

Design and Implementation of GaAs CCD/MESFET ICs for Artificial Neural Network Application

by

LIDONG CHEN

B.Sc., Anhui University, Hefei, China, 1983

M.Sc., Anhui University, Hefei, China, 1986

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical and Computer Engineering

We accept this dissertation as conforming to the required standard

Dr. H. H. L. Kwok, Supervisor (Department of Elec. & Comp. Eng.)

Dr. F. El-Guibaly, Departmental Member (Department of Elec. & Comp. Eng.)

Dr. J. M.-S. Kim, Departmental Member (Department of Elec. & Comp. Eng.)

Dr. G. C. Shojja, Outside Member (Department of Computer Science)

Dr. E. R. Fossum, External Examiner (JPL, California Inst. of Tech.)

©LIDONG CHEN, 1995

University of Victoria

*All rights reserved. This dissertation may not be reproduced in whole or in part,
by photocopy or other means, without the permission of the author.*

Supervisor: Dr. Harry H. L. Kwok

Abstract

The research work in this thesis can be divided into two parts. The first part was on the modeling and design of high-speed GaAs CMCCDs (cermet-gate charge-coupled devices), which is a part of a research collaboration between TRIUMF and the University of Victoria. The second part was a project on the design and the implementation of a prototype GaAs artificial neural network (ANN) IC using CMCCDs and MESFET technologies.

The research on high-speed GaAs CMCCDs in this thesis primarily focused on the modeling of the charge transfer properties of the device, as well as an attempt to develop a design methodology to optimize the device structure. On the modeling side, we have developed two-dimensional numerical models such that they will allow us to compute the channel potential distribution and the charge transfer efficiency for different device geometries and clocking schemes. In addition, an equivalent circuit model was developed and it allowed us to more efficiently (than the numerical model) study the transient effects of the clock waveforms using a SPICE-type simulator. Using our numerical device models and an optimization algorithm, we have also developed a design method that will allow us to optimize the device geometry for both the two-phase and the uni-phase CMCCDs. For the particular uni-phase CMCCD that is of interest to us and TRIUMF, our theoretical analysis has predicted a device performance that is in close agreement with what we measured on a fabricated CMCCD.

The second phase of our research was on the design of a prototype ANN IC using the GaAs CMCCDs and MESFET circuits. The CMCCDs were to be used as the analog storage elements of the synaptic weights as well as the binary shift-register in

the ANN circuit. For these purposes, we have tested extensively the signal-to-noise ratio and the linearity of the CMCCDs. The ANN circuit was designed based on the popular Hopfield model and can be used as an associative memory. We have developed a hybrid ANN architecture to reduce the number of the components that is required. In this architecture, each synapse is made up of a CMCCD (for the weight storage) and a transconductance amplifier (for the multiplication). The result of the neuron weight-summation is multiplexed by an activation circuit that performs a hard-limiting function. The feedback to the synapses was achieved using shift- and buffer-registers. An ANN circuit with 16 neurons and several subcircuits were fabricated using the NT/BNR 0.8 μm GaAs depletion-mode MESFET technology and this technology unfortunately is not able to support the fabrication of the CCDs. A multi-chip approach was therefore taken and the CMCCDs were fabricated using the TRIUMF GaAs technology. The individual subcircuits and CMCCDs were tested at frequencies up to 200 MHz. The entire system, however, was only tested up to 40 MHz due to the testing environment. Nevertheless, a very good agreement between measurements and simulations was observed. When it was tested as an associative memory, the ANN IC appears to be fairly robust since it can withstand an error rate up to 25%.

Examiners:

Dr. H. H. L. Kwok, Supervisor (Department of Elec. & Comp. Eng.)

Dr. F. El-Guibaly, Departmental Member (Department of Elec. & Comp. Eng.)

Dr. J. M.-S. Kim, Departmental Member (Department of Elec. & Comp. Eng.)

Dr. G. C. Shoja, Outside Member (Department of Computer Science)

Dr. E. R. Fossum, External Examiner (JPL, California Inst. of Tech.)

Contents

Abstract	ii
Contents	x
List of Figures	xi
List of Tables	xviii
Acknowledgements	xx
Dedication	xxi
1 Introduction	1
1.1 GaAs Semiconductor	1
1.2 GaAs Charge-coupled Devices	3
1.3 CCD's Application in Artificial Neural Networks	6
1.4 Motivations and Contributions of the Thesis	9
1.4.1 Motivations	9
1.4.2 Contributions	11

2	Operation and Modeling of GaAs CMCCD	13
2.1	Introduction	13
2.2	Operation of a Uni-phase CMCCD	14
2.3	CMCCD Modeling	19
2.4	Two-dimensional Model for Channel Potential in CMCCD	21
2.5	Two-dimensional Model for Signal Charge Transfer in CMCCD	24
2.6	Equivalent-circuit Model for CMCCD	27
3	Design Optimization and Simulation of GaAs CMCCDs	31
3.1	Introduction	31
3.2	Two-phase CMCCD Design Optimization	32
3.2.1	Device Structure	32
3.2.2	Optimization of the Two-phase CMCCD Geometry	33
3.3	Uni-phase CMCCD Design Optimization	39
3.3.1	Device Structure	39
3.3.2	Optimization of the Uni-phase CMCCD Geometry	41
3.4	Simulations of Charge Transfer Process	42
3.4.1	CTE versus Transfer Time	43
3.4.2	CTE versus Size of Charge Packet	44
3.4.3	CTE versus Amplitude of Clock Voltage	48
3.5	Simulations of the Effect of Clock Waveforms on CTE	49
3.5.1	CTE versus Rise Time	50

<i>CONTENTS</i>	vii
3.5.2 CTE versus Delay Time	52
3.5.3 CTE versus Fall Time	53
4 Fabrication and Measurement of GaAs CMCCD	55
4.1 Introduction	55
4.2 CMCCD Fabrication	55
4.3 CMCCD Test Setup	58
4.4 Experimental Results	60
5 Hybrid Architecture for ANN Implementation	65
5.1 Introduction	65
5.2 Concept of ANNs	66
5.3 Mathematical Description of the Hopfield ANN	69
5.4 Development of A Hybrid ANN Architecture	71
5.5 Overview of the Hybrid ANN Architecture	74
6 GaAs Analog Circuit Design of Hybrid ANN Architecture	79
6.1 Introduction	79
6.2 CMCCD Analog Memory	80
6.2.1 Structure and Operation	80
6.2.2 Performance Evaluation	81
6.2.3 Experimental Results	85
6.3 Weight-summation Circuit	89

<i>CONTENTS</i>	viii
6.3.1 Circuit Design	89
6.3.2 Performance Evaluation	92
6.3.3 Experimental Results	97
6.4 Activation Circuit	100
6.4.1 Circuit Design	100
6.4.2 Experimental Results	104
7 GaAs Digital Circuit Design of Hybrid ANN Architecture	108
7.1 Introduction	108
7.2 MESFET Shift- and Buffer-registers	108
7.2.1 Dual-inverter	108
7.2.2 MESFET Shift-register	114
7.2.3 MESFET Buffer-register	115
7.3 CMCCD Shift- and Buffer-registers	118
7.3.1 Circuit Structure	118
7.3.2 Operation of the CMCCD Shift- and Buffer-registers	120
7.4 Discussions on Power and Area Consumptions of the Shift- and Buffer-registers	122
8 Fabrication and Measurement of ANN IC	125
8.1 Introduction	125
8.2 Simulation, Physical Layout and Fabrication	127
8.2.1 Simulation	127

<i>CONTENTS</i>	ix
8.2.2 Physical Layout	135
8.2.3 Fabrication	136
8.3 Measurement of the ANN IC for Open-loop Operation	136
8.4 Measurement of the ANN IC as An Associative Memory	142
8.5 Discussions	149
8.5.1 Fabrication	149
8.5.2 Linearity of the Weight-summation Operation	150
8.5.3 Glitches in the Output Waveforms of the ANN IC	151
8.5.4 Limitations on High-frequency Measurement	151
9 Summary and Recommendations	153
9.1 Summary of the Results	153
9.2 Recommendations for Future Work	156
9.2.1 ANN Circuit Design	156
9.2.2 Synaptic Weight Loading and Learning	157
9.2.3 Large Scale Integration of the ANN	158
A Derivation of the Electric Field in CMCCD	168
B Equivalent Circuit Model of CMCCD in SPICE3e2	169
C Fletcher-Reeves Method	171

<i>CONTENTS</i>	x
D Noise Analysis of CMCCD	173
D.1 Thermal Generation Noise	173
D.2 Bulk Trapping Noise	174
D.3 Input and Output Noise	174
D.4 Signal-to-Noise Ratio	176
E Stability Analysis of the Activation Circuit	178
F NT/BNR 0.8 μm GaAs Models	180
G NT/BNR 0.8 μm GaAs Process	181

List of Figures

1.1	Basic GaAs MESFET structure.	2
1.2	Basic GaAs CMCCD structure.	5
2.1	Cross-sectional view of a uni-phase GaAs CMCCD and the signals applied to the device nodes.	15
2.2	Potential and charge distribution of the CMCCD for charge injection process.	16
2.3	Potential and charge distribution of the CMCCD for charge transfer process.	17
2.4	Potential and charge distribution of the CMCCD for charge detection process.	18
2.5	A single pixel used for the two-phase CMCCD modeling.	21
2.6	Two-dimensional abrupt doping distribution under the CMCCD transport electrode.	22
2.7	Mesh points for the finite-difference approximation.	23
2.8	Equivalent circuit of unit cell of the CMCCD.	28
2.9	Equivalent circuit of charge-dependent capacitor.	30

LIST OF FIGURES

3.1	Cross-sectional view of a single pixel of a two-phase CMCCD.	32
3.2	Ideal maximum potential profile of the two-phase CMCCD along the CCD channel.	34
3.3	Flow chart of the optimization.	36
3.4	Two-dimensional potential profiles of (a) the optimized two-phase CM- CCD and (b) the CMCCD without partially-embedded cermet-gates.	37
3.5	Maximum potential profiles of the optimized two-phase CMCCD with and without the partially-embedded cermet-gates.	38
3.6	Cross-sectional view of a single pixel of a uni-phase GaAs CMCCD.	39
3.7	Ideal maximum potential profile of the uni-phase CMCCD along the CCD channel.	40
3.8	Maximum potential profiles of the optimized uni-phase CMCCD.	42
3.9	Time evolution of a two-dimensional charge packet in the two-phase CMCCD. (a) $t=200$ ps, (b) $t=400$ ps, and (c) $t=600$ ps.	45
3.10	Time evolution of a one-dimensional charge packet in the uni-phase CMCCD.	46
3.11	CTE versus the the size of the charge packet.	47
3.12	CTE versus the peak-to-peak amplitude of the clock voltage.	48
3.13	Waveforms of the three-phase three-level stepped clocks.	49
3.14	CTE versus transfer time for different values of the rise time. Fall time is 10 ps.	51
3.15	CTE versus transfer time for different values of the delay time. Rise time is 10 ps.	52

LIST OF FIGURES

3.16 CTE versus transfer time for different values of the fall time. Rise time is 10 ps.	54
4.1 Photomicrograph of the CMCCD.	57
4.2 Block diagram of the CMCCD test setup.	58
4.3 <i>I-V</i> characteristics of the CMCCD.	60
4.4 Oscillograph of the CMCCD for 10 MHz operation.	61
4.5 Oscillograph of the CMCCD for 50 MHz operation.	61
4.6 More detailed quantitative demonstration of the performance of the CMCCD.	62
4.7 Typical amplitude response of the CMCCD.	63
4.8 Measurement of the CTE with respect to the normalized leading edge of the clock pulses.	64
5.1 Models of (a) a biological neuron and (b) an artificial neuron.	66
5.2 Hopfield ANN.	68
5.3 Parallel architecture of the ANN.	72
5.4 Serial architecture of the ANN.	73
5.5 Hybrid architecture of the ANN.	74
5.6 Block diagram of the proposed hybrid ANN architecture.	75
5.7 Timing diagram of the neuron states updating.	77
6.1 CMCCD analog memory.	80
6.2 Timing diagram of the CMCCD analog memory.	81

LIST OF FIGURES

xiv

6.3	V_{out} v.s. V_{in} of the CMCCD for different width of MESFET in the output buffer.	84
6.4	Test setup of the CMCCD analog memory.	85
6.5	Oscillograph of the waveforms of the CMCCD analog memory.	86
6.6	Signal-to-noise ratio of the CMCCD with respect to the clock frequency obtained from the measurement and the calculation.	86
6.7	Oscillograph of the CMCCD for the linearity measurement.	87
6.8	Measured linearity between the input and output of the CMCCD.	88
6.9	Block diagram of the weight-summation circuit.	89
6.10	Schematic of the transconductance amplifier.	90
6.11	I - V characteristics of the transconductance amplifier.	92
6.12	Waveforms of W_i and V_i used in the WSC simulation.	93
6.13	Simulation waveforms of the output currents of the WSC.	94
6.14	Simulation waveforms of the output currents of the WSC when the load MESFETs are not the same.	95
6.15	Output voltage of the activation circuit with respect to the number of the transconductance amplifiers.	96
6.16	Microphotograph of one transconductance amplifiers in the WSC.	97
6.17	Measured and simulated linearity of the WSC for binary weights.	98
6.17	Measured and simulated linearity of the WSC for binary weights (continued).	99
6.18	Schematic of the activation circuit.	101
6.19	(a) MESFET cascode circuit and (b) its small-signal equivalent circuit.	102

LIST OF FIGURES

6.20	Microphotograph of the activation circuit.	104
6.21	Oscillographs of the input and the output waveforms of the activation circuit at (a) 20 MHz and (b) 100 MHz.	105
6.22	Transfer characteristics of the activation circuit obtained from the measurement and the simulation.	105
6.23	Amplitude response of the activation circuit obtained from the measurement and the simulation	106
7.1	Schematic diagram of the dual-inverter.	109
7.2	Microphotograph of the dual-inverter.	109
7.3	Inverter chain used for DC analysis of the inverter.	110
7.4	Transfer characteristic of the dual-inverter.	111
7.5	Transient response of the dual-inverter.	113
7.6	Oscillograph of the input (upper trace) and output (lower trace) waveforms of the dual-inverter.	113
7.7	Diagram of the shift-register.	114
7.8	Oscillograph of the input and the output waveforms of the shift-register.	114
7.9	Diagram of the buffer-register.	115
7.10	Schematic diagram of the XOR gate.	116
7.11	Simulation waveforms of the input and the output of the XOR gate.	117
7.12	Diagram of the CMCCD shift- and buffer-registers.	119
7.13	Schematic of the output buffer.	120
7.14	Timing diagram of the CMCCD shift- and buffer-registers.	121

LIST OF FIGURES

xvi

7.15 On-chip power dissipation versus clock frequency for the two-phase and the uni-phase CMCCDs.	124
8.1 Schematic of the ANN chip.	126
8.2 Two 16-bit test patterns stored in the synapses.	127
8.3 One of the input patterns applied to the circuit.	129
8.4 Simulation waveforms of the 16-bit outputs of the ANN IC.	130
8.4 Simulation waveforms of the 16-bit outputs of the ANN IC (continued).	131
8.4 Simulation waveforms of the 16-bit outputs of the ANN IC (continued).	132
8.5 Simulation waveforms of the clocks ϕ_1 and ϕ_2 and the output of the activation circuit corresponding to the first stored pattern.	133
8.6 Simulation waveforms of the clocks ϕ_1 and ϕ_2 and the output of the activation circuit corresponding to the second stored pattern.	134
8.7 Microphotograph of the ANN chip.	135
8.8 Output characteristic of one neuron.	137
8.9 Measured linearity of the weight-summation computation for analog weights.	138
8.10 Input waveforms used in the test and the theoretical output waveforms of the ANN IC for the open-loop operation.	140
8.11 Oscillographs of the open-loop measurement.	141
8.12 Schematic of the circuit used to generate the three-level synaptic weights.	143
8.13 Oscillographs of the 16 weight signals.	144
8.14 Correct recall rate versus bit-error rate.	145

LIST OF FIGURES

xvii

8.15	Input patterns used in the ANN chip test.	147
8.16	Output patterns from the ANN chip.	147
8.17	Oscillographs of the clocks ϕ_1 and ϕ_2 , and the output of the chip for the different input patterns.	148
D.1	Input and output sections of the CCD.	175
E.1	Small-signal equivalent circuit for a two-stage operational amplifier. .	179

List of Tables

2.1	Physical constants used in the charge transfer modeling.	25
3.1	Optimization results of the two-phase CMCCD.	35
3.2	Optimization results of the uni-phase CMCCD.	41
3.3	Variation of the CTE with respect to the transfer time in the two-phase CMCCD.	43
3.4	Variation of the CTE with respect to the transfer time in the uni-phase CMCCD.	43
3.5	Average values of the fringing field under the emptying electrode for the different rise time.	51
3.6	Average values of the fringing field under the emptying electrode for the different delay time.	53
3.7	Average values of the fringing fields under the emptying electrode for the different fall time.	54
4.1	Main process parameters of the TRIUMF GaAs technology.	56
4.2	Geometric parameters of the CMCCD.	57

LIST OF TABLES

xix

4.3	Signal levels and dc bias voltages using for testing the CMCCD.	59
4.4	Measured CTEs at different clock frequencies.	63
6.1	Parameters used in computing CMCCD linearity.	83
6.2	Performance of the activation circuit.	107
7.1	DC performance of the dual-inverter.	112
7.2	Logic function of the XOR gate.	116
8.1	Coding of the three-level weights.	142
8.2	Percentage of the power consumed by the subcircuits of the ANN chip.	146
8.3	Main process parameters of the NT/BNR GaAs technology.	149
D.1	Noise and S/N of the two-phase and the uni-phase CMCCDs	177

Acknowledgements

I would like first to express my deepest thanks to my supervisor, Dr. Harry H. L. Kwok of the Department of Electrical and Computer Engineering, for introducing me to GaAs IC design, and for his generous encouragement, guidance and advice throughout this research, and in the preparation of this thesis.

I would also like to thank Dr. F. El-Guibaly, Dr. J. M. S. Kim, and Dr. G. C. Shoja, for their valuable comments in the course of this work. The graduate courses on semiconductor devices, analog neural network, VLSI, and optimization taught by Dr. Kwok, Dr. El-Guibaly, and Dr. Lu gave me a solid foundation. Thanks must to Ms. Vicky Smith, Graduate Secretary of the Department, for making sure that the program was going through right bureaucratic procedure.

There are many people to whom I am indebted for their help in dealing with the many technical aspects of this work: M. Wedlake, G. Deliyannides, A. Keddy, S. Pennathur, M. LeNoble, J. Cresswell and K. de Solla. As well I would like to thank K. Jones and P. Fedrigos, technical staff in the Department, for their assistance and technical support in chip testing.

Chip fabrications and CAD tools from TRIUMF, the Canadian Microelectronics Corporation, and NT/BNR are gratefully acknowledged.

