$</td>
<td>m</td>
<td>Drawn channel length</td>
</tr>
<tr>
<td>$V_{GS}$</td>
<td>V</td>
<td>Gate to source voltage</td>
</tr>
<tr>
<td>$V_{DS}$</td>
<td>V</td>
<td>Drain to source voltage</td>
</tr>
<tr>
<td>$V_{DS}$</td>
<td>V</td>
<td>Bulk to source voltage</td>
</tr>
<tr>
<td>$V_{FB}$</td>
<td>V</td>
<td>Flat band voltage</td>
</tr>
<tr>
<td>$\phi_B = 2. V_i \ln \frac{N_{SUB}}{n_i}$</td>
<td>V</td>
<td>Surface potential</td>
</tr>
<tr>
<td>$V_i = \frac{k.T}{q}$</td>
<td>V</td>
<td>Thermal voltage</td>
</tr>
<tr>
<td>$k = 1.38 \times 10^{-23}$</td>
<td>J/K</td>
<td>Boltzmann's constant</td>
</tr>
<tr>
<td>$T$</td>
<td>K</td>
<td>Temperature</td>
</tr>
<tr>
<td>$q = 1.6 \times 10^{-19}$</td>
<td>C</td>
<td>Electron charge</td>
</tr>
<tr>
<td>$N_{SUB}$</td>
<td>cm$^3$</td>
<td>Substrate doping</td>
</tr>
<tr>
<td>$n_i = 1.45 \times 10^{10}$</td>
<td>cm$^3$</td>
<td>Intrinsic carrier concentration</td>
</tr>
<tr>
<td>Symbol</td>
<td>Unit</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>$\gamma = \frac{(2.04 \times 10^{10})^{1/2}}{C_{ox}}$</td>
<td>V$^6$</td>
<td>Bulk threshold parameter</td>
</tr>
<tr>
<td>$\varepsilon_{si} = 1.04 \times 10^{10}$</td>
<td>F/m</td>
<td>Permittivity of silicon</td>
</tr>
<tr>
<td>$V_T = V_{TO} + \gamma \left( \phi_B - V_{BS} \right)$</td>
<td>V</td>
<td>Threshold voltage</td>
</tr>
<tr>
<td>$V_{TO} = V_{FB} + \phi_B + \gamma \sqrt{\phi_B}$</td>
<td>V</td>
<td>Zero bias threshold voltage</td>
</tr>
<tr>
<td>$\kappa = \frac{1}{1 + \frac{\gamma}{2. \sqrt{\phi_B - 5.V_t - V_{BS}}}}$</td>
<td></td>
<td>Effectiveness of the gate in Controlling the channel current</td>
</tr>
<tr>
<td>$V_0$</td>
<td>V</td>
<td>Early voltage</td>
</tr>
</tbody>
</table>

Table A.1: List of Symbols.
Appendix B

CMOS parameters for MIETEC 2.4µm double poly, double metal process
<table>
<thead>
<tr>
<th>Parameter</th>
<th>NMOS</th>
<th>PMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTO, V</td>
<td>0.86</td>
<td>-0.85</td>
</tr>
<tr>
<td>TOX, $\times 10^9$ m</td>
<td>40.29</td>
<td>42.46</td>
</tr>
<tr>
<td>NSUB, $\times 10^{15}$ cm$^3$</td>
<td>1.39</td>
<td>9.1</td>
</tr>
<tr>
<td>XJ, $\times 10^6$ m</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>LD, $\times 10^6$ m</td>
<td>0.22</td>
<td>0.35</td>
</tr>
<tr>
<td>UO, $\times$ cm$^2$/Vs</td>
<td>611.37</td>
<td>233.84</td>
</tr>
<tr>
<td>VMAX, $\times 10^3$ m/s</td>
<td>158</td>
<td>225.67</td>
</tr>
<tr>
<td>DELTA</td>
<td>0.85</td>
<td>0.96</td>
</tr>
<tr>
<td>THETA, V$^{-1}$</td>
<td>0.05</td>
<td>0.12</td>
</tr>
<tr>
<td>ETA</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>KAPPA</td>
<td>1.4</td>
<td>9.23</td>
</tr>
<tr>
<td>GAMMA, V$^{\kappa}$</td>
<td>0.26</td>
<td>0.69</td>
</tr>
<tr>
<td>NFS, $\times 10^{11}$ cm$^2$</td>
<td>1.35</td>
<td>3.92</td>
</tr>
<tr>
<td>CGBO, $\times 10^{-10}$ F/m</td>
<td>5.57</td>
<td>5.57</td>
</tr>
<tr>
<td>PB, V</td>
<td>0.65</td>
<td>0.76</td>
</tr>
<tr>
<td>CJ, $10^4$ F/m</td>
<td>0.69</td>
<td>3.1</td>
</tr>
<tr>
<td>MJ</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>CJSW, $\times 10^{-10}$ F/m</td>
<td>3.43</td>
<td>3.67</td>
</tr>
<tr>
<td>MJSW</td>
<td>0.27</td>
<td>0.38</td>
</tr>
<tr>
<td>PHI, V</td>
<td>0.62</td>
<td>0.67</td>
</tr>
<tr>
<td>RSH, $\Omega$</td>
<td>33.42</td>
<td>35.01</td>
</tr>
<tr>
<td>KP, $\times 10^6$ A/V$^2$</td>
<td>51.71</td>
<td>19.14</td>
</tr>
<tr>
<td>JS, A</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>
Appendix C

Analysis of a diode-based current-to-voltage converter

The objective of this Appendix is to derive the transfer characteristic of a diode-based current-to-voltage converter. The notation and configuration refer to Fig. 2.23. The analysis is based on the assumption that the operating conditions are such that the current flowing through an MOS device biased in weak inversion, as in Appendix A, is exponentially related to its gate-source potential and that all transistors have matched characteristics. For theoretical purposes the characteristic current and the effectiveness of the gate in controlling the channel current parameters of n- and p-channel devices are assumed to be identical. However this assumption does not necessarily hold true in practice. It is shown in the next Appendix that this limitation may be avoided by using a core of single type transistors and a feedback mechanism.

Consider the resistive branch made up of transistor $M_1$ and $M_3$, based on the Kirchhoff's current principle, the input current is related to

$$I_s^* = -I_{M1} - I_{M3} \quad \text{(C.1)}$$
where $I_{m_1}$ and $I_{m_3}$ are the drain currents of $M_1$ and $M_3$, respectively. Based on (A.18), these drain currents can be expressed as

$$I_{m_1} = I_N \exp \left( \kappa \frac{V_{GS_1}}{V_t} \right) = I_N \exp \left( \kappa \frac{V_{dd} - E_1 - V^*_Y}{V_t} \right)$$  \hspace{1cm} (C.2)

and

$$I_{m_3} = -I_N \exp \left( -\kappa \frac{V_{GS_3}}{V_t} \right) = -I_N \exp \left( -\kappa \frac{E_2 - V^*_Y}{V_t} \right)$$  \hspace{1cm} (C.3)

where $V_{GS_1}$ and $V_{GS_3}$ are respectively the gate-to-source potentials of $M_1$ and $M_3$. $V_{dd}$ is the supply voltage and $E_1$ and $E_2$ are the diodes biasing potentials. Thus combining (C.1), (C.2) and (C.3) it can readily be shown that

$$I^* = I_N \exp \left( \kappa \frac{V_{dd} - E_1 + E_2}{2V_t} \right) \exp \left( \kappa \frac{V^*_Y - V_{dd} - E_1 + E_2}{2V_t} \right)$$  \hspace{1cm} (C.4)

which reduces to

$$I^*_N = 2I_N \exp \left( \kappa \frac{V_{dd} - E_1 + E_2}{2V_t} \right) \sinh \left( \kappa \frac{V^*_Y - V_{dd} - E_1 + E_2}{2V_t} \right)$$  \hspace{1cm} (C.5)