My sincere thanks to Dr. Kwok for the financial support I received for assisting on research projects funded by NSERC and Micronet.

Finally, to my beautiful wife, Yanmei, for having the patience to endure and support me in spirit — I express my love and appreciation.

Dedication

To Yanmei

Abbreviations

ANN	Artificial neural network
B/D	Buffer/drain
B/O	Buffer/output
B/S	Buffer/source
BFL	Buffered FET logic
CCD	Charge-coupled device
CGCCD	Capacitance-gate charge-coupled device
CID	Charge injection device
CMC	Canadian Microelectronics Cooperation
CMCCD	Cermet-gate charge-coupled device
CMOS	Complementary metal-oxide semiconductor
CTE	Charge transfer efficiency
EPROM	Erasable programmable read only memory
GaAs	Gallium Arsenide
I/P	Input
IC	Integrated circuit
LSB	Less significant bit
MDAC	Multiplying D-to-A conversion
MESFET	Metal-semiconductor field-effect transistor
MNOS	Metal nitride oxide silicon

MOSFET	Metal-oxide-semiconductor field-effect transistor
MSB	Most significant bit
NT/BNR	Northern Telecom/Bell Northern Research
O/P	Output
PGA	Pins grid array
R/D	Reset/drain
R/G	Reset/gate
S/R	Signal-to-noise ratio
WSC	Weight-summation circuit
TRIUMF	Tri-University Meson Facility
VLSI	Very-large-scale integration
XOR	exclusive OR gate

Chapter 1

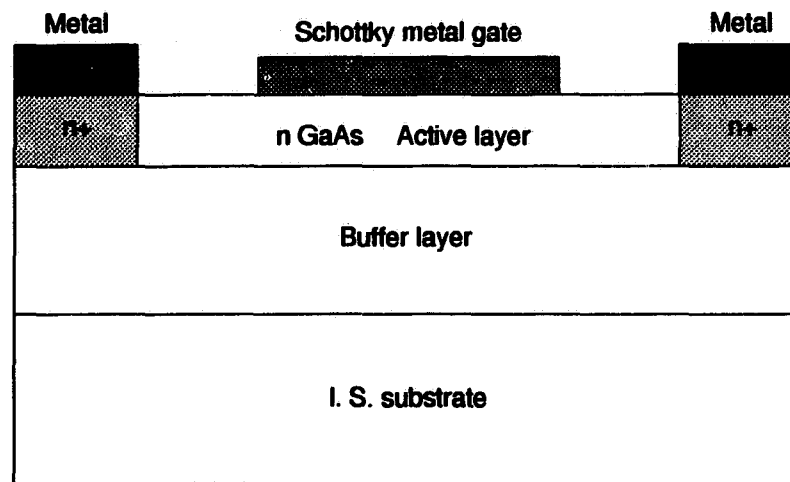
Introduction

1.1 GaAs Semiconductor

Gallium Arsenide (GaAs) is a III-V compound semiconductor particularly suitable for making high frequency devices. As far as the electrical properties are concerned, GaAs has several superior properties. One is the very high intrinsic electron mobility which is almost six times that of silicon. High electron mobility and high saturation velocity are usually required for high frequency device operation. The second superior property related to GaAs is the possibility to grow high resistivity semi-insulating GaAs. Semi-insulating GaAs is frequently used as the device substrate to produce a very low parasitic capacitance. A low parasitic capacitance also gives a low ac power dissipation. The third superior property of GaAs is the existence of a direct band-gap which is essential for radiative recombination to occur. Direct band-gap also makes GaAs a very efficient light emitter in opto-electronic applications.

The first metal-semiconductor field-effect transistor (MESFET) using an epitaxial layer of GaAs on a semi-insulating GaAs substrate was fabricated by Hooper and Lenner in 1967 [1]. It was only in the middle of the 1980's that GaAs MESFET

technology reached the point where IC production was contemplated at the industrial level. Nowadays MESFET is the predominant device in the design of GaAs ICs. Fig. 1.1 shows the basic GaAs MESFET structure and the typical process parameters. The GaAs MESFET shown has a n-type doped active layer (or channel) and its thickness is controlled by the depletion of the metal-semiconductor Schottky barrier. As observed, the gate metal (Schottky metal) is directly in contact with the channel and the drain region and the source region are highly-doped n^+ regions. The metal electrodes on top of the drain and the source form the ohmic contacts. The I - V characteristics of a GaAs MESFET are quite similar to those of the silicon MOSFET, although the superior properties of GaAs combined with the removal of the oxide layer and a low threshold voltage give the GaAs MESFET marked advantages in speed, power consumption and radiation-tolerance over the silicon MOSFET.



Thickness of active layer : $0.1 \sim 0.2 \mu\text{m}$

Doping density of active layer : $> 10^{18}/\text{cm}^3$

Figure 1.1: Basic GaAs MESFET structure.

1.2 GaAs Charge-coupled Devices

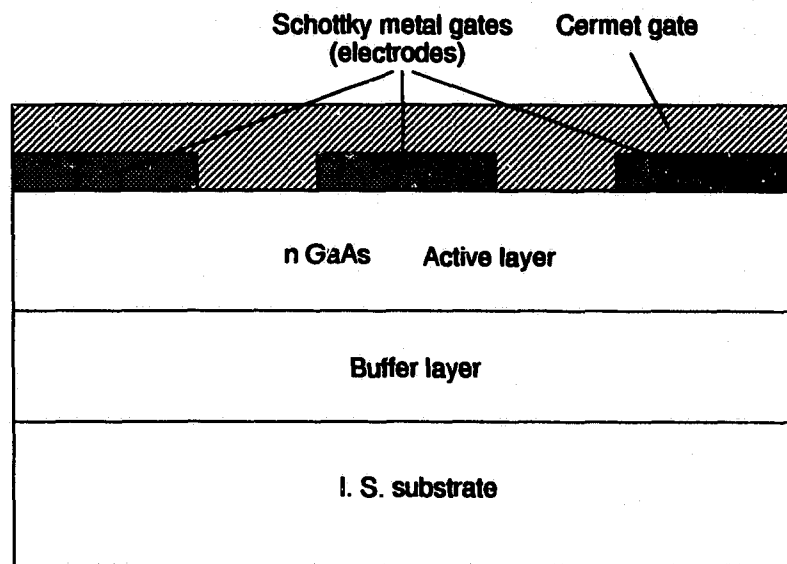
The concept of a charge-coupled device (CCD) was first proposed in 1970 by Boyle and Smith at Bell Laboratories [2, 3]. Basically, a CCD is a shift register formed by a string of transport electrodes (for simplicity, we shall call the transport electrodes as the electrodes). An analog input signal introduced electrically or optically into the CCD can be stored in the form of charge packets in the potential wells under the electrodes. A set of clocks applied to the electrodes will transfer the charge packets from under one electrode to the next. The charge packets are then detected at the output after they have sequentially passed through all the electrodes. Initially, CCDs were realized by MOS technology. In MOS CCDs, the electrodes are insulated from the semiconductor by a thin oxide layer and each electrode is, in effect, a metal-insulator-semiconductor (MIS) capacitor. Alternatively, CCDs can be fabricated using GaAs technology. The electrodes of a GaAs CCD are reverse-biased metal-semiconductor Schottky diodes instead of MIS capacitors. It has been observed that a GaAs CCD inherently has low noise and a wide dynamic range due to the fact that the signal charge packets are stored and transferred in isolation in the depleted channel of the CCD.

The first proposal of a GaAs CCD was made in 1972 by Schuermeyer *et al.* [4]. A Schottky-gate GaAs CCD employing Schottky diodes separated by approximately 1 micron wide dielectric-filled gaps was fabricated in 1977 by Kellner *et al.* [5] and in 1978 by Deyhimi *et al.* [6]. This kind of device is known as capacitive-gate CCD (CGCCD) as the GaAs surface in the interelectrode gap is covered with a dielectric material. GaAs CGCCDs operating with one-phase, two-phase, three-phase and four-phase clocking schemes have since been reported and they were found to operate in the clock frequency range from 500 MHz to 1 GHz with a charge transfer efficiency up to 0.999 [7, 8, 9, 10, 11]. In comparison, the highest clock frequency used in a silicon CCD was 180 MHz as reported by Esser and Sangster [12]. This level of performance

allows the GaAs CCDs to realize specialized analog signal processing functions at a high-speed with a reduced area and power consumption [13, 14, 15].

The GaAs CGCCD, however, sometimes suffers from a low charge transfer efficiency due to potential troughs that are formed in the interelectrode gaps [16]. Potential troughs between the adjacent potential wells will retain some of the passing electrons during the charge transfer. These electrons, at a later time, will be transferred to the output resulting in an increase in the signal dispersion. To achieve a high transfer efficiency, most CGCCDs will need submicron gaps. This oftentimes requires more sophisticated processing steps and can cause the devices to easily break-down. [16].

The requirement for a submicron gap can be avoided if a resistive-gate CCD is used. This was demonstrated by Higgins, *et al.* [17, 18], and Song and Fossum [19, 20, 22]. In a resistive-gate CCD, the GaAs surface in the interelectrode gaps is covered with a resistive material. The resistive-gate CCD is also known as cermet-gate CCD (CMCCD) as cermet has been used very frequently as the resistive-gate material in GaAs CCDs. Cermet is an insulator-metal composite containing SiO and Cr. It has a high resistance of up to mega-ohms per square and makes a low-leakage Schottky contact to the GaAs. Fig. 1.2 shows the GaAs CMCCD structure and the typical process parameters. As observed, the GaAs surface in the interelectrode gap is covered with a cermet film. The cermet film provides two important functions. Firstly, it ensures that the variation of the surface potential along the interelectrode gaps is monotonic at all clock frequencies [23]. This allows wider interelectrode gaps to be used and avoids the formation of potential troughs. The relatively short electrode also provides an increased electric field component along the direction of charge transfer under the electrode and this improves the charge transfer [17]. Secondly, since a thicker active layer is no longer needed for the removal of the potential troughs [15], a thinner active layer can be used in the CMCCD, thus making it



Thickness of active layer : $0.3 \sim 1 \mu\text{m}$
 Doping density of active layer : $10^{15} \sim 10^{17} / \text{cm}^3$

Figure 1.2: Basic GaAs CMCCD structure.

possible for a MESFET-compatible technology to be developed. The first four-phase GaAs CMCCD was demonstrated in 1982 by Higgins *et al.* at a clock frequency of 1 GHz with a charge transfer efficiency of 0.999 [17]. In 1990, LeNoble *et al.* reported a two-phase GaAs CMCCD with $2 \mu\text{m}$ electrodes and $3 \mu\text{m}$ gaps. The device used a castellated structure and was reported to operate at 46 MHz with a transfer efficiency of 0.996 [24]. In 1991 Song *et al.* demonstrated a four-phase GaAs CMCCD with $1 \mu\text{m}$ electrodes and $4 \mu\text{m}$ gaps. The device operated at the clock frequencies from 10 MHz to 1GHz and had a charge transfer efficiency of 0.999 [20].

GaAs CCDs have been successfully used in a number of signal processing applications [13, 14, 26, 27], even though the process parameters are not necessarily compatible with those for GaAs MESFETs. It can be observed from Figs. 1.1 and 1.2 that a GaAs CCD uses a relatively thick, low-doped active layers in order to avoid potential troughs [9] while a GaAs MESFET really requires a thin, highly-

doped active layer in order to obtain a high transconductance. Much effort has been made to integrate GaAs CCDs monolithically with GaAs MESFETs by using cermet gates and compromising process parameters. A GaAs VHF/UHF transversal filter was fabricated by Sovero *et al.* at Rockwell [26]. The thickness of the active layer was approximately $0.3 \mu\text{m}$. This filter used GaAs CMCCDs arranged in a pipeorgan structure to provide weighted sampling, programmable delay and summation of the analog signals. Supervisory functions were provided by the GaAs MESFET circuits integrated monolithically with the CMCCDs. The CMCCD transversal filter demonstrated a dynamic range of 60 dB for filter operation at a sampling rate of 1 GHz [27]. The two-phase CMCCD reported in [24] was also fabricated at TRIUMF using a GaAs MESFET compatible CMCCD process. The thickness of active layer was also $0.3 \mu\text{m}$. The TRIUMF process, however, was optimized for the CMCCD but not for the MESFET circuits.

1.3 CCD's Application in Artificial Neural Networks

A wide variety of technologies and design approaches has recently been investigated for VLSI implementation of artificial neural networks (ANNs) [28, 29, 30, 31]. It has been demonstrated that analog circuits are probably more suitable for the VLSI implementation of ANNs due to the similarity between the analog ANN architecture and the biological neural system, and the simplicity in the implementation [32, 33]. CCD technology has also been highlighted for its potential in realizing analog ANNs due to the low power consumption and the small chip area as well as the ability of storing and transferring analog signals [34, 35, 36, 37].

ANNs, in general, attempt to mimic, at least partially, the structures and the

functions of the human brain and the nervous system. In 1982, Hopfield proposed a simplified model of a neural network [38], which consists of an array of fully interconnected synapses and neurons that are connected to each other by the synapses. Each synapse will give a new output by multiplying the current neuron state to the connection weight stored in the synapse. The outputs of the synapses connected to the same neuron are then summed. This summation determines the neuron activity and changes the neuron output after it has passed through an activation function [39].

In the analog implementation of the Hopfield ANNs, the activation function of a neuron is frequently represented by an amplifier and a synapse by a multiplier. Since the neural network contains a large number of synapses, one of the essential problems in the implementation is to develop the storage for the analog synaptic weights. In most applications, the synaptic weights are required to be programmable and adjustable for the adaptive learning, and they can be stored in either an analog or a digital form.

When the synaptic weights are stored in a digital memory, a D-to-A converter is required for each synapse. CCDs have been successfully used to implement multiplying D-to-A conversion (MDAC) in the charge domain. A CCD programmable signal processor for computing vector-matrix products based on charge domain MDAC was proposed in 1990 by Chiang at MIT [34]. For neural network applications, the processor computes in a pipeline fashion the weight-summation required by the neural network. Using a $2\ \mu\text{m}$ CCD rule, the 8-bit CCD MDAC occupied a $200 \times 220\ \mu\text{m}^2$ chip area. Another CCD MDAC-based signal processor using the wire transfer technique [40] was reported in 1991 by Fossum *et al.* at Columbia University [41]. The size of the 10-bit pipeline MDAC is $130 \times 300\ \mu\text{m}^2$ for a $2\ \mu\text{m}$ process.

The chip area used for the storage of the synaptic weights can be reduced significantly if the weights are stored in analog form. A common storage technique is to store a weight value as a charge packet in a capacitor [43]. However, such a dynamic

storage technique requires refreshing due to charge decay, and refreshing analog signals also requires an elaborate support circuitry that may offset a lot of the advantage gained in the smaller size of the storage cell. Floating-gate transistor can be used as a programmable non-volatile analog memory [44, 45]. The weight value is set by programming the threshold voltage of the floating-gate transistor, and it is possible that both the weight storage and the analog multiplication can be implemented simultaneously using the same circuit [46]. Special fabrication techniques such as an ultra thin tunneling oxide layer and the use of a high voltage to excite the electrons across the gate oxide barrier represent the major drawbacks of the floating-gate analog memory.

A neural network IC combining MNOS devices for the storage of the synaptic weights with CCD technology was reported in 1989 by Sage *et al.* at MIT [35], and a similar IC was reported in 1991 by Neugebauer *et al.* at Caltech [47]. In these ICs, CCD technology was used to fabricate the matrix of charge injection device (CID) elements which store the charge packets encoding the synaptic weights. These ICs compute the fully parallel vector-matrix products and the typical size of a storage cell is $25 \times 25 \mu\text{m}^2$.

A conventional CCD behaves like a serial analog memory and it is possible to load or program the synaptic weights into a CCD either electrically or optically and read them serially. A semi-parallel CCD neural processor computing vector-matrix products was proposed in 1990 by Agranat *et al.* at Caltech [37]. The neural processor contains 256 neurons and 65536 programmable analog synapses. The vector-matrix products are computed by a simultaneous combination of the horizontal and the vertical shifting of the charge packets in the CCD array.

1.4 Motivations and Contributions of the Thesis

1.4.1 Motivations

We have discussed the potential of CCDs for high-speed analog signal processing and ANN implementations. The purpose of this research work is to design high performance GaAs CMCCDs and to use them with GaAs MESFET circuits in the design of an ANN chip functioning as an associative memory.

When the CMCCD is used as a component in analog signal processing, the main considerations in the CMCCD design are focused on the charge transfer efficiency, the operating speed and the charge-handling capacity (to provide a better signal-to-noise ratio). In most cases, however, these parameters are interdependent. For example, the charge-handling capacity as determined by the depth of the potential well can be increased by increasing the thickness of the active layer [25]. However, the depth of the potential well also has some bearing on the formation of the potential trough. This can in turn reduce the charge transfer efficiency. The potential trough is also related to the lengths of the electrode and the gap, which determine the operating speed of the CMCCD (we shall show this in Chapter 3).

A design methodology, therefore, needs to be developed to optimize the CMCCD structure so that a compromise among the charge transfer efficiency, the operating speed and the charge-handling capacity is achieved. We found that there was no such design methodology reported when we started this work. The design methodology should provide great flexibilities in the clocking schemes and the device parameters, and should be efficient in computation. To develop the CMCCD design methodology, numerical models are required to compute the channel potential distribution and the signal charge transfer in the CMCCD [53].

When the CMCCD operates at high frequency, the clock waveforms have finite rise

time and fall time due to the parasitic effects. It has been shown that the effectiveness of the charge transfer is determined by the fringing field, while the clock voltages make an important contribution to the fringing field. Therefore, it is necessary to investigate how the clock waveforms affect the charge transfer efficiency. To effectively investigate this, a standard SPICE-type simulator needs to be used. Again, according to our knowledge, there was no reported work on this research. In addition, an equivalent circuit model of the CMCCD was developed earlier [55] and we could use it to fulfill this investigation.

In most of the ANN implementations, a fully connected parallel architecture has been used in order to make the ANNs most similar to the biological model of the neural network [42]. It should be pointed out, however, that a parallel architecture consumes a large chip area not only for the electronic neural cells but also for the massive interconnections. The latter is limited by the wiring of a large number of interconnections on the two-dimensional surface of the wafer. In fact, a parallel architecture limits the size and complexity of the neural networks, reduces the reliability of the device and suffers from massive power dissipation because N amplifiers and N^2 multipliers are required.

CCDs have been used in neural signal processing with special capabilities of storing analog synaptic weights and reducing massive interconnections. The semi-parallel CCD neural processor proposed in [36] tried to solve the problems suffered from by the parallel architecture. This neural processor itself, however, suffers from the following problems. Firstly, the number of multipliers, adders and activation functions are all equal to the number of neurons (N). This means that the area and the power consumed by the neural cells will not be reduced compared to the parallel structure. Secondly, the charge summation is completed serially in N clock pulses, and this will result in charge decay and consequently will limit the number of the neurons that can be integrated. In addition, the neural processor was only tested using the

binary weights in terms of the vector-matrix product operation. To overcome those problems in [36], we need to develop a new architecture for the ANN implementation. Moreover, the implemented ANN should demonstrate its functionality of associative memory using the analog synaptic weights.