Hence

$$V^*_Y = \frac{V_{dd} - E_1 + E_2}{2} + \frac{V_t}{\kappa} \sinh^{-1} \left( \frac{I^*_N}{2I_N} \right)$$  \hspace{1cm} (C.6)

where $I_N$ is the quiescent bias current given by

$$I_N = I_N \exp \left( \kappa \frac{V_{dd} - E_1 + E_2}{2V_t} \right)$$  \hspace{1cm} (C.7)
(C.6) shows that $V^*_y$ is the sum of a quiescent and a dynamic component. Note that the former can be adjusted to $V_{dd}/2$ if $E_1 = E_2$.

Given the symmetry of the scheme, it can be readily derived that

$$V_y = \frac{V_{dd} - E_1 + E_2}{2} + \frac{V_1}{\kappa \cdot \sinh^{-1}\left(\frac{I_s}{2.1\chi}\right)}$$  \hspace{1cm} (C.8)

Since $I^*_s = -I'_s = i_s/2$, then it can be concluded that

$$v_y = V^*_y - V_y = \frac{2V_1}{\kappa \cdot \sinh^{-1}\left(\frac{i_s}{4.1\chi}\right)}$$  \hspace{1cm} (C.9)
Appendix D

Analysis of a diode-based current-to-voltage converter incorporating a feed-back mechanism

The aim of this Appendix is to derive an analytical expression for the transfer characteristic of the diode-based current-to-voltage converter depicted in Fig. 3.8 (a). The analysis is based on the simplified model of a PMOS device operated in the subthreshold mode of conduction (3.22). It has also been assumed that all devices have similar features. The notation and configuration refer to Fig. 3.8 (a).

Since it was assumed that all the transistors in the resistive scheme shown in Fig. 2.23 had identical characteristics, it can therefore be appreciated that the operation of the core diode group of the resistive load incorporating the feedback arrangement is described by (C.7) and (C.9) as

\[ v_o = \frac{2V_t}{K} \sinh^{-1}\left(\frac{i_o}{4.1x}\right) \]  

\[ \text{(D.1)} \]
where

\[ I_x = I_x \cdot \exp \left( \frac{\kappa \cdot V_{dd} - [E_1 + E_2]}{2 \cdot V_t} \right) \]  \hspace{1cm} (D.2)

where \( V_{dd} \) is the supply voltage and \( E_1 \) and \( E_2 \) are respectively the potentials across diodes \( M_5 \) and \( M_6 \). An expression for \( V_{dd} - [E_1 + E_2] \) in (D.2) may be derived as follow.

Consider the shunt device \( M_7 \), its drain current may be expressed as

\[ I_{M7} = -I_x \cdot \exp \left( -\kappa \cdot \frac{V_{GS7}}{V_t} \right) = -I_x \cdot \exp \left( \kappa \cdot \frac{V_{dd} - E_1 - E_2}{V_t} \right) \]  \hspace{1cm} (D.3)

where \( V_{GS7} \) is the gate-to-source potential of \( M_7 \) and \( E_2 \) is the potential across diodes \( M_5 \) and \( M_6 \). It may be noticed that the biasing current of the core diode group is given by

\[ I_c = I_{M7} - I_{M6} \]  \hspace{1cm} (D.4)

where \( I_{M6} \) is the drain current of \( M_6 \) and can be expressed as

\[ I_{M6} = -I_x \cdot \exp \left( -\kappa \cdot \frac{V_{GS6}}{V_t} \right) = -I_x \cdot \exp \left( \kappa \cdot \frac{E_2}{V_t} \right) \]  \hspace{1cm} (D.5)

where \( V_{GS6} \) is the gate-to-source potential of \( M_6 \). Thus combining (D.3), (D.4) and (D.5), it can readily be shown that

\[ V_{dd} - [E_1 + E_2] = E_\alpha - \frac{V_t}{\kappa} \cdot \ln \left( 1 + \frac{I_c}{I_{M7}} \right) \]  \hspace{1cm} (D.6)

Given that the current source \( I_x \) biases both diodes \( M_8 \) and \( M_9 \), the sum of their gate-to-source potentials can be expressed as

\[ E_\alpha = \frac{2 \cdot V_t}{\kappa} \cdot \ln \left( \frac{I_x}{I_{M7}} \right) \]  \hspace{1cm} (D.7)

Then combining (D.1), (D.2), (D.6) and (D.7) it can be derived that the differential potential generated by the resistive load is given by
\[ v_o = \frac{2V_1}{K} \sinh^{-1} \left( \frac{i_o}{4.1 \cdot \sqrt{1 + \frac{1}{I_{D7}}}} \right) \]
Appendix E

Switch-induced error voltage on an elementary PMOS-based sample-and-hold circuit

The aim of this appendix is to determine an analytical expression that evaluates the level of switch-induced error voltage on the elementary PMOS-based sample-and-hold circuit depicted in Fig 4.9. The analysis is based on the lumped models developed by Sheu and Hu [89] and assumes that half of the channel carriers exit at the hold node and a linear variation of the gate voltage between the "ON" and "OFF" steady states (i.e. ground and supply voltage $V_{dd}$). During the switching-off transient of the PMOS device two distinct periods may be identified.

In the first phase, when the absolute value of the gate-to-source potential is greater than that of $V_T$, the PMOS transistor is biased in the strong inversion region. This conductive state is represented by the equivalent lumped model given in Fig. E.1 (a). $C_G$ and $C_{ov}$ are respectively the gate and gate-drain overlap capacitances which may be expressed as

\begin{align*}
C_G &= \text{[Expression]} \\
C_{ov} &= \text{[Expression]}
\end{align*}
where $C_{ox}$ is the gate capacitance per unit area, $W$ and $L$ are respectively the drawn width and drawn length of the transistor and $L_D$ is the lateral diffusion distance.

Applying Kirchhoff's current principle at the drain node, it can be deduced that

$$I_{hold} = I_{ov} + I_G - I_D$$  
(E.3)

Assuming the input signal to be constant

$$I_{hold} = C_{hold} \cdot \frac{dV_o}{dt} = C_{hold} \cdot \frac{d(V_w + V_D)}{dt} = C_{hold} \cdot \frac{dV_D}{dt}$$  
(E.4)

Since the gate signal is a ramp which starts rising at the origin from ground toward supply voltage at a rate $U$

$$V_G = U.t$$  
(E.5)

the sum of the gate and gate-drain overlap capacitance currents can be expressed as

$$I_{ov} + I_G = \left( C_{ov} + \frac{C_G}{2} \right) \cdot \frac{d(V_G - V_D - V_w)}{dt}$$  
(E.6)

which may be approximated to

$$I_{ov} + I_G = \left( C_{ov} + \frac{C_G}{2} \right) \cdot \frac{dV_G}{dt} = \left( C_{ov} + \frac{C_G}{2} \right) U$$  
(E.7)

if one realistically assumes that $|d V_G / dt| \gg |d V_D / dt|$. Operated in the strong inversion region, the drain current of the PMOS transistor is given in Appendix A by

$$I_D = K_F \cdot \left\{ (V_T - [V_G - V_w]) \cdot V_D + \frac{v_D^2}{2} \right\}$$  
(E.8)
where $K_p$ is the transconductance parameter of the PMOS device and $V_T$ its threshold voltage which is given by

$$V_T = V_{TO} - \gamma \left( \sqrt{V_{dd} - V_w - \Phi_B} \right)$$  \hspace{1cm} (E.9)

Since the drain-to-source potential remains small with respect to $V_T - (V_G - V_w)$, (E.8) can be approximated to

$$I_D = K_p (V_{TW} - U.t) \cdot v_D$$  \hspace{1cm} (E.10)

where $V_{TW} = V_T + V_w$. Thus combining (E.3), (E.4), (E.7) and (E.10), during the conductive phase, the behaviour of the basic PMOS-based sample-and-hold circuit is described by the following differential equation

$$C_{hold} \cdot \frac{dv_D}{dt} = \left( C_{ov} + \frac{C_G}{2} \right) \cdot U - K_p (V_{TW} - U.t) \cdot v_D$$  \hspace{1cm} (E.11)

A solution to this equation may be obtained by utilising the well-known method of variation of parameters and is given by

$$v_D(t) = \sqrt{\frac{\pi \cdot U \cdot C_{hold}}{2 \cdot K_p}} \cdot \left( \frac{C_{ov} + \frac{C_G}{2}}{C_{hold}} \right) \cdot \exp \left( \frac{K_p}{2 \cdot U \cdot C_{hold}} \cdot [U.t - V_{TW}]^2 \right) \cdot \left[ \text{crf} \left( \sqrt{\frac{K_p}{2 \cdot U \cdot C_{hold}} \cdot V_{TW}} \right) - \text{crf} \left( \sqrt{\frac{K_p}{2 \cdot U \cdot C_{hold}} \cdot [V_{TW} - U.t]} \right) \right]$$  \hspace{1cm} (E.12)

This expression is valid until the gate-to-source potential reaches the limit of the above threshold conduction. At this time, $t_i = V_{TW} / U$, the first period is over and the error voltage is given by

$$v_D(t_i) = \sqrt{\frac{\pi \cdot U \cdot C_{hold}}{2 \cdot K_p}} \cdot \left( \frac{C_{ov} + \frac{C_G}{2}}{C_{hold}} \right) \cdot \text{crf} \left( \sqrt{\frac{K_p}{2 \cdot U \cdot C_{hold}} \cdot V_{TW}} \right)$$  \hspace{1cm} (E.13)
The equivalent lumped model of the circuit, in the subsequent turn-off phase, is given in Fig. E.1 (b). Since the transistor is biased in the weak inversion, the drain current and the gate capacitances have a negligible influence, the operation of the circuit is thus dictated by the following differential equation

\[
C_{\text{hold}} \frac{dV_D}{dt} = C_{ov} \cdot U \tag{E.14}
\]

for which the solution is

\[
V_D(t') = \frac{C_{ov}}{C_{\text{hold}}} \cdot U \cdot t' + \delta \tag{E.15}
\]

where \( \delta = V_D(t' = 0) \) is the magnitude of the switch-induced error voltage at the end of the first period. Thus the complete solution is

\[
V_D(t') = \frac{C_{ov}}{C_{\text{hold}}} \cdot U \cdot t' + \sqrt{\frac{\pi \cdot U \cdot C_{\text{hold}}}{2 \cdot K_p}} \cdot \left( \frac{C_{ov} + C_G}{2 \cdot C_{\text{hold}}} \right) \cdot \text{erf} \left( \sqrt{\frac{K_p}{2 \cdot U \cdot C_{\text{hold}}} \cdot V_{TW}} \right) \tag{E.16}
\]

This expression is valid for the gate-to-source voltage rising from \( V_{TW} \) to the supply voltage.

At the time \( t' = t_2 = (V_{dd} - V_{TW}) / U \), the switching-off transient is complete and the total amount of switch-induced error voltage on the holding capacitor is

\[
V_{DT} = \sqrt{\frac{\pi \cdot U \cdot C_{\text{hold}}}{2 \cdot K_p}} \cdot \left( \frac{C_{ov} + C_G}{2 \cdot C_{\text{hold}}} \right) \cdot \text{erf} \left( \sqrt{\frac{K_p}{2 \cdot U \cdot C_{\text{hold}}} \cdot V_{TW}} \right) + \frac{C_{ov}}{C_{\text{hold}}} \cdot (V_{dd} - V_{TW}) \tag{E.17}
\]
Figure E.1: Equivalent lumped models of the basic PMOS-based sample-and-hold circuit. (a) conductive phase, (b) turn-off phase.
Appendix F

Switch-induced error voltage on a dummy-compensated PMOS-based sample-and-hold circuit

An analytical expression which approximates the magnitude of the switch-induced error voltage on the dummy-compensated PMOS-based sample-and-hold circuit depicted in Fig. 4.10 is derived in this appendix. The analysis is based on the lumped models suggested by Sheu et al. [89] and assumes that the width of the switching device is twice that of the dummy transistor, the impedances on the drain and source side of the switch are identical and the clock signals have opposite slopes and ramp simultaneously in a linear fashion. Three different modes of operation may be identified.

At the origin of the switching-off phase, the switching transistor is biased in the strong inversion while the dummy device is weakly inverted, the equivalent lumped model of the dummy-compensated scheme is shown in Fig. F.1 (a), where all the parameters have been defined in the previous appendix. From the Kirchhoff's current law
\[ I_{\text{hold}} = I_{ov} + I_G + I_G^D - I_D \]  

\(^{(F.1)}\)

Based on the analysis presented in the previous appendix, and assuming that the clock signal of the dummy device falls as

\[ V_G^D = V_{dd} - U.t \]  

\(^{(F.2)}\)

it can readily be shown that, in the early stage of the switching, the transient characteristic is described by

\[ \frac{d V_D}{d t} = \frac{C_G}{2} \cdot U \cdot K_P \cdot (V_{TW} - U.t) \cdot V_D \]  

\(^{(F.3)}\)

The solution of this differential equation is

\[ V_D(t) = \sqrt{\frac{\pi \cdot U \cdot C_{\text{hold}}}{2 \cdot K_P}} \cdot \left( \frac{C_G}{2 \cdot C_{\text{hold}}} \right) \cdot \exp \left( \frac{K_P \cdot U \cdot \left[ t - \frac{V_{TW}}{U} \right]^2}{2 \cdot C_{\text{hold}}} \right) \cdot \left\{ \text{erf} \left( \sqrt{\frac{K_P}{2 \cdot U \cdot C_{\text{hold}}} \cdot V_{TW}} \right) - \text{erf} \left( \sqrt{\frac{K_P}{2 \cdot U \cdot C_{\text{hold}}} \cdot (V_{TW} - U) \cdot t} \right) \right\} \]  