As mentioned in Sections 1.1 and 1.2, GaAs technology is excellent in realizing high-speed electronic circuits for signal processing. It has been noted, however, that almost all of the ANNs implemented so far have used CMOS technology, and there was no report on the GaAs implementation of the ANN according to the author's literature survey. If the ANN is implemented using GaAs technology and the synaptic weights are stored in the GaAs CMCCDs, the high-speed charge transfer in the CMCCDs would reduce the computation time of the ANN and increase the number of the neurons to be integrated.

1.4.2 Contributions

Specific contributions of this thesis are the following:

- Development of two-dimensional numerical models to compute the potential distribution and the signal charge transfer in the CMCCDs.
- Development of a design methodology to optimize the structure of a CMCCD using the Fletcher-Reeves algorithm and the potential distribution model. The method was applied successfully to the design of a two-phase and a uni-phase CMCCDs [59, 60]. The simulations were performed using the charge transfer model and a comparison of the devices was made [61].
- Investigation of the effect of the clock waveforms on the charge transfer efficiency. This has been done successfully by using SPICE3e2 incorporated with the equivalent circuit model and the analytical fringing field expression [62].

- Fabrication (TRIUMF process) and measurement of the uni-phase GaAs CMCCD. The experimental results have verified the simulation results.
- Development of a hybrid architecture for the implementation of the Hopfield ANNs [63]. This approach merges parallel multiplication/summation and serial neuron state updating and only uses one adder and one activation function, thus reducing significantly the chip area, the power dissipation and the complexity of the interconnections compared to the parallel architecture and the semi-parallel architecture proposed in [36].
- Design of the subcircuits of the hybrid ANN architecture using CMCCD and MESFET technologies. We have shown by experiment that the CMCCD serial analog memory can be used for the synaptic weight storage with very low noise, large dynamic range, and low on-chip power dissipation. The experimental results have verified the simulation of the MESFET subcircuits.
- Fabrication (Northern Telecom/Bell Northern Research process) and measurement of a prototype GaAs MESFET ANN IC containing 16 neurons. We have successfully demonstrated, using a multi-chip approach, the functionality of the chip as an associative memory.

Chapter 2

Operation and Modeling of GaAs CMCCD

2.1 Introduction

Functionally, a GaAs CMCCD is similar to an analog shift-register. It has an input section, a string of electrodes, and an output section. In order to transfer charge from under one electrode to the next, a set of clocks applied to the electrodes is required. The clocks can be either four-phase, three-phase, two-phase, or uni-phase. In general, the smaller the number of clock phases, the less the complexity of the clock driving circuit. In this chapter, the operation of a uni-phase CMCCD is described. Two two-dimensional numerical models for computing the channel potential distribution and the signal charge transfer are developed. An equivalent circuit model of CMCCDs is also discussed.

2.2 Operation of a Uni-phase CMCCD

Fig. 2.1 shows a cross-sectional view of a uni-phase GaAs CMCCD together with the signals applied to the device nodes. The CMCCD consists of a n-type active layer (or n-channel) and a semi-insulating GaAs substrate. Initially, the input ohmic contact and the output ohmic contact are both biased positively with respect to the voltages applied to the control gates and the electrodes so that the channel region under the gates G_1 , G_2 , G_3 and the electrodes are all depleted of electrons. The clocks ϕ_1 and ϕ_{1A} have voltage levels of $0/-4$ V and $+1/-3$ V respectively, and *Bias1* and *Bias2* are set at -2 V and -1 V respectively. These voltages are required to transfer signal charge packets in the channel to the device output where they are detected.

The input voltage V_{in} applied to G_1 is converted to a charge packet using the fill-and-spill method which has been widely used due to its advantages of low-noise and better conversion linearity [25]. Fig. 2.2 shows the sequence of events that occurs in the voltage-to-charge conversion by the fill-and-spill method. The different times (i.e., t_1 , t_2 and t_3) appearing in Fig. 2.2 correspond to those appearing in Fig. 2.1. Besides the input voltage V_{in} applied to G_1 , a reference voltage V_{ref} is applied to G_2 , and an input sampling pulse $V_{I/P}$ to the input ohmic contact. At $t = t_1$, $V_{I/P}$ is high and ϕ_1 is low. Therefore, the region in the channel extending from the right-hand edge of the input ohmic contact to the left-hand edge of G_2 is depleted. At $t = t_2$, $V_{I/P}$ is lowered to a voltage between the potentials under G_1 and the first electrode ϕ_1 . This causes the thermally generated electrons to flow into and to fill the potential well under G_2 from the input ohmic contact. Note that the voltage in the first electrode ϕ_1 also provides a barrier to the electrons under G_2 . At $t = t_3$, $V_{I/P}$ is returned to a high voltage and the electrons under G_1 as well as the excess electrons under G_2 is removed from the channel through the input ohmic contact. This creates a well-defined charge packet consisting of electrons under G_2 . It can be shown that

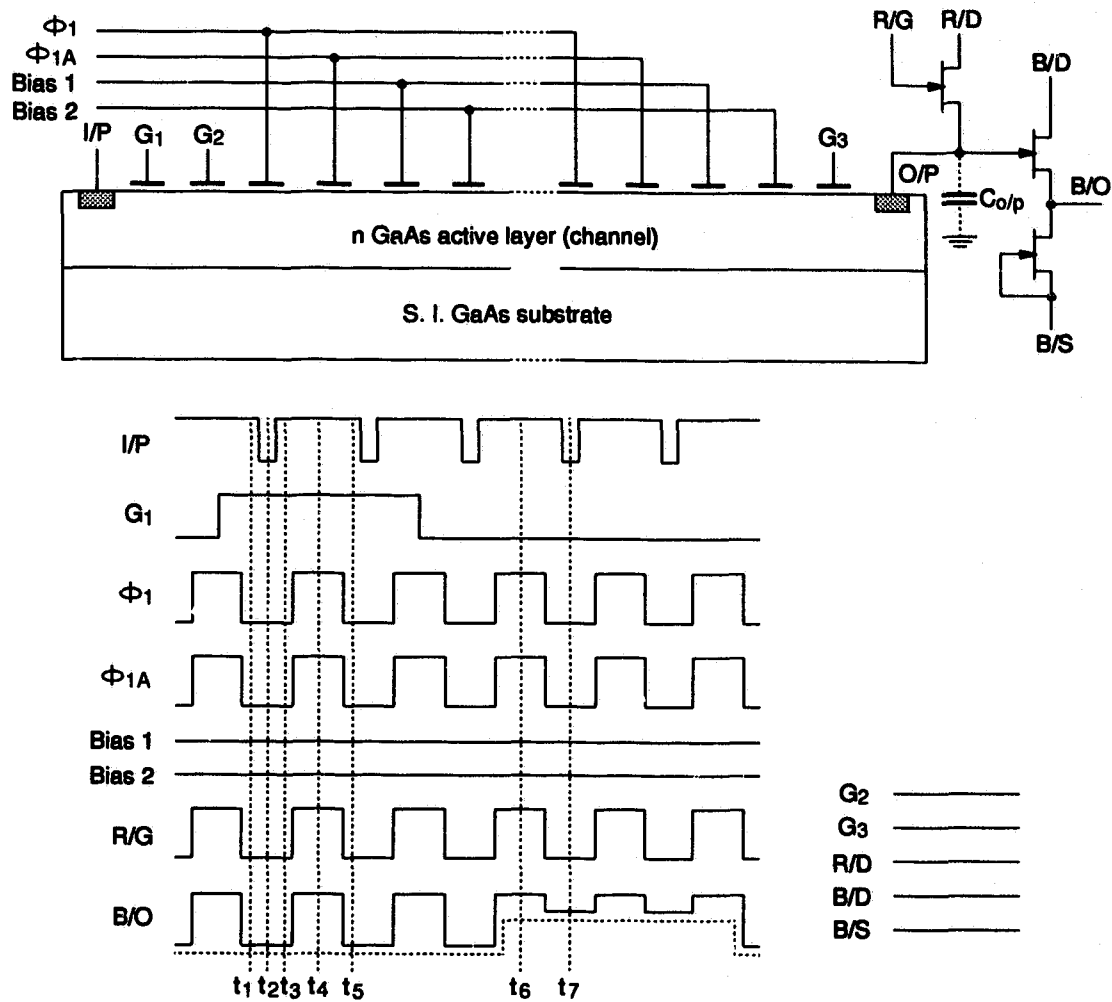


Figure 2.1: Cross-sectional view of a uni-phase GaAs CMCCD and the signals applied to the device nodes.

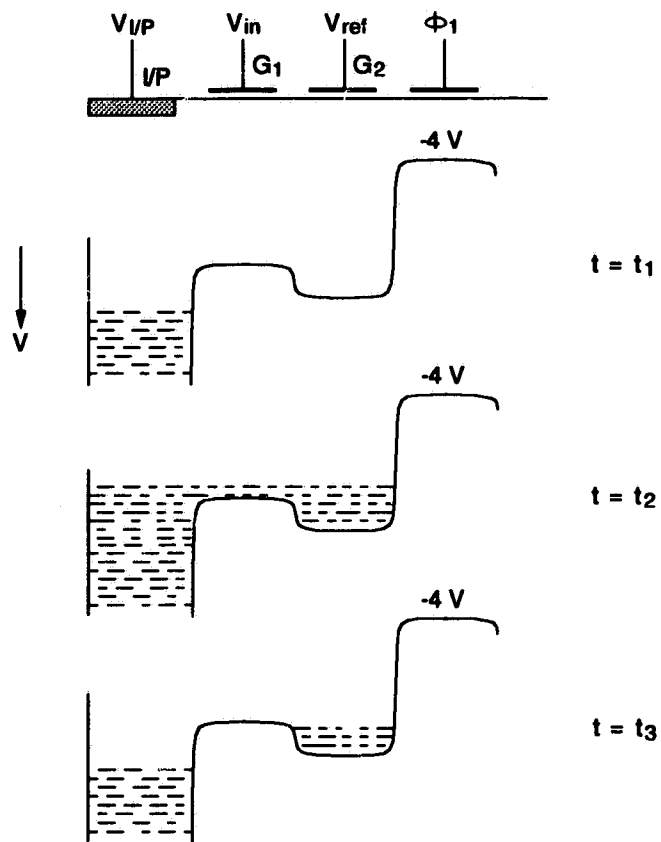


Figure 2.2: Potential and charge distribution of the CMCCD for charge injection process.

the size of the charge packet is proportional to $V_{ref} - V_{in}$ [25].

The charge packet residing in the potential well under G_2 is transferred to the transport region during the positive cycle of ϕ_1 and ϕ_{1A} . Fig. 2.3 shows the actual transfer of a charge packet across one pixel of the CMCCD. At $t = t_4$, the charge packet

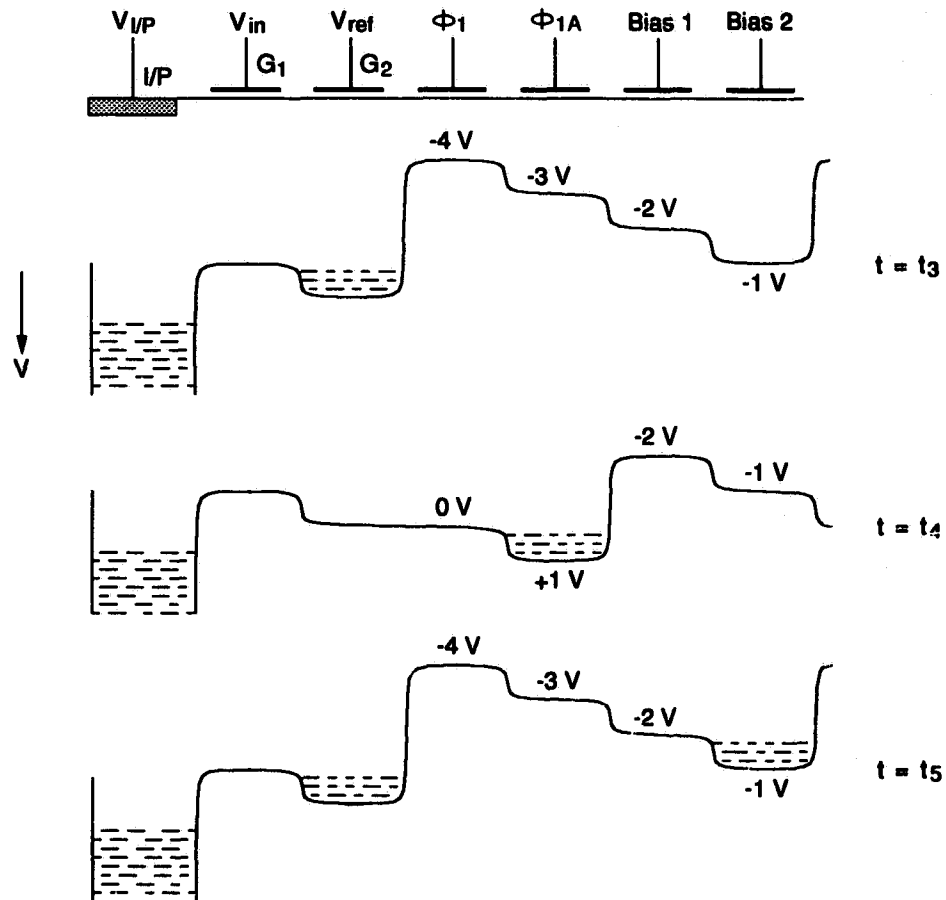


Figure 2.3: Potential and charge distribution of the CMCCD for charge transfer process.

under G_2 moves to the potential well under the electrode ϕ_{1A} where the potential is maximal. At $t = t_5$, both ϕ_1 and ϕ_{1A} are at high voltage levels and the charge packet under the electrode ϕ_{1A} moves to the potential well under the electrode Bias2. As

indicated in Fig. 2.3, the direction of the charge transfer is determined by the direction of the electric field due to the difference in the voltages applied to the adjacent transport electrodes.

A charge packet transferred to the potential well under the final electrode ϕ_{1A} moves into the output ohmic contact during the negative cycle of ϕ_1 and ϕ_{1A} , as shown in Fig. 2.4. Prior to the charge detection, the output ohmic contact of the

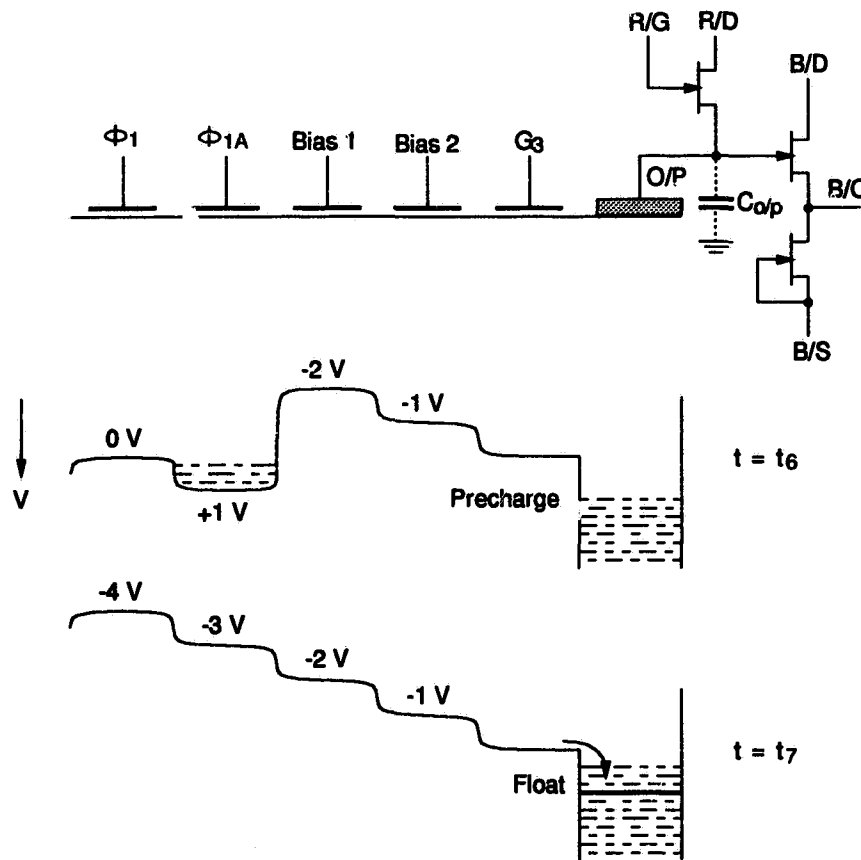


Figure 2.4: Potential and charge distribution of the CMCCD for charge detection process.

CMCCD is first precharged to a high voltage by turning on the reset MESFET using the same voltage pulse as ϕ_1 applied to R/G ($t = t_6$). The reset MESFET is disabled

on the negative cycle of ϕ_1 and ϕ_{1A} . This allows the output ohmic contact to float at its precharged level. At $t = t_7$, the electrons passing through the potential well under G_3 exit the CMCCD through the floating output ohmic contact, charging the parasitic capacitance $C_{O/P}$ and driving the output ohmic contact to a voltage negative with respect to its precharged level. The signal produced at the output ohmic contact is also buffered by a MESFET source follower. As observed from the waveform of B/O in Fig. 2.1, the output of the CMCCD is the delayed input signal modulated by the clock feedthrough.

It should be pointed out that the operations of two-phase, three-phase and four-phase CMCCDs are similar to the operation of the uni-phase CMCCD, and therefore the charge transfer mechanisms and the modeling of the devices with these clocking schemes are essentially the same.

2.3 CMCCD Modeling

It is generally accepted that the operation of a CMCCD can be accurately modeled by the Poisson equation and the continuity equation [25]. To study and design a CMCCD we must first compute the channel potential distribution by solving the Poisson equation. Another reason for computing the potential distribution is to determine the fringing field. The fringing fields under the electrodes are crucial for the charge transfer. As we shall see later, the computation of the fringing fields is quite indispensable if we want to model the signal charge transfer. The fringing field can be obtained by computing the gradient of the channel potential.

To evaluate the performance of a CMCCD, we need to model the signal charge transfer. This can be done by solving the continuity equation. Three types of charge transfer mechanisms are considered, namely, the self-induced drift effect, the drift effect due to the fringing field, and the thermal diffusion of the electrons.

Sometimes a CMCCD needs to be simulated using a standard circuit simulator like SPICE. This necessitates the development of an equivalent circuit model for the GaAs CMCCD. As we shall see in the next chapter, the effect of the clock waveforms on the charge transfer can be efficiently determined by the simulation using SPICE3e2 incorporated with an equivalent circuit model of CMCCDs.