\(^{(F.4)}\)

which is valid until the gate-to-source potential of the dummy transistor falls to \(V_{TW}\) at which point the dummy device enters into the strong inversion regime. At this time, \(t_i = (V_{dd} - V_{TW}) / U\), the level of error voltage is

\[ V_D(t_i) = \sqrt{\frac{\pi \cdot U \cdot C_{\text{hold}}}{2 \cdot K_P}} \cdot \left( \frac{C_G}{2 \cdot C_{\text{hold}}} \right) \cdot \exp \left( \frac{K_P}{2 \cdot U \cdot C_{\text{hold}}} \cdot [V_{dd} - 2 \cdot V_{TW}]^2 \right) \cdot \left\{ \text{erf} \left( \sqrt{\frac{K_P}{2 \cdot U \cdot C_{\text{hold}}} \cdot V_{TW}} \right) - \text{erf} \left( \sqrt{\frac{K_P}{2 \cdot U \cdot C_{\text{hold}}} \cdot [2 \cdot V_{TW} - V_{dd}]} \right) \right\} \]  

\(^{(F.5)}\)

and the second period commences. During this phase both switching and dummy devices are strongly inverted. This state is modelled in Fig. F.1 (b) and described by the following system equation

\[ C_{\text{hold}} \cdot \frac{d V_D}{d t} = - I_D = - K_P \cdot (V_{TW} - U.t') \cdot V_D \]  

\(^{(F.6)}\)
for which the solution is

\[ \nu_D(t') = \delta \exp\left(\frac{K_p}{2U_C_{hold}} \left[ U(t') - V_{TW}\right]^2\right) \]  

(F.7)

where \( \delta \) is determined by the level of switch-induced voltage at the origin of this period (F.5). Since \( \nu_D(t' = 0) = \nu_D(t_1) \), it can readily be shown that

\[ \delta = \sqrt{\frac{\pi U_C_{hold}}{2K_p}} \left( \frac{C_G}{2C_{hold}} \right) \exp\left( \frac{K_p}{2U_C_{hold}} \left[ (V_{dd} - 2V_{TW})^2 - V_{TW}^2 \right] \right). \]

\[ \text{erf}\left( \sqrt{\frac{K_p}{2U_C_{hold}}} V_{TW} \right) - \text{erf}\left( \sqrt{\frac{K_p}{2U_C_{hold}}} [2V_{TW} - V_{dd}] \right) \]  

(F.8)

Substituting (F.8) back into (F.7), the complete solution of (F.6) is

\[ \nu_D(t') = \sqrt{\frac{\pi U_C_{hold}}{2K_p}} \left( \frac{C_G}{2C_{hold}} \right) \exp\left( \frac{K_p}{2U_C_{hold}} \left[ (V_{dd} - 2V_{TW})^2 - V_{TW}^2 + [U(t') - V_{TW}]^2 \right] \right). \]

\[ \text{erf}\left( \sqrt{\frac{K_p}{2U_C_{hold}}} V_{TW} \right) - \text{erf}\left( \sqrt{\frac{K_p}{2U_C_{hold}}} [2V_{TW} - V_{dd}] \right) \]  

(F.9)

which is valid until the switching device reaches the turn-off state (i.e. \( V_G = V_{TW} \)). The time required to reach this state is \( t_2 = \frac{2V_{TW} - V_{dd}}{U} \). At this point the magnitude of the error is

\[ \nu_D(t_2) = \sqrt{\frac{\pi U_C_{hold}}{2K_p}} \left( \frac{C_G}{2C_{hold}} \right) \exp\left( \frac{K_p}{2U_C_{hold}} \left[ V_{dd}^2 + 2V_{TW} - 3V_{dd}V_{TW} \right] \right). \]

\[ \text{erf}\left( \sqrt{\frac{K_p}{2U_C_{hold}}} V_{TW} \right) - \text{erf}\left( \sqrt{\frac{K_p}{2U_C_{hold}}} [2V_{TW} - V_{dd}] \right) \]  

(F.10)

and the switching device enters the final phase of turn-off. During this phase, for which the model is depicted in Fig F.1 (c), the gate-overlap capacitor of the switching device tends to increase the error voltage while the gate and gate-overlap capacitors of the dummy transistor over-compensate for this effect. Based on a technique similar to that
used to evaluate the error signal during the turn-off state of the basic sample-and-hold scheme (see previous appendix for details), it can be deduced that, when both clock signals \( \phi \) and \( \phi' \) have reached their final steady states, the total amount of switch-induced error voltage on the capacitor is

\[
\nu_{DT} = \sqrt{\frac{\pi U C_{\text{hold}}}{2.4L}} \cdot \frac{C_G}{2C_{\text{hold}}} \cdot \exp \left( \frac{K_p}{2U C_{\text{hold}}} \cdot \left\{ V_{dd}^2 + 2V_{TW} - 3V_{dd} \cdot V_{TW} \right\} \right) \\
\left\{ \text{erf} \left( \sqrt{\frac{K_p}{2U C_{\text{hold}}} \cdot V_{TW}} \right) - \text{erf} \left( \sqrt{\frac{K_p}{2U C_{\text{hold}}} \cdot [2V_{TW} - V_{dd}] \right) \right\} - \frac{C_{ox}}{2C_{\text{hold}}} (V_{dd} - V_{TW})
\]

(F.11)
Figure F.1: Equivalent lumped models of the PMOS-based dummy-compensated sample-and-hold circuit. (a) \(0 \leq V_G \leq V_{dd} - (V_T + V_w)\); (b) \(V_{dd} - (V_T + V_w) < V_G \leq V_T + V_w\); (c) \(V_T + V_w < V_G \leq V_{dd}\).
Appendix G

Yield analysis of the digital-to-analogue converter

The yield analysis presented in this appendix is derived in connection with the structure of the digital-to-analogue converter (DAC) presented in Fig. 4.12 and is a direct extension of the method presented by Lakshmikumar et al. [81].

The gain error and the integral linearity of the DAC are mainly determined by the accuracy of the individual current sources. While the gain error can be compensated for by adjusting the biasing current \( I_p \), the integral linearity depends on the matching properties of the n-channel MOS transistors \( (M_1-M_{10}) \). It is well known that, in CMOS technology, process parameter deviations are the results of random variations which can be characterised by the probability density function known as the normal (or Gaussian) distribution [65]. Thus, as shown in [81], the variance of the normalised output of the DAC, for any given input digital word, can be expressed as a function of the variance of the unit current source

\[
\sigma_\tilde{z}^2 = \frac{\tilde{z}(1-\tilde{z})}{I_F \cdot (I_{DAC} + I_{DAC}^-)} \cdot \sigma_{I_p}^2
\]  

(G.1)
where \( I_F \) and \( \sigma_{IF}^2 \) are respectively the mean and variance of the unit current source. \( I_{DAC}^- \) is the mean value of the DAC's output and \( I_{DAC}^\prime \) its analogue complement which can respectively be formulated as

\[
I_{DAC}^- = \frac{I_F}{4} \sum_{i=1}^{\frac{g}{2}} 2^{-i} \cdot b_i
\]  

\( G.2 \)

and

\[
I_{DAC}^\prime = \frac{I_F}{4} \sum_{i=1}^{\frac{g}{2}} 2^{-i} \cdot (1 - b_i)
\]  

\( G.3 \)

where \( b_i \) is the \( i^{th} \) digital input of the DAC and is either 0 or 1. The expected value of the DAC's output, normalised to full scale, may be shown to be

\[
\bar{z} = \frac{I_{DAC}^-}{I_{DAC}^- + I_{DAC}^\prime}
\]  

\( G.4 \)

Substituting (G.2) and (G.3) into (G.4), it can readily be shown that

\[
\bar{z} = \frac{j}{255}
\]  