Among the four clocking schemes, the two-phase CMCCD is the only one that requires an asymmetric electrode geometry to achieve unidirectional charge flow, and it also requires an embedded cermet layer. In other words, the structure of the two-phase CMCCD is most complicated. In this work, we have developed the numerical models mainly based on a two-phase CMCCD structure. The corresponding computer programs can also be used for the other three kinds of CMCCDs by changing the geometric parameters of the device. Fig. 2.5 shows a two-dimensional plane of a single pixel used for the two-phase CMCCD modeling. The plane is coincident with the central axis of the CMCCD so that the potential $\psi(x, y)$ and the charge density $n(x, y)$ are considered to be invariant along the axis normal to this plane. Cartesian coordinates are defined as shown with the origin located at the left upper corner.

We assume, for simplicity in computation, that the CMCCD consists of two regions, namely, a n-type GaAs active layer ($0 \leq x \leq x_t$, $0 \leq y \leq y_m$) and a semi-insulating GaAs substrate ($x_t \leq x \leq x_m$, $0 \leq y \leq y_m$). The n-type active layer is uniformly doped and fully depleted. The semi-insulating GaAs substrate is slightly p-type and depleted at its upper surface.

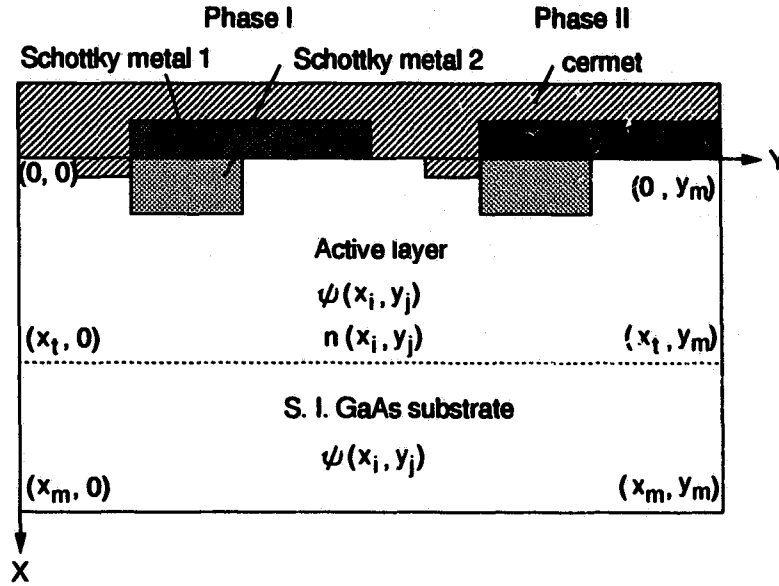


Figure 2.5: A single pixel used for the two-phase CMCCD modeling.

2.4 Two-dimensional Model for Channel Potential in CMCCD

In the absence of any signal charge, the channel potential, $\psi(x, y)$, of a GaAs CMCCD, can be obtained by solving the Poisson equation, i.e.,

$$\nabla^2 \psi(x, y) = -\frac{qN_d}{\epsilon_s} \quad (2.1)$$

in the active layer, and by solving the Laplace equation, i.e.,

$$\nabla^2 \psi(x, y) = 0 \quad (2.2)$$

in the semi-insulating substrate. N_d is the donor density in the active layer, ϵ_s is the permittivity of GaAs, and q is the electron charge. The x direction is chosen to be the depth of the CMCCD, and the y direction is the length of the CMCCD along which the charge packets are transferred. The two-dimensional solution of Eqs. (2.1)

and (2.2) is determined using the abrupt doping approximation shown in Fig. 2.6. The boundary conditions are given by:

1. the potential along the GaAs surface under the transfer electrodes is given by $V_g - V_{bi}$, where V_g is the applied gate voltage and V_{bi} is the built-in voltage;
2. the surface potential in the gap region is given by the linear interpolation of the potentials at the adjacent electrodes [24];
3. the bottom surface of the semi-insulating substrate is assumed to be grounded;
4. periodicity of the channel potential is assumed for every pixel;
5. the potentials and their derivatives are assumed to be continuous across the channel and the substrate.

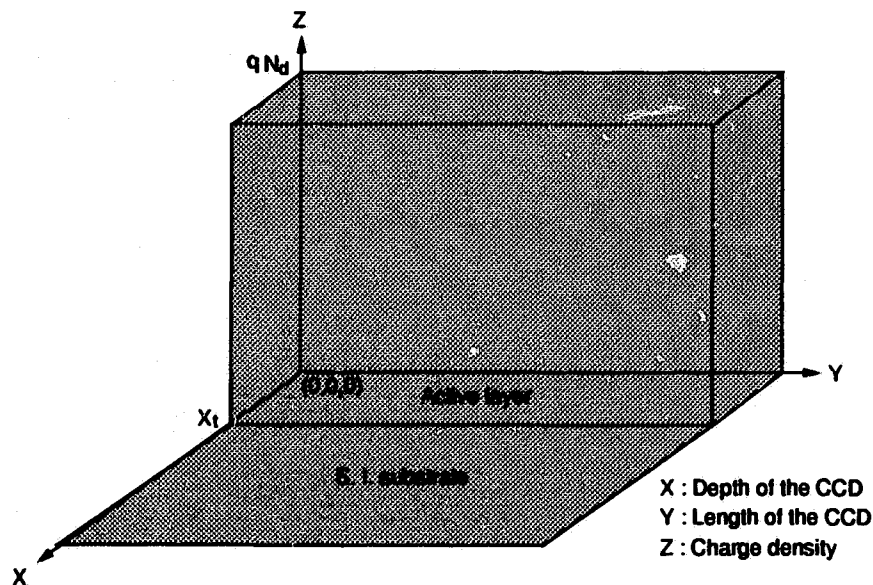


Figure 2.6: Two-dimensional abrupt doping distribution under the CMCCD transport electrode.

In the computations, Eqs. (2.1) and (2.2) are discretized on a finite difference grid superimposed onto a pixel. Fig. 2.7 shows the rectangular mesh adopted for the finite-difference approximation [53]. A grid spacing Δx is superimposed along the depth of the CCD and a grid spacing Δy along the CCD channel. The discretized forms of Eqs. (2.1) and (2.2) are given by:

$$\frac{\psi_{i+1,j}}{\Delta x^2} + \frac{\psi_{i-1,j}}{\Delta x^2} + \frac{\psi_{i,j+1}}{\Delta y^2} + \frac{\psi_{i,j-1}}{\Delta y^2} - 2\left(\frac{\psi_{i,j}}{\Delta x^2} + \frac{\psi_{i,j}}{\Delta y^2}\right) = -\frac{qN_d}{\epsilon_s} \quad (2.3)$$

$$\frac{\psi_{i+1,j}}{\Delta x^2} + \frac{\psi_{i-1,j}}{\Delta x^2} + \frac{\psi_{i,j+1}}{\Delta y^2} + \frac{\psi_{i,j-1}}{\Delta y^2} - 2\left(\frac{\psi_{i,j}}{\Delta x^2} + \frac{\psi_{i,j}}{\Delta y^2}\right) = 0 \quad (2.4)$$

The values of $\psi(i, j)$ are computed at each of the grid points using Eqs. (2.3) and (2.4) and the boundary conditions. Iteration is performed till the difference between the values at each point obtained in the successive iterations is below a preset tolerance which is chosen to be 10^{-4} V in this work.

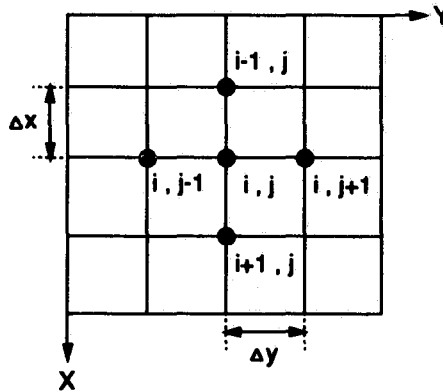


Figure 2.7: Mesh points for the finite-difference approximation.

A computer program *ccd_pot* was developed in C to calculate $\psi(x, y)$ from Eqs. (2.3) and (2.4). The inputs of *ccd_pot* include a vector of the geometric parameters of the device, the clock voltages, and the grid spacings. The grid spacing Δx is chosen to be $0.01 \mu\text{m}$ and Δy to be $0.125 \mu\text{m}$ respectively. We have examined the accuracy of our computation by running *ccd_pot* for a two-phase CMCCD for different values of

Δx and Δy and have found that the further reduction in the grid spacings by half would only improve the accuracy of $\psi(i, j)$ by an average value of 0.8%.

2.5 Two-dimensional Model for Signal Charge Transfer in CMCCD

When a signal charge packet is fed to the input section of a CMCCD, it will move from under one electrode to the next due to the clock voltages. A two-dimensional charge transfer model similar to the one-dimensional model proposed in [54] was developed to study the performance of the CMCCD in terms of charge transfer efficiency. The effects of the surface-state traps and the interface traps were neglected because the charge packets are transferred primarily in the depleted channel of the CMCCD. The generation and recombination of carriers were also ignored. The charge transfer process is described by the continuity equation, i.e.,

$$\frac{\partial n_s(x, y, t)}{\partial t} = \frac{1}{q} \nabla J(x, y) \quad (2.5)$$

where the current density $J(x, y)$ is given by:

$$J(x, y) = qn_s(x, y)v(E(x, y)) - qD(E(x, y)) \nabla n_s(x, y) \quad (2.6)$$

$n_s(x, y, t)$ is the area charge density along the x and y directions. D is the electron diffusivity and is defined as [21]:

$$D(E) = D_0 + D_1 e^{-[\ln(|E|) - \ln(E_p) / \ln(A)]^2} \quad (2.7)$$

$v(E)$ is the electron velocity given by:

$$v(E) = \frac{\mu_0 E}{\sqrt{1 + (\frac{\mu_0 E}{v_s})^2}} \quad (2.8)$$

where

$$v_s = v_1 e^{-\frac{|E|}{E_1}} + \frac{v_2}{1 + \left(\frac{|E|}{E_2}\right)^B} \quad (2.9)$$

The physical constants used in the above equations are the same as those used in [53].

They are listed in Table 2.1.

Parameter	Unit	Value
v_1	cm/sec	4.77×10^7
v_2	cm/sec	3.24×10^7
E_1	V/cm	1644
E_2	V/cm	130.5
B		0.32
D_0	V/cm	129.5
D_1	cm ² /sec	312
E_p	V/cm	3394.8
A	V/cm	1.82
μ_0	cm ² /V.sec	5000

Table 2.1: Physical constants used in the charge transfer modeling.

Eqs. (2.6) suggests that the lateral electric field $E(x, y)$ is responsible for the charge transfer. The electric field is comprised of two terms (see Appendix A):

$$E(x, y) = E_f(x, y) + \frac{q}{\epsilon_s} \left(\frac{n_s(x, y)}{N_d} - x_t \right) \nabla n_s(x, y) \quad (2.10)$$

where x_t is the thickness of the active layer. The first term in Eq. (2.10) is the fringing field which is the derivative of $\psi(x, y)$, the solution of Eqs. (2.1) and (2.2) under the empty well condition (i.e., there is no signal charge in the channel) in a fully depleted channel. The second term of Eq. (2.10) is the self-induced drift field resulting from

the mutual repulsion of the signal charge. In the interelectrode gap, the electric field is given by:

$$E(x, y) = E_{applied} + \frac{q}{\epsilon_s} \left(\frac{n_s(y)}{N_d} - x_t \right) \nabla n_s(x, y) \quad (2.11)$$

where $E_{applied}$ is the applied electric field underneath the cermet filling gap and is assumed to be constant [24].

As can be seen from Eqs. (2.6) and (2.10), a charge packet is transferred from the emptying electrode to the collecting electrode due to three separate mechanisms: the self-induced drift effect, the drift effect due to the fringing field, and the thermal diffusion of the electrons. Physically, the self-induced drift is a charge-to-charge repulsion effect, and it is the main driving force of transferring charge in the early stage of the transfer. In the final stage of the transfer, because the charge under the emptying electrode is very small, the self-induced drift effect becomes very small, and the thermal diffusion which results in an exponential decay of the charge will become more important in the charge transfer. The fringing field is due to the difference of the voltages on the adjacent electrodes. Since the fringing field exists even in the absence of the signal charge, it becomes the dominant mechanism in transferring the last bit of the signal charge when the self-induced drift effect becomes negligible. The magnitude of the fringing field is determined by the geometric parameters and the amplitudes of the clock voltages.

In the finite difference approximation, the time-dependent charge density at the grid point (i, j) is updated in every iteration using the finite difference form of Eq. (2.5) and is given by:

$$n_s(i, j, t + \Delta t) = n_s(i, j, t) + \frac{1}{q} \left\{ \frac{J(i - \Delta x, j) - J(i, j)}{\Delta x} + \frac{J(i, j - \Delta y) - J(i, j)}{\Delta y} \right\} \Delta t \quad (2.12)$$

and the finite difference approximation of Eq. (2.6) is given by:

$$J(i, j) = qn_s(i, j)v(E(i, j)) - qD(E(i, j))$$

$$\times \left\{ \frac{n_s(i + \Delta x, j) - n_s(i, j)}{\Delta x} + \frac{n_s(i, j + \Delta y) - n_s(i, j)}{\Delta y} \right\} \quad (2.13)$$

A computer program *ccd_trsf* was developed in *C* to calculate $n_s(i, j, t)$ at each space grid point and time grid point for a given number of pixels and a given transfer time. In this way CMCCD can be simulated numerically using *ccd_trsf* to evaluate its performance in terms of charge transfer efficiency. The inputs of *ccd_trsf* include a vector of the geometric parameters of the device, the channel potential $\psi(x, y)$ obtained under the empty well condition, the transfer time, the clock voltages and the peak signal charge density. It was found that the two-dimensional charge transfer model given by Eqs. (2.13) and (2.12) improved the accuracy of computation by an average of 3% over the one-dimensional model developed in [53] where the electrons in the charge packet were assumed to locate at the potential maximum along the CMCCD channel.

2.6 Equivalent-circuit Model for CMCCD

The key elements of the CMCCD are the transport electrodes. Based on the one-dimensional charge transfer model of the CMCCD, the electrodes can be modeled as distributed capacitors [55]. Charge transfer is facilitated by current sources linking the capacitors. Fig. 2.8 shows an one-dimensional equivalent circuit model of a unit cell of the CMCCD. Each pixel of the CMCCD is divided into a number of such unit cells. The current sources are given by the following equations [55]:

$$\begin{aligned} i_{1,j} &= \frac{v(E_f(j))}{\Delta y} Q_j && \text{under electrodes} \\ &= \frac{v(E_{\text{applied}})}{\Delta y} Q_j && \text{under gaps} \\ i_{2,j} &= \frac{D(E_f(j))}{\Delta y^2} (Q_j - Q_{j+\Delta y}) && \text{under electrodes} \end{aligned} \quad (2.14)$$

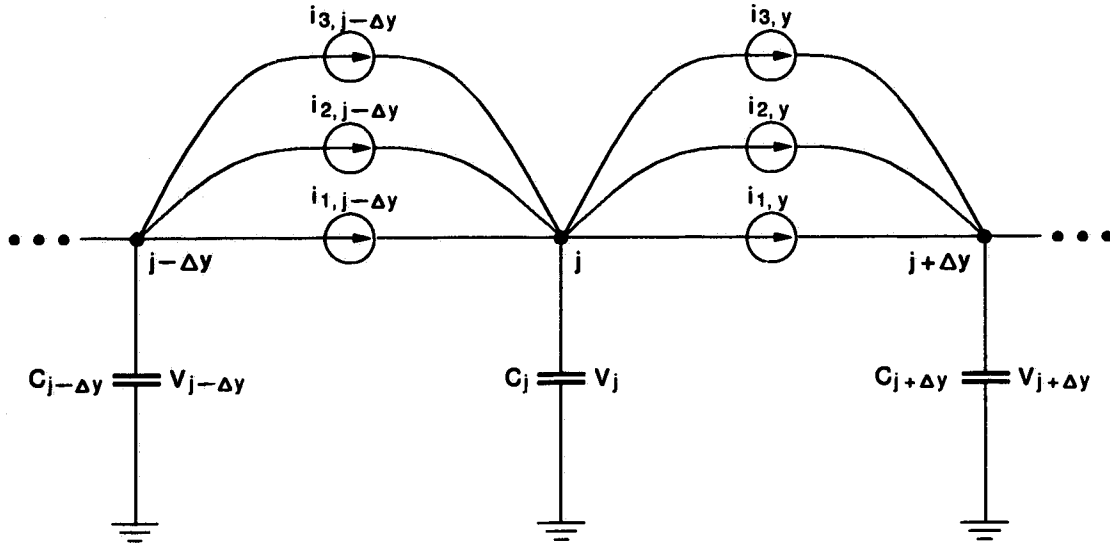


Figure 2.8: Equivalent circuit of unit cell of the CMCCD.

$$= \frac{D(E_{\text{applied}})}{\Delta y^2} (Q_j - Q_{j+\Delta y}) \quad \text{under gaps} \quad (2.15)$$

$$i_{3,j} = \frac{\mu' V_j}{\Delta y^2} (Q_j - Q_{j+\Delta y}) \quad \text{under electrodes}$$

$$= \frac{\mu' V_j}{\Delta y^2} (Q_j - Q_{j+\Delta y}) \quad \text{under gaps} \quad (2.16)$$

where Q_j represents the charge stored at the position j , $E_f(j)$ is the fringing field at position j , $v(E)$ is the field-dependent electron drift velocity, $D(E)$ is the field-dependent electron diffusivity, V_j is the voltage across the capacitor C_j , and μ' is the effective mobility which is given by:

$$\mu' = \frac{v(E(j))}{E(j)} \quad (2.17)$$

In [55], the fringing field was first calculated numerically from the potential distribution, and then put into the equivalent circuit model. If the clock voltages are changed, the fringing field (also the potential distribution) must be calculated again. This offsets the efficiency of the equivalent circuit model and make it impractical.

Only if the fringing field is expressed analytically, the equivalent circuit model can be used practically and flexibly. The fringing field under the k th electrode can be expressed analytically as [56]:

$$E_f(j) = \frac{2\Delta V}{L} \sum_{n=1, n/3 \neq \text{int.}}^{\infty} (-1)^{n+1} \left[\left(\sin \frac{n\pi L_g}{3L} \right) / \left(\frac{n\pi L_g}{3L} \right) \right] \cos \frac{2\pi n j}{L_1} \exp\left(-\frac{2\pi n x_t}{3L}\right) \quad (2.18)$$

where ΔV is the voltage difference on the adjacent electrodes, and is given by:

$$\Delta V = \begin{cases} V_k - V_{k-1} & \text{for } E_f(j) \text{ under the left-half of the electrode} \\ V_{k+1} - V_k & \text{for } E_f(j) \text{ under the right-half of the electrode} \end{cases} \quad (2.19)$$

j is the position along the length of the electrode in increments of Δy , L_1 is the electrode length, L_g is the gap length, $L = L_1 + L_g$, and x_t is the channel thickness. In SPICE3e2, $E_f(j)$ can be easily realized using a nonlinear dependent source function.

The charge-dependent capacitors are of the form [53]:

$$C_j = \frac{C_0}{1 - k_1 Q_j} \quad (2.20)$$

where $C_0 = \epsilon_s w \Delta y / x_t$ and $k_1 = (q w \Delta y N_d x_t)^{-1}$. Similar to Eq. (2.18), Eq. (2.20) can be realized in SPICE using the nonlinear dependent source function. Fig. 2.9 shows the equivalent circuit of the charge-dependent capacitor. The capacitor C_j with a terminal voltage V_j is associated with a charge $Q_j = C_j V_j$. Substituting in the value of the capacitance as a function of charge, V_j can be expressed:

$$V_j = \frac{Q_j(1 - k_1 Q_j)}{C_0} \quad (2.21)$$

To realize the capacitor, a current meter (a zero-voltage source) is used to sense the current flowing into the capacitor. A voltage-controlled current source mimics the sensed current and feeds it to a test capacitor with a capacitance of 1 F. A voltage-controlled voltage source given by Eq. (2.21), with Q_j being replaced by the voltage across the 1 F capacitor, gives the terminal voltage of the charge-dependent capacitor.

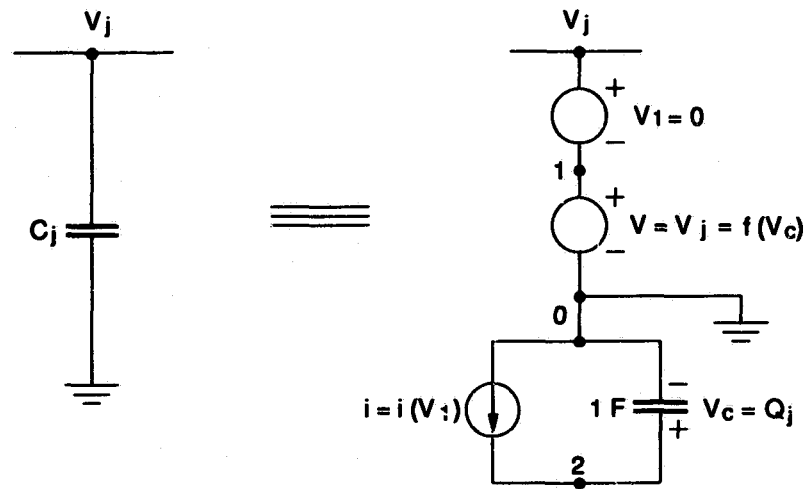


Figure 2.9: Equivalent circuit of charge-dependent capacitor.

The voltage at node 2 represents the charge stored in the original capacitor that is modeled.

It was found by simulation that the differences in charge transfer efficiency between an analytical model and the equivalent circuit model vary between 5% and 20% [55], and the equivalent circuit model, though less accurate, appears to be computationally far more efficient, especially if an analytical expression of the fringing field is used. A SPICE3e2 code for the equivalent circuit model for a unit cell of the CMCCD is given in Appendix B.

Chapter 3

Design Optimization and Simulation of GaAs CMCCDs

3.1 Introduction

It has been noted that the two-phase CMCCD has the most complicated geometric structure since asymmetric electrodes are required to achieve the unidirectional charge flow [25]. In this chapter, a design methodology is developed to optimize the structure of a two-phase CMCCD, and applied to the design of a uni-phase CMCCD. The simulations of the two structures are performed to evaluate the performance in terms of the charge transfer efficiency. The effect of the clock waveforms on the charge transfer efficiency is also simulated. The numerical models and the equivalent circuit model for the CMCCD described in Chapter 2 are used in the design and simulation.

3.2 Two-phase CMCCD Design Optimization

3.2.1 Device Structure

Fig. 3.1 shows a cross-sectional view of a single pixel of a two-phase CMCCD with an asymmetric electrode structure (also known as castellated structure). The dopant density of the active layer N_d is 5×10^{16} . This dopant density was used in the TRIUMF process [24], from which our CMCCD were made. N_d , therefore, is not considered as a variable in our design. The electrodes are divided into two sections with lengths of L_1 and L_2 respectively. They will be called subelectrode 1 and subelectrode 2

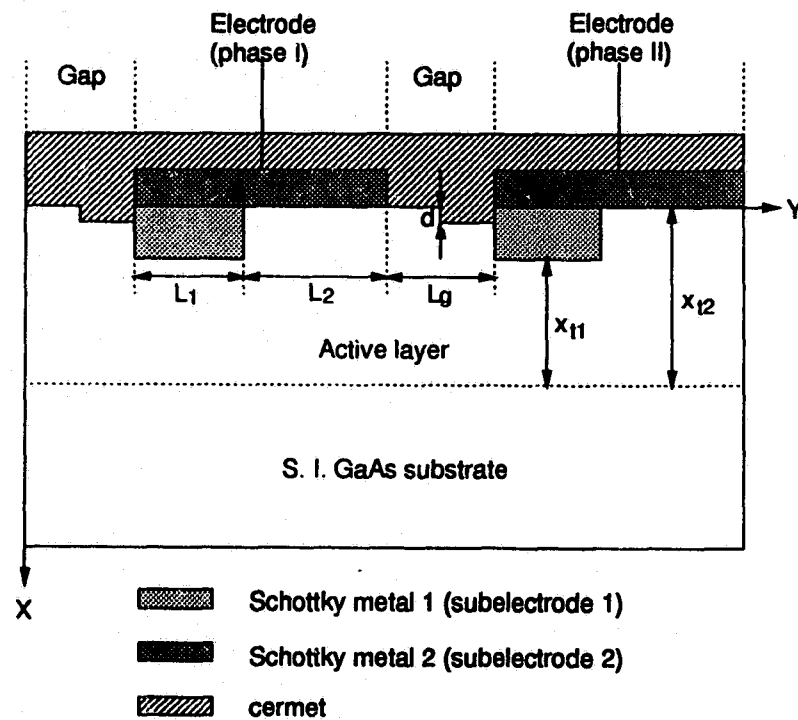


Figure 3.1: Cross-sectional view of a single pixel of a two-phase CMCCD.

and are distinguished from each other by the thicknesses of the active layer, i.e., x_{t1} and x_{t2} . The signal charge is initially stored under subelectrode 2 where the channel

potential is a maximum. Subelectrode 1 has a lower potential and it forms a barrier for the stored charge. Unlike the structure reported in [24] where the cermet layer is planar over the gaps with a length of L_g , the cermet layer appearing in Fig. 3.1 is partially-embedded (50%) into the active layer to a depth of d . The purpose of this, as we shall see, is to ensure that there are no potential troughs under the gaps.

It should be pointed out that although the structure shown in Fig. 3.1 is a two-phase CMCCD, it is possible to change the structure to suit uni-phase, three-phase or four-phase CMCCDs. This can be achieved by letting $x_{t1} = x_{t2}$, $d = 0$, and adding one or more electrodes. In this way, the optimized design methodology described below will be suitable for any of the four clocking schemes.

3.2.2 Optimization of the Two-phase CMCCD Geometry

The two-dimensional channel potential distribution obtained from the program *ccd_pot* is a function of the geometric parameters: L_1 , L_2 , L_g , d , x_{t1} , and x_{t2} (see Fig. 3.1). In optimization we assumed that the electrons always reside at the potential maximum, and only the maximum potential profile along the y direction (the charge transfer direction) is of interest, i.e., $\psi_{max}(y) = MAX\{\psi(x, y)\}$ for all x . This assumption reduces the computing time significantly. Ideally the potential profile of a CMCCD should be flat near the potential maximum ψ_{peak} in order to increase charge storage and it should have steep edges to give a high fringing field. This is illustrated in Fig. 3.2, where ψ_{peak} is the peak potential in the channel when the clock is 0 V, and ψ_{BH} is the barrier height, i.e., the depth of the potential well which determines the amount of charge stored. For a clock voltage of 5 V, we assume $\psi_{peak} = 2.5$ V and $\psi_{BH} = 1.5$ V.

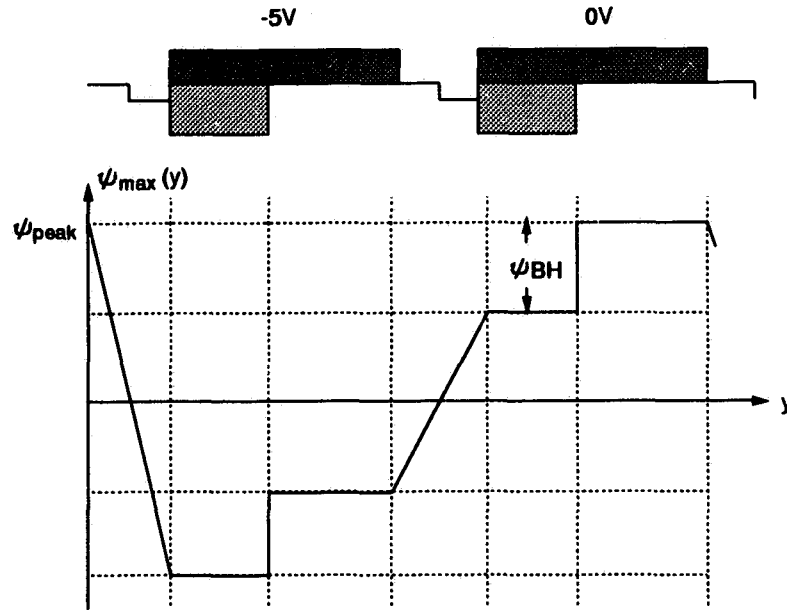


Figure 3.2: Ideal maximum potential profile of the two-phase CMCCD along the CCD channel.

3.2.2.1 Objective Function for Optimization

Let $\psi_{max}(j)$ be the maximum potential profile along the y direction and $\psi_{max-i}(j)$ be the ideal potential profile. For a given clock voltage, $\psi_{max}(j)$ is determined by a geometric parameter vector $\mathbf{X} = [L_1, L_2, L_g, d, x_{t1}, x_{t2}]$ and $\psi_{max-i}(j)$ is determined by some of the elements in \mathbf{X} , namely, L_1, L_2 , and L_g . An objective function is then defined as an error function of the following form:

$$e(\mathbf{X}) = \frac{1}{N_G} \sum_{j=0}^{N_G-1} (\psi_{max}(j) - \psi_{max-i}(j))^2 \quad (3.1)$$

where N_G is the number of the grids along y direction. The optimization can then be described so as to find $\mathbf{X} = [L_1, L_2, L_g, d, x_{t1}, x_{t2}]$, such that $e(\mathbf{X})$ is minimized. Note that L_1, L_2 , and L_g must be the integer multiples of Δy , and d, x_{t1} , and x_{t2} must be the integer multiples of Δx .

3.2.2.2 Implementation of Optimization

It should be noted that the objective function $e(\mathbf{X})$ of Eq. (3.1) is computed numerically rather than analytically. Therefore, the computation of the second order derivatives should be avoided to reduce the computing time. The Fletcher-Reeves method [57], a conjugate-direction-approach-based algorithm, was used in the optimization. The evaluations of $e(\mathbf{X})$ and its gradient and the use of a line-search approach are required by the Fletcher-Reeves method. The gradient of $e(\mathbf{X})$ is expressed by the finite difference approximation. A more detailed description of the Fletcher-Reeves algorithm is given in Appendix C. Fig. 3.3 shows the flow chart of the optimization. A computer program *ccd_opt* was developed in C based on the flow chart to implement the optimization. In the beginning of the iteration, the starting point is to set $\mathbf{X}_0 = [2, 3, 3, 0.05, 0.24, 0.30]$ (μm) which was found to cause a potential trough in the channel.

3.2.2.3 Optimization Results

Table 3.1 lists the results of the optimization, where \mathbf{X}^* is the solution of the error function. The clock voltage is 5 V. The two-dimensional potential profiles of the optimized two-phase CMCCD with and without partially-embedded cermet-gates are plotted in Fig. 3.4 (a) and (b). A cross sectional view of Fig. 3.4, i.e., the

\mathbf{X}^* (μm)						$e(\mathbf{X}^*)$
L_1	L_2	L_g	d	x_{t1}	x_{t2}	
2.3750	3.2500	1.2500	0.0800	0.2400	0.3500	0.0736

Table 3.1: Optimization results of the two-phase CMCCD.

one-dimensional maximum potential profile along the CMCCD channel is plotted in

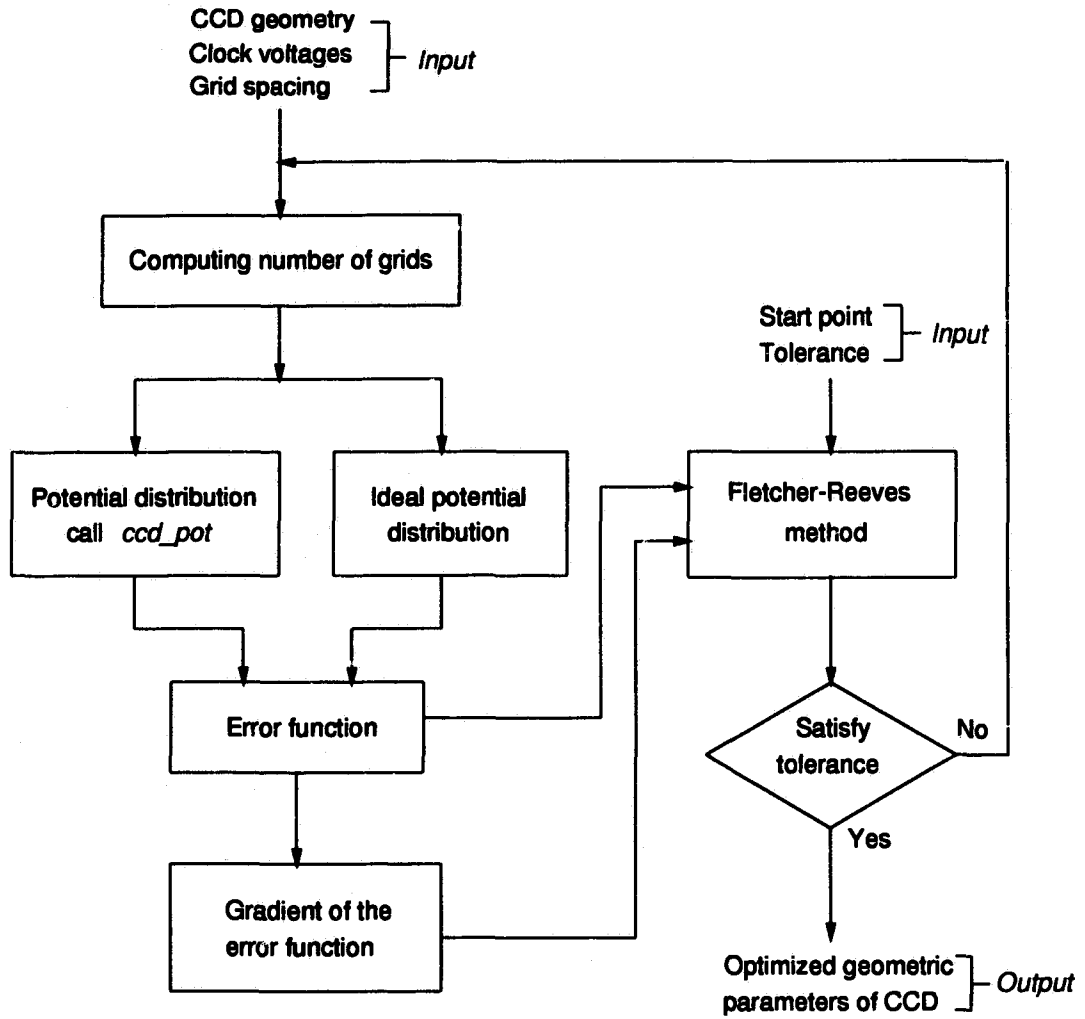


Figure 3.3: Flow chart of the optimization.

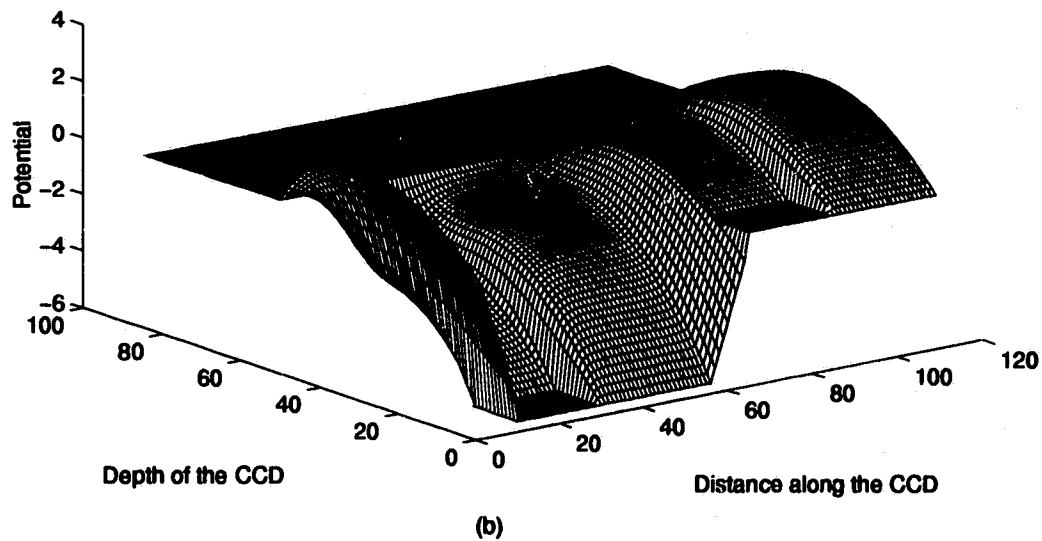
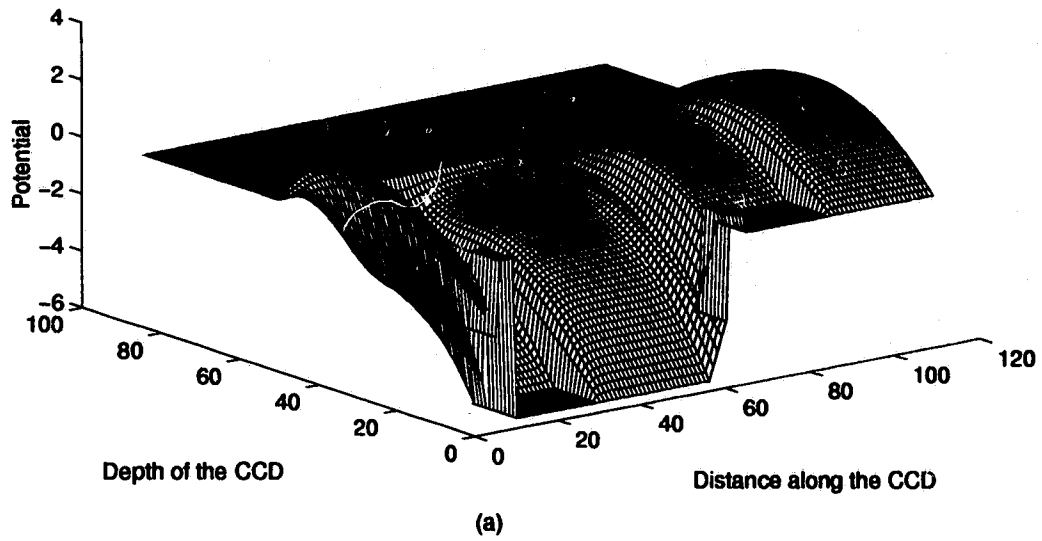


Figure 3.4: Two-dimensional potential profiles of (a) the optimized two-phase CM-CCD and (b) the CMCCD without partially-embedded cermet-gates.