\( G.5 \)

where \( j \) is a decimal representation of the input digital word and can take any value between 0 and 255. Thus, substituting (G.2), (G.3) and (G.5) into (G.1), the variance of the normalised DAC's output may be re-written as

\[
\sigma_z^2 = \frac{j(255 - j)}{255^2 \left( 63 + \frac{1}{2} + \frac{1}{4} \right)} \frac{\sigma_{IF}^2}{I_F^2}
\]  

\( G.6 \)

Note that the variance is maximum (i.e. \( d \sigma_z^2 / d j = 0 \)) when the digital input is half that of the full scale and nil for a maximum and minimum output. Based on a Gaussian distribution, the probability that the normalised output of the DAC, for any given input digital word, has an integral linearity of \( \pm 1 \) LSB is given by
\[ P\left( \frac{z}{255} - \frac{1}{255} < z < \frac{z}{255} + \frac{1}{255} \right) = \frac{1}{\sqrt{2\pi} \cdot \sigma_z} \int_{-\frac{1}{255}}^{\frac{1}{255}} \exp\left( -\frac{1}{2} \left( \frac{z - z}{\sigma_z} \right)^2 \right) dz \]  

(G.7)

which may be simplified to

\[
P\left( \frac{z}{255} - \frac{1}{255} < z < \frac{z}{255} + \frac{1}{255} \right) = \text{erf}\left( \frac{1}{255 \cdot \sqrt{2} \cdot \sigma_z} \right) \]  

(G.8)

The circuit yield of the DAC is obtained [81] by multiplying the probabilities that each of the 256 possible outputs have less than ±1 LSB error

\[
G = \prod_{j=1}^{254} \text{erf}\left( \frac{1}{255 \cdot \sqrt{2} \cdot \sigma_z} \right) \]  

(G.9)

Substituting (G.6) into (G.9), the circuit yield of the DAC is

\[
G = \prod_{j=1}^{254} \text{erf}\left( \sqrt{\frac{63 + \frac{1}{2} + \frac{1}{4}}{2 \cdot (255 - j) \cdot \sqrt{2} \cdot \frac{\sigma_{1p}}{1F}}} \right) \]  

(G.10)
Appendix H

Published papers
CMOS subthreshold-mode I/V converter for analogue neural network applications

D. Couë and G. Wilson

Indexing terms: Neural networks, CMOS integrated circuits

An improved I/V converter for analogue neural network applications is presented, using devices operating in the subthreshold mode of conduction. The proposed scheme employs diode connected transistors biased by a network incorporating feedback. Analysis and simulation suggest that the modified structure offers a substantial reduction in the sensitivity to supply voltage variations. As the proposed scheme uses only NMOS transistors, it can be implemented in a standard single well process.

Introduction: Analogue neural networks (ANNs) are typically organised as parallel layers of processing units (neurons) interconnected by elements defined as synapses. Synaptic multiplications in an ANN are often implemented using linear transconductance multipliers [1, 2] allowing the neuronal summation operation to be achieved by a simple physical connection of the synaptic outputs. The summed output current is then converted to an equivalent potential, as required by the following layer of synapses, using a nonlinear resistive load.

Where ANN architectures require a large number of resistive loads, as for example when the backpropagation algorithm is implemented on-chip [3] or reconfigurability is involved [4], it is highly desirable, if not essential, that the silicon area associated with each load-resistor circuit, as determined by the number and size of the transistors involved, be minimised, and in particular that power dissipation be kept to a minimum. In this context, designs exploiting the low current levels associated with subthreshold operation are receiving much attention. However, because the drain current in a subthreshold mode device is an exponential function of the gate-source voltage, the I/V characteristics of load resistors are highly sensitive to the biasing arrangements. This letter illustrates the sensitivity problem and describes a load-resistor configuration for which the sensitivity to power supply variations is substantially improved.

Sensitivity problem: In several reported ANN designs [3-5], I/V conversion has been achieved using the CMOS arrangement of biased diodes as shown in Fig. 1. Note that, for a fully differential scheme, the resistor circuit consists of four matched transistors M1-M4 and two biasing voltages E1, E0. With the devices operating in the subthreshold mode of conduction, the output difference voltage is given by

\( v_0 = v_1 - v_2 = \frac{2V_s}{\pi} \sinh \left( \frac{\phi}{2} \right) \)  

where \( V_s = kTq \approx 26mV \) at room temperature, \( \phi \) is a measure of the effectiveness of the gate in controlling the channel current and \( \phi \) is the input current.

If the qessient bias current given by

\( I = I_L \exp \left( \frac{V_{ds} - (E_0 + E_1)}{2V_s} \right) \)  

where \( V_{ds} \) is the supply voltage and \( I \) is the characteristic current for the transistors. Given that the sin function can be approximated as

\( \sinh(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots \)  

for \( |x| < \infty \) then for \( |I| \lesssim 2I \) the effective driving point resistance may be approximated to the first order as

\[ r_D = \frac{V_0}{I_0} \approx \frac{V_1}{2kT} - \frac{V_1}{2kT} \exp \left( \frac{V_{ds} - (E_0 + E_1)}{2V_s} \right) \]

It will be appreciated therefore that this resistance is highly sensitive to variations in both the supply and bias sources. The following section describes a modified resistive load incorporating a feedback mechanism which substantially improves the sensitivity to power supply and biasing variations and is suitable for NMOS processes.

Proposed scheme: After rearranging the positions of the bias sources \( E_0 \) and \( E_1 \) as shown in the circuit diagram of Fig. 2, the core diode group \((M_1-M_4)\) and the associated biasing elements can be implemented using only NMOS devices. The biasing potentials are generated by two diode connected transistors \((M_5, M_6)\) and are controlled by a feedback arrangement comprising transistors \(M_7-M_9\), together with the current source \(I_F\). Since the potential across the diodes \(M_7-M_9\), is nominally fixed by \(I_F\), variations in the supply \(V_{dd}\) are distributed across the gate-source junctions of the shunt transistors \(M_1-M_4\) and the biasing transistor \(M_6\). With the proviso that these transistors have matching characteristics, it follows that half of the supply voltage variations will appear across the device \(M_6\). The symmetry of the scheme also follows that an equal variation will appear across the diode connected transistor \(M_7\), with the biasing potential across the core diode group buffered from supply variations the load resistance sensitivity is correspondingly reduced.

The reduction in sensitivity can be demonstrated with the assumption that the operating conditions are such that the NMOS current transistors are related to their gate-source potentials as

\[ I(t) = I_F \exp \left( \frac{V_{gs}}{V_T} \right) \]

Reprinted from ELECTRONICS LETTERS 23rd May 1996 Vol. 32 No. 11 pp. 990-991
It can be shown that, for well matched devices, the potentials across diodes \( M_1 \) and \( M_4 \) can be expressed as
\[
E_n = [V_{dd} - (E_1 + E_2)] + \frac{V_L}{\kappa} \ln \left( 1 + \frac{I_D^1 + I_D^2}{I_D^T} \right)
\]
(6)

However, \( E_n \) is also related to the bias current by the following expression:
\[
I_D = I_s \exp \left( \frac{E_n}{2V_T} \right)
\]
(7)

Combining eqns. 1, 6 and 7, the output difference voltage may be shown to be
\[
u_o = \frac{2V_T}{\kappa} \sinh \left( \frac{V_o}{4V_T} \right) \left( 1 + \frac{I_C}{I_D^T} \right)
\]
(8)

where \( I_{dm} \) is the drain current for device \( M_1 \) and \( I_s \) is the current biasing the core diode group. It can be seen that the proposed scheme offers a resistive characteristic which is set by the bias current and the current ratio \( I/I_{dm} \). Since \( I_s \) is nominally independent it follows that the sensitivity of the resistance value is determined by the static current ratio \( I/I_{dm} \). Any given level of insensitivity can be achieved by ensuring that the shunt current \( I_m \) is sufficiently larger than the core group current \( I_s \). Improved sensitivity is thus acquired at the expense of an increase in the static current consumption. Given the subthreshold mode of operation and the advantages of an NMOS implementation, the static current burden should not present a problem.

The behaviour of the proposed \( I/V \) converter has been assessed via a series of PSPICE simulations based on a 3V power supply using level 3 models with parameters for a 2.4μm CMOS process. The circuit was simulated for power supply voltage variation of ±100mV, with a bias current \( I_s \) of 2nA and an aspect ratio of \( W/L = 10/1 \) for all transistors. The results suggest a sensitivity improvement >40:1 is readily achievable, and confirm that as the static current ratio \( I/I_{dm} \) decreases the sensitivity of the nominal resistance to power supply variations improves.

© IEE 1996

References
5 COLE, D., and WILSON, O.: 'A four-quadrant subthreshold mode multiplier for analogue neural network applications', IEEE Trans. Neural Netw., to be published

270
A Four-Quadrant Subthreshold Mode Multiplier for Analog Neural-Network Applications

Dominique Coué and George Wilson

Abstract—A new four-quadrant CMOS analog multiplier is presented, based on devices operating in the subthreshold mode of conduction. The proposed circuit is a cross-coupled quad structure in which differential multiplication is obtained by driving the gate and bulk (back gate) terminals of the devices. Analysis and simulation have shown that the new structure has the characteristics required for the design of very large scale integration (VLSI) analog neural networks. Although operating at subthreshold current levels, reasonable speed can nevertheless be obtained since voltage swings are in the range of a few \( V_T \). The behavior of the basic multiplier has been assessed experimentally using transistor-arrays and simulation studies on a network including 11 neurons and 31 synapses indicate a useful level of functionality.

I. INTRODUCTION

THE four-quadrant multiplier is an important building block for a large number of signal processing applications, and particularly in analog neural networks (ANN’s). Neural networks are typically organized as parallel layers of processing units (neurons) interconnected by elements defined as synapses [1] as illustrated in Fig. 1. The output of a synapse is the product of its input (output of previous layer) and a weight. The function performed by the neural network is determined by its topology and the weights associated with each interconnection. Applications typically require a large number of interconnected neurons and therefore synaptic connections i.e., multipliers. It is desirable therefore, if not essential, that multiplying elements use a minimum number of active devices and dissipate minimum power. In MOS technologies the synaptic element may be designed using transistors operating in the saturation, linear, or subthreshold regions. Subthreshold operation has the advantage that current levels are typically lower than devices biased into strong inversion, but operating speeds are diminished due to the reduced ability to charge/discharge capacitive elements. Nevertheless, subthreshold mode of operation may be attractive, because of the lower power levels involved and with voltage swing requirements in the order of few \( V_T \) the relatively low device capacitances allow reasonable speeds to be achieved.

Several reported synaptic designs [2]–[5] exploit the MOS version of the Gilbert cell [6] (see Fig. 2) which utilizes six transistors arranged as stacked differential pairs. A fixed tail current \( I_D \) is distributed between the differential pairs \( P_1 \) and \( P_2 \) via a third differential pair \( P_3 \) in a manner determined by the difference voltage, \( \nu_Y = V_{X_1} - V_{X_2} \). The difference output current of the Gilbert structure is given by

\[
I_D = (I_3 + I_4) - (I_2 + I_4) = I_D \cdot \tanh \left( \frac{\nu_Y}{2V_T} \right) \cdot \tanh \left( \frac{\nu_X}{2V_T} \right)
\]

where \( \nu_X = V_{X_1} - V_{X_2} \), \( V_T = kT/q \), is approximately 26 mV at room temperature and \( \kappa \) is a measure of the effectiveness of the gate in controlling the channel current. Given that the \( \tanh \) function can be expanded as

\[
tanh x = x - \frac{1}{3} x^3 + \frac{2}{15} x^5 - \frac{17}{315} x^7 + \cdots \quad \text{for} \quad |x| < \frac{\pi}{2}
\]

where

\[
I_D = (I_3 + I_4) - (I_2 + I_4)
\]

Fig. 2. MOS version of the Gilbert multiplier.

\[
I_D = I_D \cdot \tanh \left( \frac{\nu_Y}{2V_T} \right) \cdot \tanh \left( \frac{\nu_X}{2V_T} \right)
\]

1. INTRODUCTION

THE four-quadrant multiplier is an important building block for a large number of signal processing applications, and particularly in analog neural networks (ANN's). Neural networks are typically organized as parallel layers of processing units (neurons) interconnected by elements defined as synapses [1] as illustrated in Fig. 1. The output of a synapse is the product of its input (output of previous layer) and a weight. The function performed by the neural network is determined by its topology and the weights associated with each interconnection. Applications typically require a large number of interconnected neurons and therefore synaptic connections i.e., multipliers. It is desirable therefore, if not essential, that multiplying elements use a minimum number of active devices and dissipate minimum power. In MOS technologies the synaptic element may be designed using transistors operating in the saturation, linear, or subthreshold regions. Subthreshold operation has the advantage that current levels are typically lower than devices biased into strong inversion, but operating speeds are diminished due to the reduced ability to charge/discharge capacitive elements. Nevertheless, subthreshold mode of operation may be attractive, because of the lower power levels involved and with voltage swing requirements in the order of few \( V_T \) the relatively low device capacitances allow reasonable speeds to be achieved.

Several reported synaptic designs [2]–[5] exploit the MOS version of the Gilbert cell [6] (see Fig. 2) which utilizes six transistors arranged as stacked differential pairs. A fixed tail current \( I_D \) is distributed between the differential pairs \( P_1 \) and \( P_2 \) via a third differential pair \( P_3 \) in a manner determined by the difference voltage, \( \nu_Y = V_{X_1} - V_{X_2} \). The difference output current of the Gilbert structure is given by

\[
I_D = (I_3 + I_4) - (I_2 + I_4)
\]

Fig. 1. Architecture of a neural network.

\[
I_D = I_D \cdot \tanh \left( \frac{\nu_Y}{2V_T} \right) \cdot \tanh \left( \frac{\nu_X}{2V_T} \right)
\]

where \( \nu_X = V_{X_1} - V_{X_2} \), \( V_T = kT/q \), is approximately 26 mV at room temperature and \( \kappa \) is a measure of the effectiveness of the gate in controlling the channel current. Given that the \( \tanh \) function can be expanded as

\[
tanh x = x - \frac{1}{3} x^3 + \frac{2}{15} x^5 - \frac{17}{315} x^7 + \cdots \quad \text{for} \quad |x| < \frac{\pi}{2}
\]
II. THE PROPOSED MULTIPLIER

An NMOS version of the proposed four quadrant-multiplier is shown in Fig. 3 and consists of a current source and differential pairs $P_1$ and $P_2$ of matched transistors operating in the subthreshold mode of conduction. One differential input is applied as in the Gilbert structure while the second appears between the bulk terminals of $M_1(M_2)$ and $M_2(M_2)$. It should be noted that using the well as a multiplier input has previously been described in connection with single-quadrant multipliers [3]. The proposed scheme thus represents a natural extension of the back gate effect to yield the increased functionality associated with full four-quadrant multiplication.