Fig. 3.5. As can be seen from Fig. 3.5, there is no potential trough in the active layer for the optimized structure, whereas there exists a potential trough under the gap when the partially-embedded cermet in the right-half side of the gap is removed (i.e., $d = 0$). The barrier height ψ_{BH} in the optimized device is 1.3 V compared to 1.5 V in the ideal case, indicating a close to the maximal charge handling capability. The optimization results suggest that increasing the depth of the active layer under the electrode results in an increase in charge handling capability, while reducing the depth of the active layer under the gap (through a partially-embedded cermet-gate) is helpful to removal the potential trough.

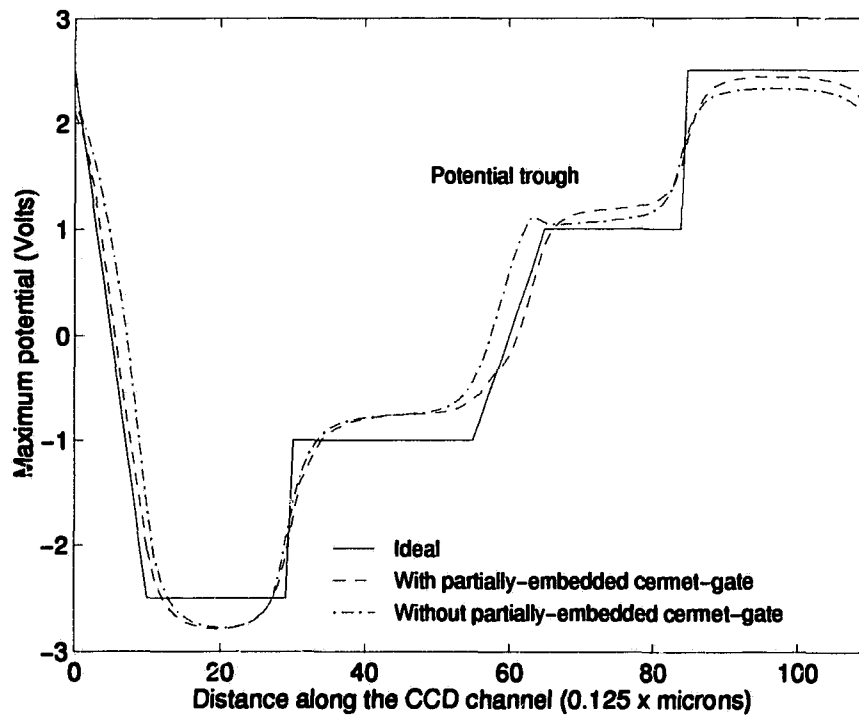


Figure 3.5: Maximum potential profiles of the optimized two-phase CMCCD with and without the partially-embedded cermet-gates.

3.3 Uni-phase CMCCD Design Optimization

3.3.1 Device Structure

The uni-phase CMCCD has a simpler structure than the two-phase CMCCD in the sense that the uni-phase CMCCD uses a planar electrode structure and only one-phase clocks. Fig. 3.6 shows a cross-sectional view of a single pixel of the uni-phase CMCCD. The doping density of the active layer N_d is also $5 \times 10^{16} / \text{cm}^3$. In order to

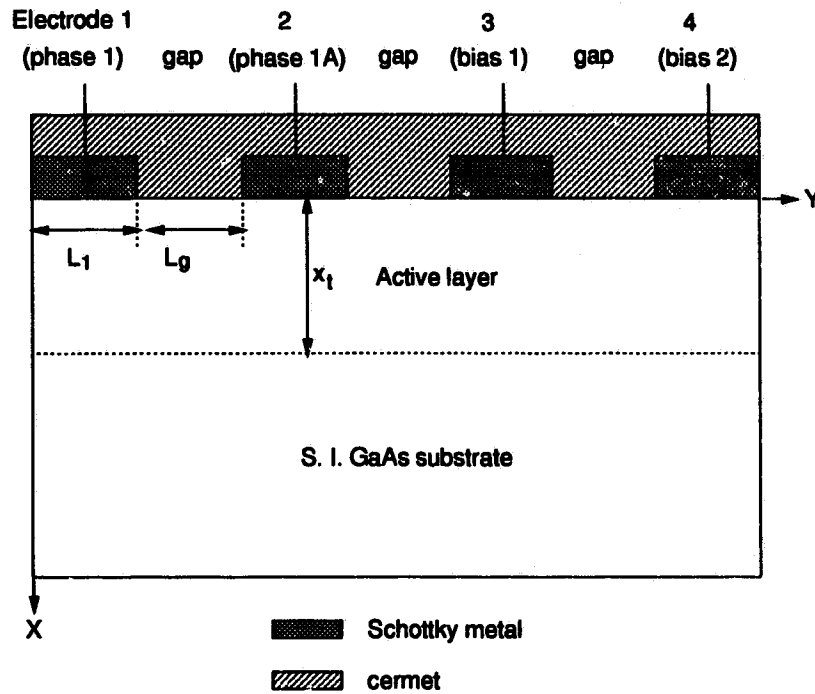


Figure 3.6: Cross-sectional view of a single pixel of a uni-phase GaAs CMCCD.

develop a built-in electric field within the CMCCD channel to direct the charge flow, four electrodes are used to form one pixel. Two of the electrodes are directly clocked and two are dc biased. Fig. 3.7 shows the ideal one-dimensional maximum potential profile along the CCD channel where the clock voltage is 4 V. It can be seen from Fig. 3.7 that at $t = t_1$ the signal charge is under the second electrode (ϕ_{1A}) and at t

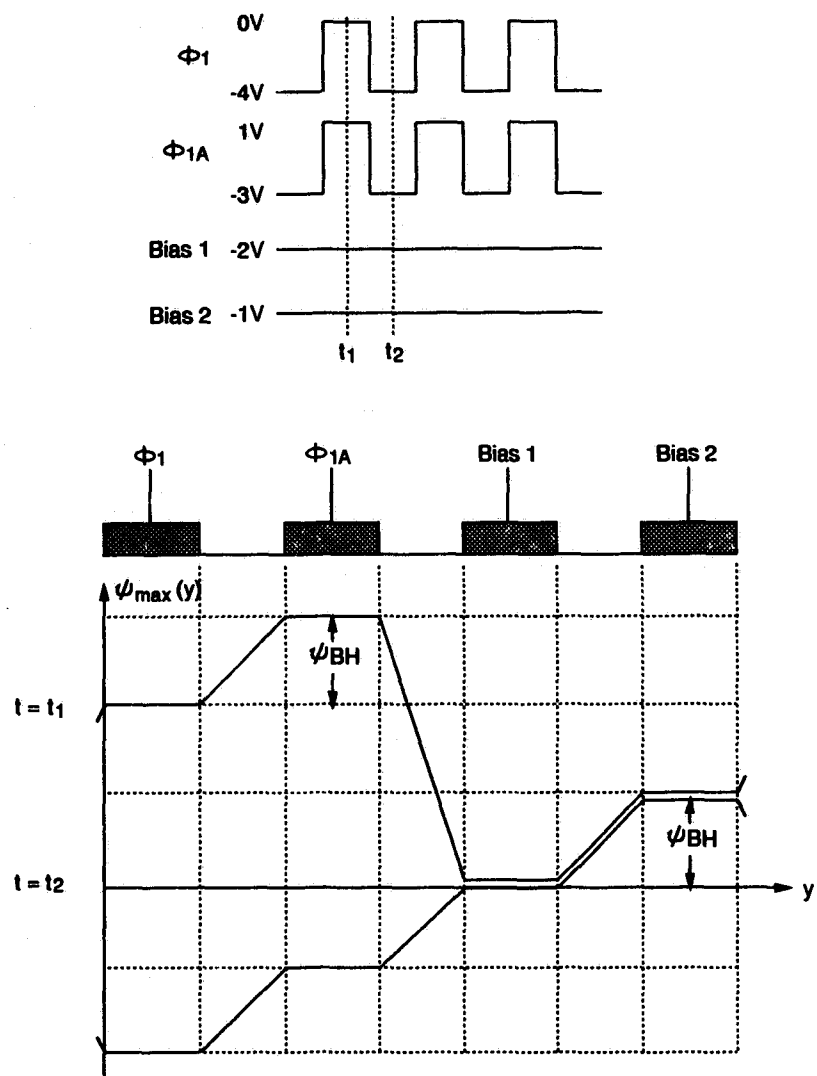


Figure 3.7: Ideal maximum potential profile of the uni-phase CMCCD along the CCD channel.

= t_2 the charge moves to the fourth electrode (*Bias2*) where the potential becomes maximal. In both cases the ideal barrier height ψ_{BH} is chosen to be 1 V.

3.3.2 Optimization of the Uni-phase CMCCD Geometry

The geometric parameter vector for the uni-phase CMCCD is $\mathbf{X} = [L_1, L_g, x_t]$. The objective function of the optimization is the same as in Section 3.1.2, and the same program *ccd_opt* is used to implement the optimization. The clock voltage is 4 V. The starting point is to set $\mathbf{X}_0 = [3, 3, 0.4]$ (μm) which was also found to cause a potential trough in the channel. Table 3.2 lists the optimization results. The

\mathbf{X}^* (μm)			$e(\mathbf{X}^*)$
L_1	L_g	x_t	
3.0000	3.0000	0.2800	0.0525

Table 3.2: Optimization results of the uni-phase CMCCD.

optimization results suggest that a large gap (equal to the length of the electrode) can be used in the CMCCD. This agrees with the design of a resistive-gate CCD where even larger gaps (4 μm gaps and 1 μm electrodes) were used [20]. Fig. 3.8 shows the maximum potential profile along the CCD channel of the optimized uni-phase CMCCD. The solid line in Fig. 3.8 represents the situation when the electrode voltages are (0, 1, -2, -1) (V) while the dashed line when the electrode voltages are (-4, -3, -2, -1) (V). Due to the planar structure of the uni-phase CMCCD, we do not expect to see a potential trough in Fig. 3.8. As can be seen from Fig. 3.8, the barrier height ψ_{BH} in the optimized structure is about 0.8 V compared to 1 V in the ideal case, indicating a reduced charge handling capability.

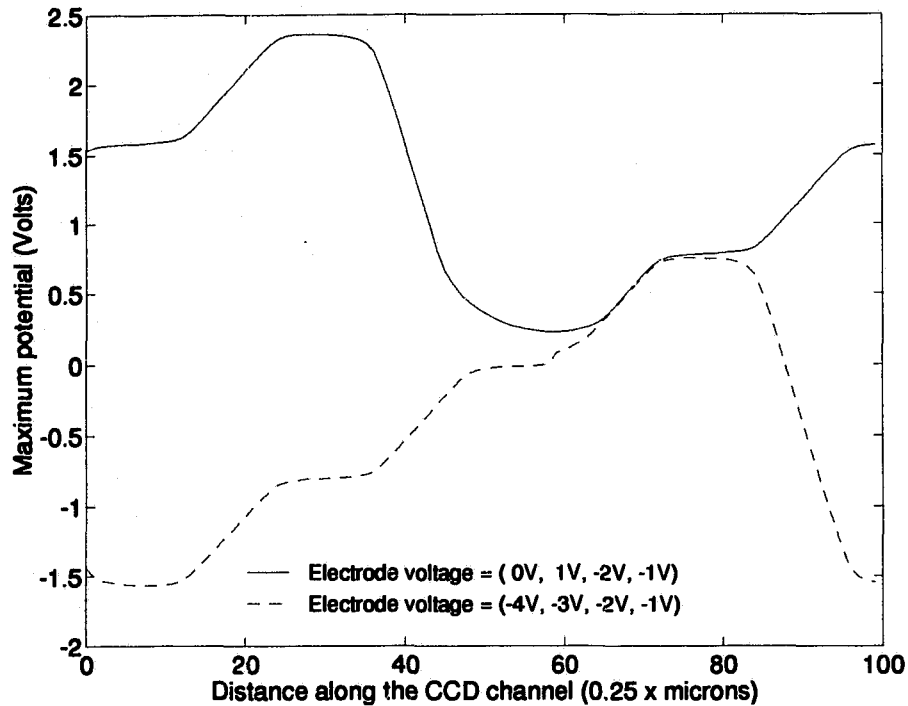


Figure 3.8: Maximum potential profiles of the optimized uni-phase CMCCD.

3.4 Simulations of Charge Transfer Process

One of the important properties of a CCD is the charge transfer efficiency (CTE) which is defined as the ratio of Q_{tf}/Q_o , where Q_{tf} is the amount of charge transferred forward to the collecting electrode and Q_o is the initial charge under the emptying electrode. The charge transfer processes of the optimized two-phase and uni-phase CMCCDs were simulated using the charge transfer program *ccd.trsf* [59]. The initial signal charge distribution is assumed to be a symmetric exponentially decaying distribution function. Simulations were performed to determine the variation of the CTE as a function of the transfer time, the size of the charge packet, and the amplitude of the clock voltage.

3.4.1 CTE versus Transfer Time

The CTEs at different times during the charge transfer are listed in Table 3.3 for the two-phase CMCCD and in Table 3.4 for the uni-phase CMCCD. The peak area charge density Q_{peak} is assumed to be 10^{12} /cm² for the two-phase CMCCD and 5×10^{11} /cm² for the uni-phase CMCCD and the clock voltages are 5 V and 4 V respectively. The simulation results suggest that in the two-phase CMCCD, 99.87

Transfer time (ps)	100	200	300	400	500
CTE	0.6894	0.9913	0.9943	0.9987	0.9989

Table 3.3: Variation of the CTE with respect to the transfer time in the two-phase CMCCD.

Transfer time (ps)	200	400	600	800	1000
CTE	0.6490	0.8531	0.9514	0.9884	0.9990

Table 3.4: Variation of the CTE with respect to the transfer time in the uni-phase CMCCD.

percent of the initial charge has been transferred from the emptying electrode to the collecting electrode after 400 ps and beyond that time, the CTE increased less significantly. The main reason for this is that near the end of the charge transfer, the self-induced drift under the collecting electrode becomes significant. This tends to counter the charge flow due to the fringing field. Consequently, near the end of the charge transfer, the effective electric field is very small, and the charge transfer practically stops. In addition, it is noted that for the uni-phase CMCCD it takes

a longer time for the charge transfer. This is due to the larger separation between one potential well and the next ($12 \mu\text{m}$ in the uni-phase CMCCD versus $6.875 \mu\text{m}$ in the two-phase CMCCD). Our simulation result suggests that the two-phase CMCCD may potentially operate at the clock frequency of 1.25 GHz with a CTE up to 0.9987 and the uni-phase CMCCD at the clock frequency of 500 MHz with a CTE up to 0.9990.

3.4.2 CTE versus Size of Charge Packet

To investigate the charge handling capability of the two-phase and the uni-phase CMCCDs, the time evolutions of the signal charge were simulated with different values of Q_{peak} . Fig. 3.9 shows the time evolution of a two-dimensional charge packet in the two-phase CMCCD, where Q_{peak} is $10^{12} / \text{cm}^2$; the transfer time is 400 ps; and the clock voltage is 5 V. Fig. 3.10 shows the time evolution of an one-dimensional charge packet in the uni-phase CMCCD, where Q_{peak} is $5 \times 10^{11} / \text{cm}^2$; the transfer time is 1000 ps; and the clock voltage is 4 V. Fig. 3.11 shows a plot of the variation of the CTE versus Q_{peak} for the two-phase and the uni-phase CMCCDs. When Q_{peak} increases from $10^{10} / \text{cm}^2$ to $10^{12} / \text{cm}^2$, the CTE drops very slowly. On the other hand, when Q_{peak} is increased beyond $10^{12} / \text{cm}^2$, the CTE reduces more rapidly. This is because the size of the charge packet now exceeds the charge handling capability of the device and the reverse charge flow becomes greatly enhanced [66]. It is noted that the charge holding capability of the two-phase CMCCD is greater than that of the uni-phase CMCCD and this can be explained by the fact that the barrier height in the two-phase CMCCD is 1.3 V while in the uni-phase CMCCD is 0.8 V.