For an NMOS transistor, the subthreshold current $I_{DS}$ is given by

$$I_{DS} = I_D \cdot \exp \left( \frac{1}{\eta} \cdot \frac{V_{DS}}{V_T} \right) \cdot \exp \left( \frac{\kappa \cdot V_{DS}}{V_T} \right) \cdot \left( 1 - \exp \left( -\frac{V_{DS}}{V_T} \right) + \frac{V_{DS}}{V_T} \right)$$

(4)

where $V_{DS}$ is the gate-to-source potential, $I_D$ is the drain-to-source current, $V_T$ is the Early voltage and $I_0 = W/L \cdot I_{DS}$, where $I_{DS}$ is a characteristic current. With the assumptions that the drain-source quiescent potential $V_{DS} \geq 4 \cdot V_T$ but remains low enough to disregard Early effect, (4) reduces to

$$I_{DS} = I_D \cdot \exp \left( \frac{\kappa \cdot V_{DS}}{V_T} \right) \cdot \exp \left( \frac{1}{\eta} \cdot \frac{V_{DS}}{V_T} \right).$$

(5)

Fig. 3. Circuit diagram of the proposed multiplier.

Fig. 4. Static characteristics of the analog multiplier: (a) $I_D$ against $V_{DS}$ and (b) $I_D$ against $V_T$.

then for $|\nu X| \leq V_T/\kappa$, (1) may be approximated to a first order as

$$i_0 = \frac{I_0 \cdot a^2}{4 \cdot V_T^2} \cdot \nu X \cdot \nu Y.$$  

(3)

Fig. 5. Experimental static characteristics of the proposed multiplier.

It can be seen that the Gilbert circuit operates as a linear transconductance multiplier for input voltage differences less than $V_T/\kappa$. In the next section it will be shown that it is possible to obtain a similar relationship in an arrangement in which the bias current is distributed across two differential pairs by modulating their bulk potentials. In Section III, simulation and experimental results for the multiplier are given and in Section IV, it is shown how the proposed structure can be used in the design of a neural network.
It should be emphasized that the normal operation of MOS transistor requires the source/bulk and drain/bulk junctions be reverse-biased, i.e., $V_{sb} \leq 0$ and $V_{db} \leq 0$.

Given that the sum of the device currents at the common source node can be expressed as

$$I_b = I_s \cdot \exp \left( \frac{-V_s}{V_t} \right) \cdot \left[ \exp \left( \frac{1 - \kappa}{V_t} \cdot \frac{V_{sb}}{V_t} \right) + \exp \left( \frac{\kappa \cdot V_{sb}}{V_t} \right) \right]$$

the difference current output becomes

$$i_0 = I_s \cdot \tanh \left[ \frac{\kappa \cdot V_{sb}}{2 \cdot V_t} \cdot \tanh \left( \frac{1 - \kappa}{V_t} \cdot \frac{V_{sb}}{V_t} \right) \right].$$

Note that (1) and (7) differ only in the coefficient of the factor, $\nu_s/2 \cdot V_t$ which is $1 - \kappa$ for the proposed scheme and $K$ for the Gilbert cell. For a typical $\kappa$ value of 0.8 [9], this variation has the effect of increasing the linear range of the $Y$ differential input of the proposed multiplier to $V_t/(1 - \kappa) \approx 130 \text{ mV}$. Thus over the signal ranges $i_x/V_x < V_t/K$ and $V_t < V_t/(1 - \kappa)$, (7) can be approximated to a first order as

$$i_0 = I_s \cdot \kappa \cdot \frac{(1 - \kappa)}{4 \cdot V_t^2} \cdot \nu_s \cdot V_t$$

and is similar to the result obtained for the Gilbert cell.

1) **Limits of Operation:** To ensure proper operation of the proposed circuit, the source/bulk junctions of transistors $M_1$-$M_4$ should not be forward biased. This requires that

$$V_t + \nu_Y \leq V_s + \nu_s$$

where $V_t$ is the quiescent bulk common-mode input level, $\nu_Y$ is the dynamic bulk input voltage, $V_s$ and $\nu_s$ are the quiescent and dynamic source potentials, respectively. The quiescent and dynamic source voltage can be obtained by rearranging (6) as follows:

$$V_s = V_t \cdot \left[ \ln \left( \frac{4 \cdot I_s}{I_b} \right) + \frac{\kappa \cdot V_x}{V_t} + \frac{(1 - \kappa) \cdot \nu_Y}{V_t} \right]$$

and

$$\nu_s = V_t \cdot \left[ \ln \left( \frac{1 + \exp \left( \frac{\kappa \cdot \nu_s}{V_t} \right)}{2} \right) + \ln \left( \frac{\kappa \cdot V_x}{V_t} + \frac{1 + \exp \left( \frac{1 - \kappa}{V_t} \cdot \nu_Y \right)}{2} \right) \right].$$

where $V_x$ and $\nu_s$ are the quiescent and dynamic gate potentials, respectively. Combining (9), (10), and (11), the bulk dynamic input range can be expressed as

$$|\nu_Y| \leq V_t \cdot \left[ \ln \left( \frac{4 \cdot I_s}{I_b} \right) + \frac{\kappa \cdot (V_x - V_Y)}{V_t} \right]$$

and an upper limit for $\nu_Y$ can be obtained by setting $\nu_s$ at $\nu_s = -10 \cdot V_t$ (which represents a practical lower bound) giving

$$|\nu_Y|_{\text{max}} = V_t \cdot \left[ \ln \left( \frac{4 \cdot I_s}{I_b} \right) + \kappa \cdot \frac{(V_x - V_Y)}{V_t} - 0.7 \right].$$
COUE AND WILSON: A FOUR-QUADRANT SUBTHRESHOLD MODE

Fig. 7. Feedforward neural network with the activation function distributed over the following synapses.

Therefore, in the context of neural-network design, $I_s, V_X, V_Y$ and the aspect ratio of the transistors ($W/L$), would be chosen in order to obtain a bulk dynamic input range of a few $V_T$.

III. SIMULATION AND EXPERIMENTAL RESULTS

1) Simulation Results: The behavior of the proposed multiplier has been assessed via a series of PSpice simulations based on a 2 V power supply using Level 3 models with parameters for a 1.5 $\mu$m CMOS process. The simulations typically show that for a bias current of 100 nA and an aspect ratio of $W/L = 6/1$ for the transistors $M_1-M_4$, a bulk dynamic input range of 150 mV can be obtained with the common mode voltage of the gate and bulk terminals set to 1 V and 100 mV, respectively.

The family of simulated static characteristics in Fig. 4(a) and (b) shows the differential output current as a function of the differential inputs $V_X, V_Y$ over the range $-50$ mV to $50$ mV in equal 20 mV increments, respectively, and confirm that the multiplier generates differential output currents related...
Fig. 9. Load.

Fig. 10. Sigmoid activation function.

IV. DESIGN AND SIMULATION OF AN ANALOG NEURAL NETWORK

1) Design Method: For a feedforward network (see Fig. 6), the neuron transfer function can be expressed in the form

\[ y_j = \sum w_{jk} x_k + b \]

where \( f(\cdot) \) is a nonlinear activation (threshold) function and \( x_k \) the output of the \( k \)th neuron in layer \( m \). The inputs to the neuron are usually referred to as connections, and the \( w_{jk} \) as weights. Optionally, a bias \( b \) is added to the weighted sum. Each input is either driven by the output of a neuron of the previous layer \( x_k \) or an external input. In the following section we describe an ANN implementation in which the inherent tanh function of the transcondiitance multiplier described earlier is exploited to perform both the activation function and multiplication.

The modified ANN structure is based on the observation that the activation function of a neuron can be distributed over each of the forward layer synapses without affecting the overall behavior, as shown in Fig. 7. We shall refer to the combined thresholding and synaptic weighting as a thresholding-synapse (TS) block. The output and input of the TS block are related as

\[ f(S_j) = f(\frac{\sum w_{jm} x_j + b}{f_{\text{out}(\text{tanh})}}) \]

where \( f(\cdot) \) is the nonlinear activation function and \( f_{\text{out}(\text{tanh})} \) the output of the tanh function.
in the layer $m$ and the driven neuron in the layer $m+1$, respectively.

2) Circuit Implementation: An analog circuit implementation of the neural architecture shown in Fig. 7 can be realized using the transconductance-thresholding-synapse (TTS) shown in Fig. 8. The summation function can be achieved simply by connecting TTS outputs to a common bus bar. With the assumption that the designer has access to a twin-well process, the summed current can subsequently be converted to an equivalent balanced potential, as required by the following layer TTS blocks, using the load arrangement shown in Fig. 9.

It has been demonstrated (14) that in the linear range of $V_{ij}(|V_{ij}| < 2 \cdot \sqrt{2/(1 - \kappa)})$ the difference output current of a TTS can be expressed in the form

$$I_{pj} = g \cdot V_{oj} \cdot \tanh \left( \frac{V_{oj}}{2 \cdot V_j} \right).$$

(17)

3) Load: The summed current $I_{sj}$ is converted to a voltage $V_{sj}$ using a load comprising four transistors operating as diodes. For matched transistors, the difference output voltage is given by

$$V_{sj} = \frac{2 \cdot V_j}{\kappa} \cdot \sinh^{-1} \left( \frac{I_{sj}}{4 \cdot I_S} \right).$$

(18)

where $I_S = I_S \cdot \exp(\kappa \cdot V_{dd} - 2E/2 \cdot V_j)$; $V_{dd}$ is the supply voltage and $E$ is the biasing voltage of the diodes. Combining (17) and (18) the normalized output and input variables are related as

$$P_j = \frac{g}{4 \cdot I_S} \cdot V_{oj} \cdot \tanh(\sinh^{-1}(S_j))$$

(19)

where $P_j = I_{pj}/4 \cdot I_S$ and $S_j = I_{sj}/4 \cdot I_S$.

When this expression is related to (16) it can be seen that the value of the synaptic weight is given by $w_j = g/4 \cdot I_S \cdot V_{oj}$ and the activation function is $\tanh(\sinh^{-1}(\cdot))$. The $\tanh(\sinh^{-1}(\cdot))$ function is similar to the most common form of activation function such as the sigmoid function (hyperbolic tangent), as shown in Fig. 10.

An analog implementation of a feedforward neural network using TTS and load circuits is illustrated in Fig. 11. It can be noticed that the activation function of the output layer is achieved using a load and a TTS circuit for which the weight is set to one. It may also be noted that the input data of the ANN will have to be preprocessed to account for the influence of the $\tanh$ functions on the input layer.

4) Design Aspects: It can be noted in (19) that, by proper choice of the bias current $I_b$ and the current $I_S$, the maximum weight can be set to a desired value. The maximum weight value was set to 10, when the bias current was 250 nA and $I_S = 3.125$ nA.

The speed of a TTS is determined by the current it can source/sink at its output and by the resistive-capacitive load it has to drive. A simplified small-signal equivalent circuit of a TTS and load is shown in Fig. 12. The current $I$ is controlled by the output current of the TTS, $R$ represents the resistive load, and $C_L$ is the total capacitance at the TTS output

$$C_L = N \cdot C_i + C_e$$

(20)

where $C_i$ is the average input capacitance of a TTS (including wiring capacitances), $N$ is the number of TTS inputs con-
connected at the output of the driving TTS-load, and $C_L$ is the capacitance of the current-to-voltage converter.

If $I_{\text{max}}$ is the maximum current available at the output of a TTS, and $C_L$ is assumed to be linear, then the rate of change of the output voltage is defined by the nonlinear differential equation

$$I_{\text{max}} = C_L \frac{dV}{dt} + 4 \cdot I_S \cdot \sinh \left( \frac{\kappa \cdot V}{2 \cdot V_i} \right).$$

(21)

For a maximum output current $I_{\text{max}} = I_b/2 = 125$ nA (single-ended), a solution of (21) suggests a settling time to within 5% of error of about 0.65 $\mu$s for a capacitive load of 0.25 pF ($N \approx 10$).

The choice of $W/L$ ratios for the transistors in the transconductance-thresholding-synapse are made keeping in mind five different issues, mainly, the accuracy of computation, the area of the TTS, the speed of operation, the maximum weight value, and the power dissipation.

5) Simulation: The utility of the proposed multiplying structure has been assessed via a neural-network simulation in PSpice of the function $f(x) = 0.8 \cdot \sin^{-1} \left( \frac{V}{V_i} \right)$. (22)

The structure of the network implementing (22) is shown in Fig. 13 and comprises 11 neurons and 31 synapses arranged in three layers. It may be noted that an extra TTS circuit has been added to each neuron to bias the weighted sum. The inputs to these bias synapses were set to unity, i.e., $V_b = 2 \cdot V_i / \kappa$.

The neural network was trained using the error backpropagation algorithm with the learning rate set to 0.2 [10]. Whereas the weights of a network to be trained are usually initialized at small random values, in this case, the training time was shortened using initial weights set to values given by the trained digital equivalent. Fig. 14 shows the trained output of the network after 80 iterations, together with the desired function and the corresponding error which has an RMS value of approximately 6%.

It will be appreciated that a physical silicon implementation would not necessarily exhibit the performance predicted by simulation for several reasons including modeling imperfections and process/device tolerances [8–9]. However, these nonidealities can be compensated for during training of the neural-network system [5].

In practice it would also be necessary to provide additional dynamic analog weight storage circuitry to facilitate weight adjustment in each synapse during the learning process. These weights can be stored as potentials on storage capacitors, and set, via access switches, by circuits that convert values stored in digital memory into analog signals [11]. It may be noted that the weights cannot be stored precisely due to charge leakage through parasitic paths and charge injection from the access switches. However, switch charge injection can be reduced to few millivolts using compensation techniques [12] and the error introduced by the charge leakage can be limited by refreshing the weight voltages periodically [13].

V. CONCLUSION

An analog multiplier based on transistors operating in the subthreshold mode of conduction has been presented. Analysis and simulation have shown that the new scheme produces a difference output current related to the product of difference voltages via a tanh function. It has been demonstrated that the new multiplier can be applied to the design of analog very large scale integration (VLSI) neural networks. The proposed circuit is area efficient, has a low power dissipation and the nonlinearity due to the tanh function can be used to perform the activation function. Although operating at subthreshold current levels, reasonable speeds can be obtained since voltage swings are in the range of a few $V_i$.

REFERENCES

COUE AND WILSON: A FOUR-QUADRANT SUBTHRESHOLD MODE