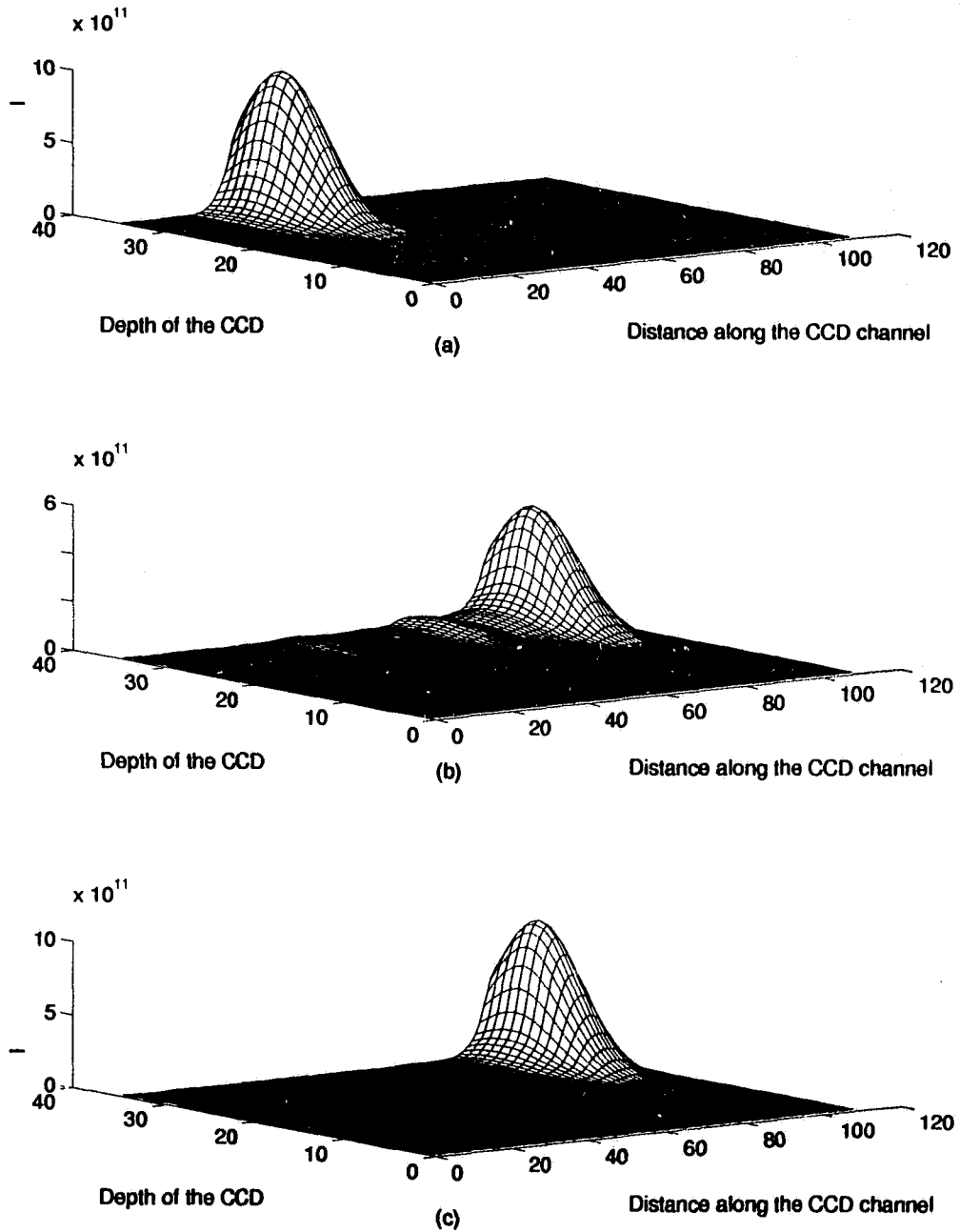


Figure 3.9: Time evolution of a two-dimensional charge packet in the two-phase CMCCD. (a) $t=200$ ps, (b) $t=400$ ps, and (c) $t=600$ ps.

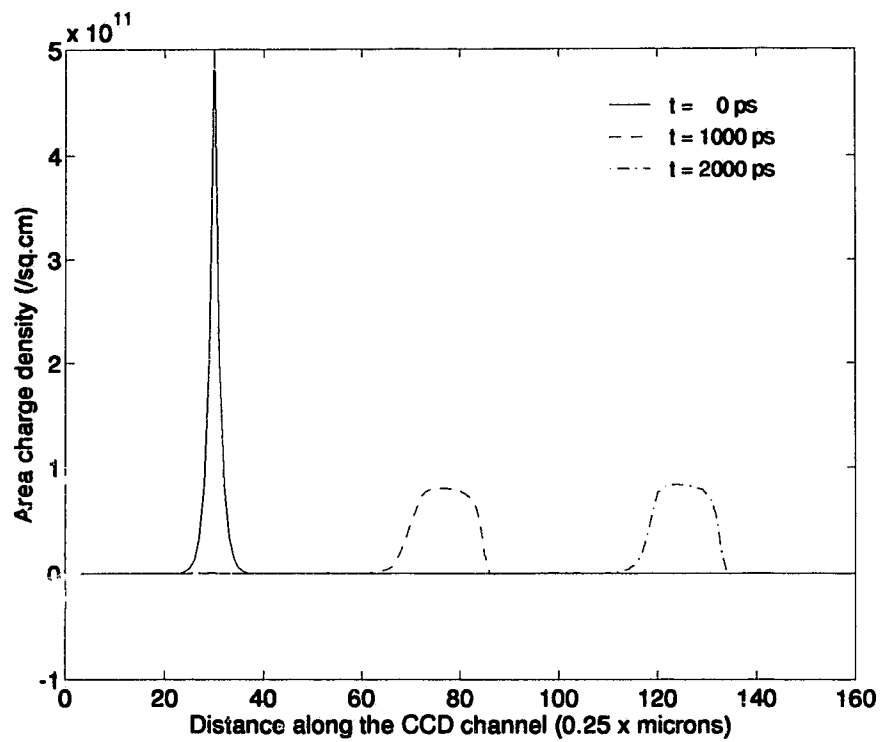


Figure 3.10: Time evolution of a one-dimensional charge packet in the uni-phase CMCCD.

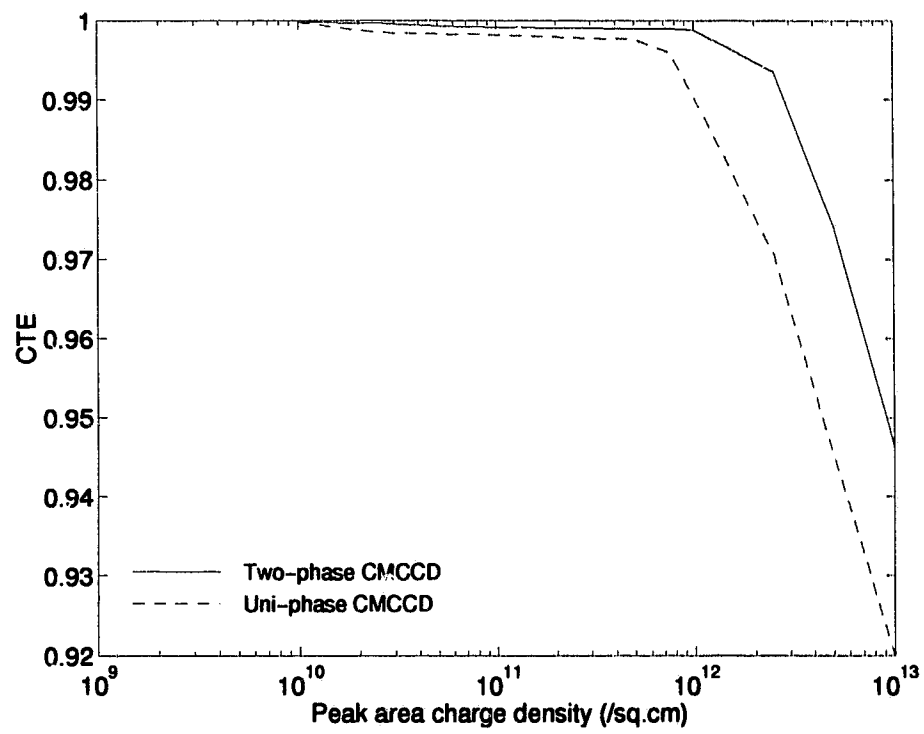


Figure 3.11: CTE versus the the size of the charge packet.

3.4.3 CTE versus Amplitude of Clock Voltage

In a CMCCD, charge transfer is due to the clock voltages applied to the electrodes. The amplitudes of the clock voltages, therefore, affect directly the fringing field and hence the CTE. A plot of the variation of CTE with respect to the peak-to-peak amplitude of the clock voltage is shown in Fig. 3.12, where Q_{peak} is 10^{12} /cm² for the two-phase CMCCD and 5×10^{11} /cm² for the uni-phase CMCCD. The transfer time is 400 ps for the two-phase CMCCD and 1000 ps for the uni-phase CMCCD. It is observed that the CTE decreases when the amplitude of the clock voltage is increased beyond 6 V. This result agrees with a similar observation reported in [20].

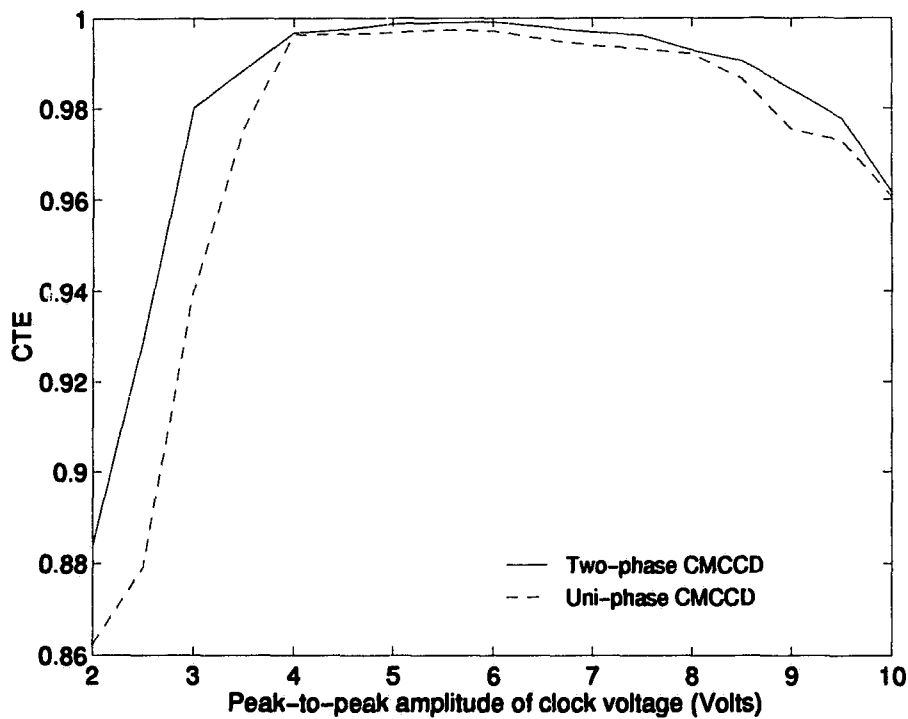


Figure 3.12: CTE versus the peak-to-peak amplitude of the clock voltage.

3.5 Simulations of the Effect of Clock Waveforms on CTE

Charge transfer in a CCD is also affected by the clock waveforms, especially at high frequency. It was noted that the fringing field under the emptying electrode largely determines the effectiveness of the charge transfer to the collecting electrode, and the difference of the clock voltages at the adjacent electrodes makes an important contribution to the value of the fringing field. For practical device operation, the clock waveforms will have finite rise time, delay time and fall time (as shown in Fig. 3.13) due to the parasitic effect. Using the equivalent circuit model of the GaAs

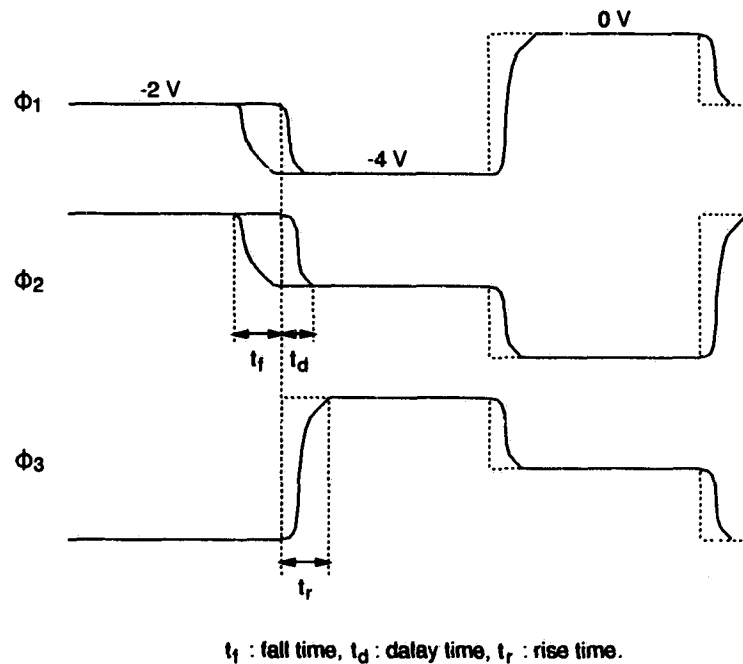


Figure 3.13: Waveforms of the three-phase three-level stepped clocks.

CMCCD described in Section 2.4, we developed a SPICE3e2 program to investigate how the rise time, the delay time and the fall time of the clock waveforms could

affect the CTE [62]. In the simulation, a three-phase clocking scheme was used as the clock voltages in a three-phase CMCCD represent the most complicated three-level waveforms. The electrodes and gaps are both $2 \mu\text{m}$ in length and $100 \mu\text{m}$ in width and they are divided into 16 segments (i.e., $\Delta y = 0.125 \mu\text{m}$). The thickness of the active layer is assumed to be $0.40 \mu\text{m}$ and charge transfer time is 200 ps. The peak area charge density is $5 \times 10^{11} / \text{sq.cm}$ and the clock voltage levels are -4 V , -2 V and 0 V , respectively.

3.5.1 CTE versus Rise Time

Figure 3.14 gives the time evolution of the CTE of the CMCCD for different values of the rise time of the clock voltage. With an increase in the rise time, the transfer process was slowed down considerably. The main reason for this is that the rise time of the clock voltage on the collecting electrode determines how fast the potential well will be formed. Since the voltage on the collecting electrode takes a finite time to change from -4 V to 0 V , charge cannot be transferred from the emptying electrode to the collecting electrode instantaneously. This means that a substantial amount of charge may remain under the emptying electrode during this transition. The slower the rise time, the larger the amount of charge that remains under the emptying electrode. In addition, the holding time of the peak potential ($= 0 \text{ V}$ in the simulation) during charge transfer will be proportionally reduced, and this also reduces the CTE. The effect of the rise time on the fringing field is given in Table 3.5 which shows that the average fringing field under both the left-half and the right-half of the emptying electrode are indeed smaller for a faster rise time.

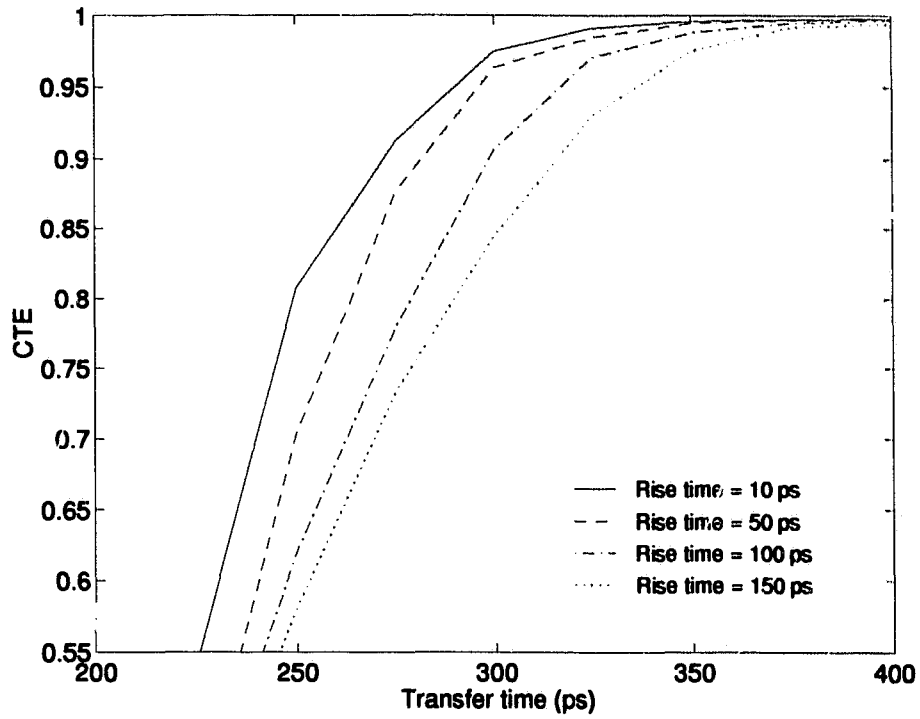


Figure 3.14: CTE versus transfer time for different values of the rise time. Fall time is 10 ps.

Rise times (ps)	E_l (KV/cm)	E_r (KV/cm)
10	3.642	2.508
50	3.315	2.448
100	2.818	2.355
150	2.438	2.096

Table 3.5: Average values of the fringing field under the emptying electrode for the different rise time.

3.5.2 CTE versus Delay Time

CTE is less affected by the delay time than the rise time and this is shown in Fig. 3.15. As the voltage at the collecting electrode rises rapidly to a high value (for a rise time of 10 ps), the potential well will begin to form, while at the same time, the voltages at the emptying electrode and the preceding electrode go down slowly. It is observed that the voltage difference between the preceding electrode and the emptying electrode remains the same during the delay time and hence the fringing field under the left-

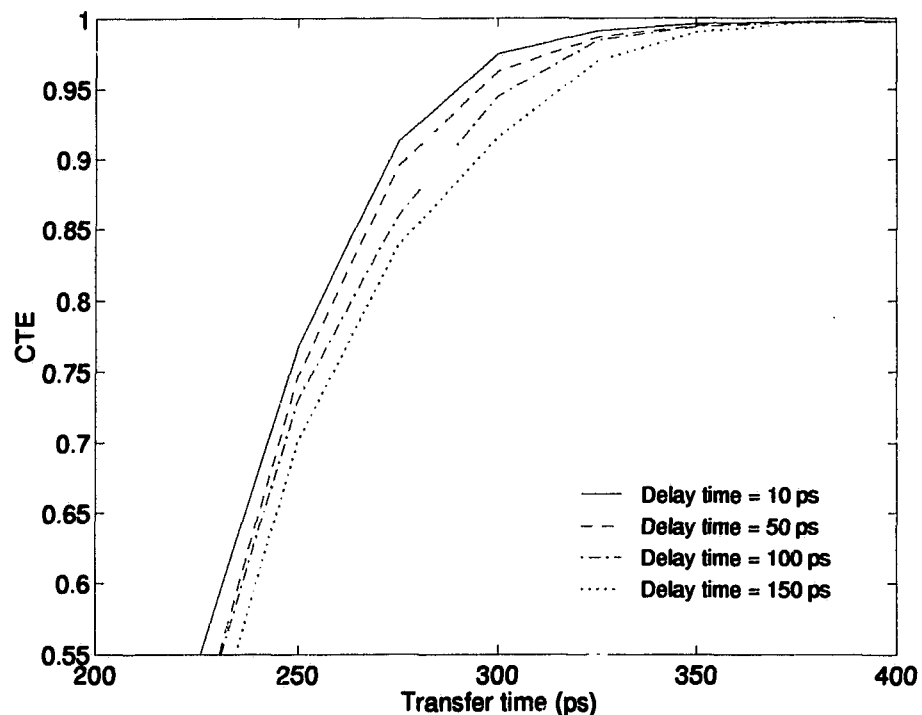


Figure 3.15: CTE versus transfer time for different values of the delay time. Rise time is 10 ps.

half of the emptying electrode reduces slightly, resulting in a reduction in CTE. Note that the charge packet is transferred from left to right. Table 3.6 gives the average fringing field under the emptying electrode for the different delay time.

Delay times (ps)	E_i (KV/cm)	E_r (KV/cm)
10	3.642	2.508
50	3.642	2.240
100	3.642	1.955
150	3.642	1.653

Table 3.6: Average values of the fringing field under the emptying electrode for the different delay time.

3.5.3 CTE versus Fall Time

Fig. 3.16 shows the variation of CTE versus the transfer time for different values of the fall time when the rise time is 10 ps. The fall time almost has no effect on the charge transfer. This is because most of the charge has already been transferred from the emptying electrode to the collecting electrode during the fall time. It can be observed that as the voltages on the emptying electrode and the collecting electrode go down together, the voltage difference between these two electrodes remains unchanged, and therefore the fringing field under the right-half of the emptying electrode does not change either. Only under the left-half of the emptying electrode does the fringing field go down. But it is obvious that the fringing field under the right-half of the emptying electrode affects the charge transfer much more significantly than the left-half when the signal charge moves from left to right. Table 3.7 gives the average fringing field under the emptying electrode for the different fall time.

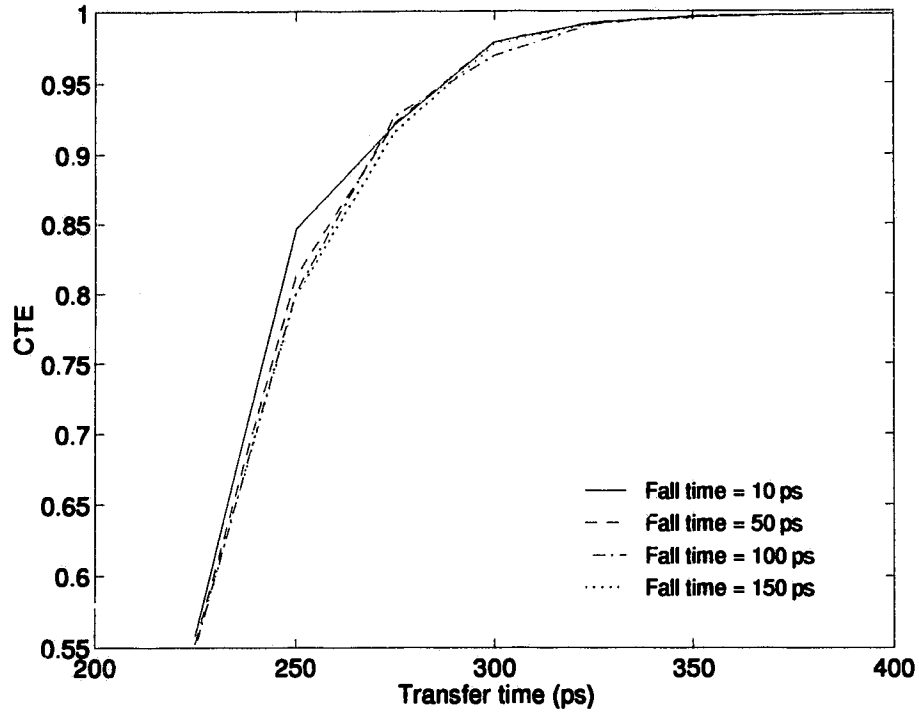


Figure 3.16: CTE versus transfer time for different values of the fall time. Rise time is 10 ps.

Fall times (ps)	E_l (KV/cm)	E_r (KV/cm)
10	2.750	2.051
50	2.420	2.051
100	2.053	2.051
150	1.655	2.051

Table 3.7: Average values of the fringing fields under the emptying electrode for the different fall time.

Chapter 4

Fabrication and Measurement of GaAs CMCCD

4.1 Introduction

A uni-phase CMCCD with an optimized geometric structure was fabricated using the TRIUMF (Tri-University Meson Facility) GaAs process. The fabrication used a 2 μm minimum design rule and required six mask layers. The device was tested at the clock frequency ranging from 100 KHz to 100 MHz. In this chapter, the fabrication of the device is briefly described, and the experimental results are presented and discussed. More specific test results on noise and linearity will be given in Chapter 6.

4.2 CMCCD Fabrication

The TRIUMF GaAs process was designed and optimized for CMCCD fabrication. It is also possible to integrate monolithically MESFET output circuits with the CMCCDs. The wafer used for the CMCCD fabrication was an undoped semi-insulating

GaAs onto which an epitaxial layer had been grown. The epitaxial layer consists of a 1 μm thick undoped buffer layer and a 0.3 μm thick uniformly doped n-type active layer. The fabrication of the GaAs CMCCD required six mask layers; one for each of the following steps: the ohmic contacts, the isolated active regions, the metal/GaAs Schottky gates, the cermet/GaAs Schottky gates, the interconnect vias and the second level metallization. Conventional contact lithography was used. Table 4.1 lists the main process parameters of the TRIUMF GaAs technology.

Parameter	Value
Active layer thickness	0.3 μm
Epitaxial doping density	$5 \times 10^{16} / \text{cm}^3$
Pinch-off voltage	-2.0 V
Min. design rule	2 μm
Cermet gate	Yes

Table 4.1: Main process parameters of the TRIUMF GaAs technology.

Fig. 4.1 shows the microphotograph of the 128 pixel, uni-phase GaAs CMCCD fabricated at TRIUMF. The input section, located on the left-hand end of the CMCCD, consists of an input ohmic contact and two control gates G_1 and G_2 . This input structure allows for charge injection by the fill-and-spill method. The output section is used for charge detection and it consists of a control gate G_3 , an output ohmic contact, a reset MESFET and a source follower formed by two depletion-mode MESFETs of 30 μm wide. The CMCCD chip was mounted in a 28-pin ceramic flat package of LCC28TS2. Table 4.2 lists the geometric parameters of the CMCCD.

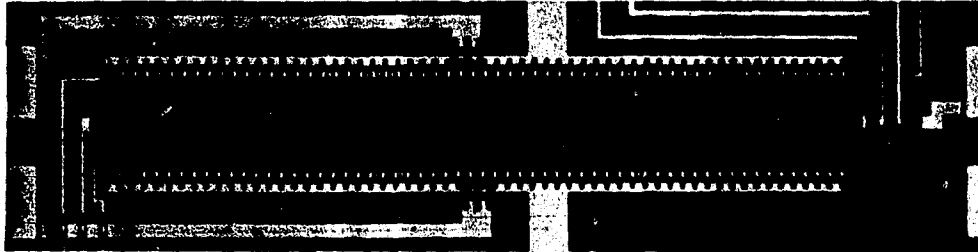


Figure 4.1: Photomicrograph of the CMCCD.

Parameter	Value (μm)
Width of CCD channel	100
Length of electrode	3
Length of gap	3
Length of G_1 , G_2 and G_3	3
Length of ohmic contact	5
Space between G_1 and G_2	3
Space between G_1 and input ohmic contact	2
Space between G_2 and first electrode	3
Space between G_3 and last electrode	3
Space between G_3 and output ohmic contact	2

Table 4.2: Geometric parameters of the CMCCD.

4.3 CMCCD Test Setup

Fig. 4.2 shows the test setup for the CMCCD. A HP-8130A dual-channel pulse generator and a HP-8131A signal-channel pulse generator were used to generate the clocks ϕ_1 and ϕ_{1A} and the sampling pulses $V_{I/P}$. A HP-8112A pulse generator was used to synchronize the clocks and $V_{I/P}$. A Philip PM-5134 function generator was used for the input signals. The fill-and-spill method was used for charge injection by feeding the input signal to G_1 and biasing G_2 at a dc level. The signal levels and the dc bias voltages used for testing the CMCCD are listed in Table 4.3.

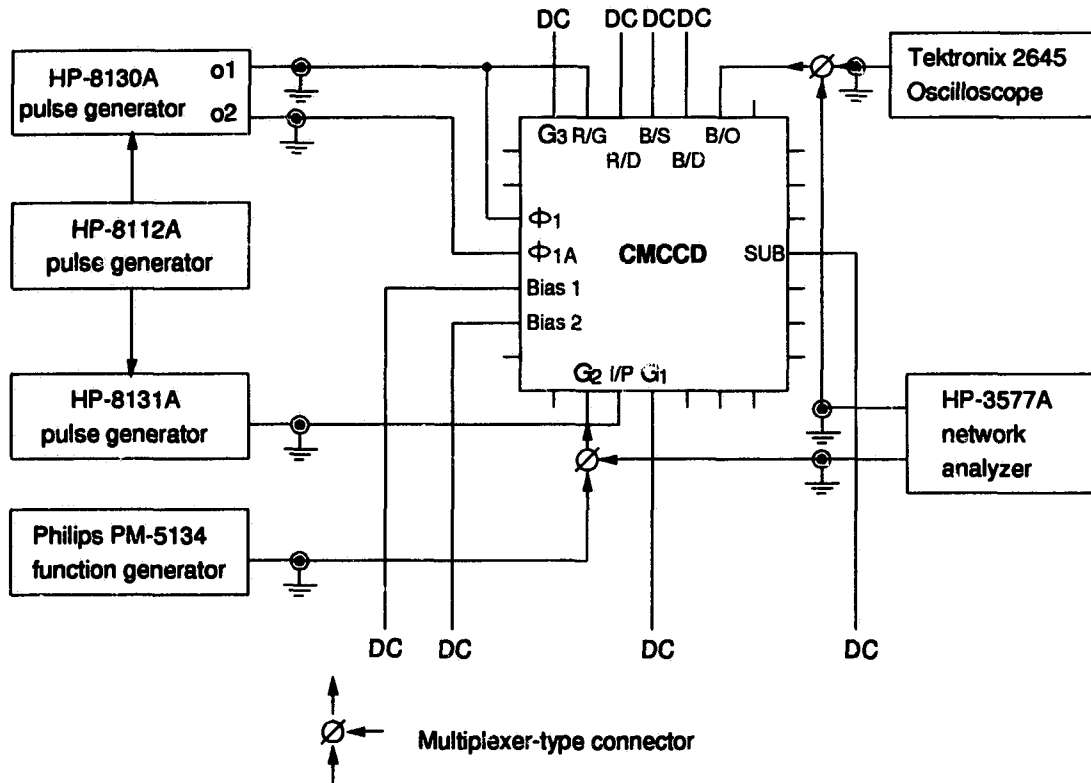


Figure 4.2: Block diagram of the CMCCD test setup.

Gate	Signal level (Volts)
IP	-2 to +1 (pulses)
G ₁	-1 to 0 (input)
G ₂	0
ϕ_1	-4 to 0 (clock)
ϕ_{1A}	-3 to +1 (clock)
Bias 1	-2
Bias 2	-1
G ₃	0
RG	-4 to 0 (clock)
RD	+3
BD	+5
BS	-5

Table 4.3: Signal levels and dc bias voltages using for testing the CMCCD.

4.4 Experimental Results

The I - V characteristics of the CMCCD were measured prior to testing the charge transfer in order to determine if the gates and the transport electrodes were functional. The input ohmic contact and the output ohmic contact of the CMCCD were used as the drain and the source, respectively. All of the electrodes and the three control gates were connected together to the gate voltage. A HP-4145A semiconductor parameter analyzer was used to measure the I - V characteristics which is shown in Fig. 4.3, where ΔV_g is 0.25 V. The measured threshold voltage is approximately -1 V and the saturation current is approximately $6.5 \mu\text{A}$ when the gate voltage was 0.5 V.

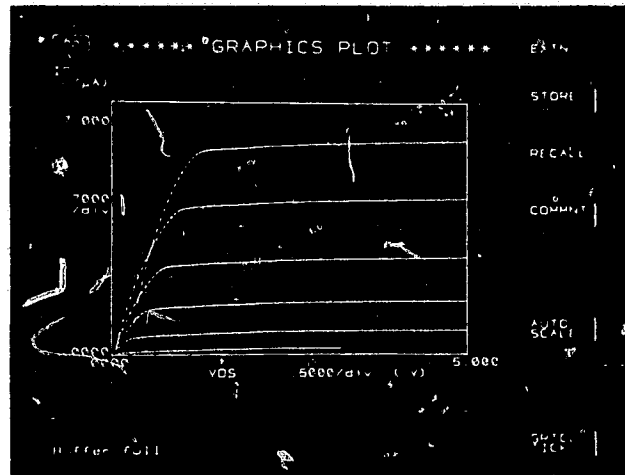


Figure 4.3: I - V characteristics of the CMCCD.

The GaAs CMCCD was tested at the clock frequency ranging from 100 KHz to 100 MHz by feeding its input (gate G_1) with either a sawtooth signal or a pulse signal. The measured input and output waveforms of the CMCCD are shown in Figs. 4.4 and 4.5 where the CMCCD operated at 10 MHz and 50 MHz, respectively. The oscillographs display the CMCCD input waveform along the upper trace and the processed CMCCD output waveform along the bottom trace. The CMCCD output waveform at 10 MHz contains the delayed input waveform modulated by the clock

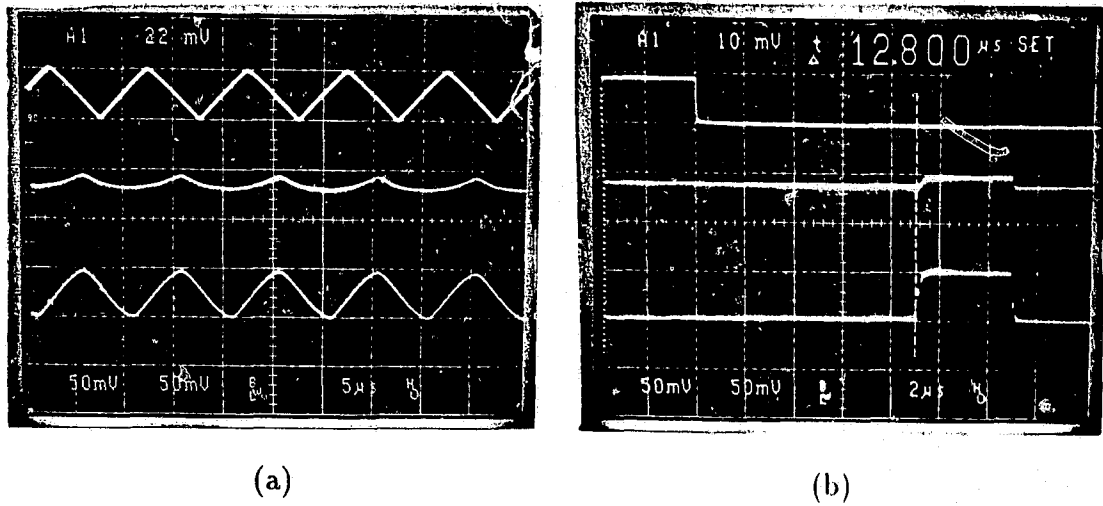


Figure 4.4: Oscillograph of the CMCCD for 10 MHz operation.

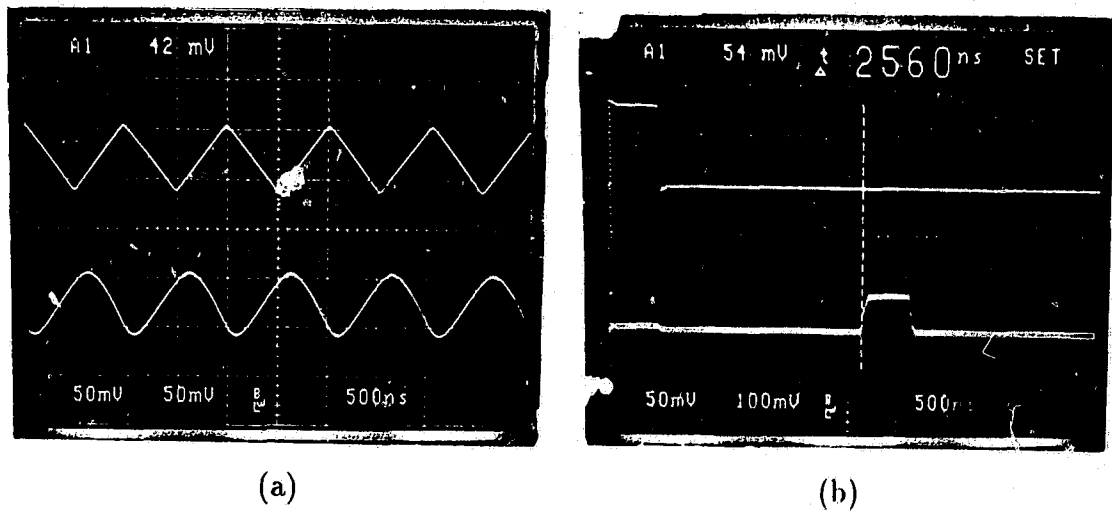


Figure 4.5: Oscillograph of the CMCCD for 50 MHz operation.

feedthrough. The clock feedthrough can be removed using a low-pass filter, as shown in the oscillographs at 50 MHz. The pulse responses of Figs. 4.4 and 4.5 show the signal delay at the output to be $12.80 \mu\text{s}$ and $2.56 \mu\text{s}$ at 10 MHz and 50 MHz, respectively. Such measurements agree with the theoretical calculations of the delay time of a CMCCD which is given by:

$$t_{\text{delay}} = \frac{N_{\text{pixel}}}{f_c} \quad (4.1)$$

where N_{pixel} is the number of pixels of the CMCCD and f_c is the clock frequency.

A more detailed quantitative demonstration of the performance of the CMCCD is shown in Fig. 4.6, where the waveforms of the input signal, the output signal and the clock are displayed. The spikes near the positive and the negative transitions of the output is the result of the coupling of the reset pulse to the output ohmic contact through the parasitic capacitance C_{GS} of the reset MESFET. This happens all the time for the single-transistor pass-gate.

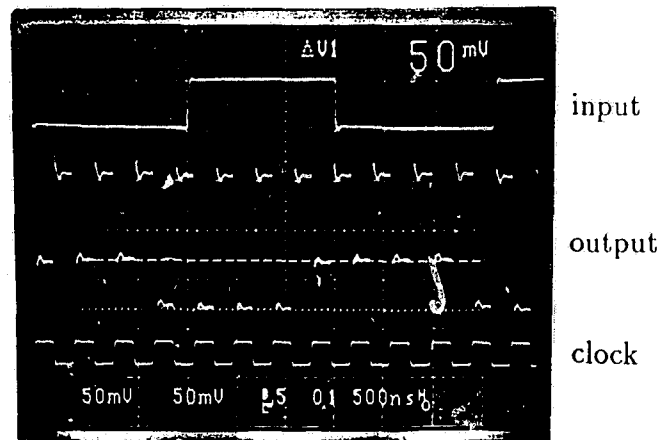


Figure 4.6: More detailed quantitative demonstration of the performance of the CMCCD.

The CTE, denoted by η , of the CMCCD was determined by fitting the data to

the following equation [69]:

$$G_v = \exp[-N_T(1 - \eta)(1 - \cos \frac{2\pi f}{f_c})] \quad (4.2)$$

where G_v is the amplitude response, N_T is the number of transfers through the CMCCD, f is the input signal frequency and f_c is the clock frequency. Fig. 4.7 shows the typical amplitude response of the CMCCD obtained from a HP-3577A network analyzer. The measured CTEs at different clock frequencies are listed in Table 4.4.

The small degradation in the CTE at low frequency, as observed at 1 MHz, is at-

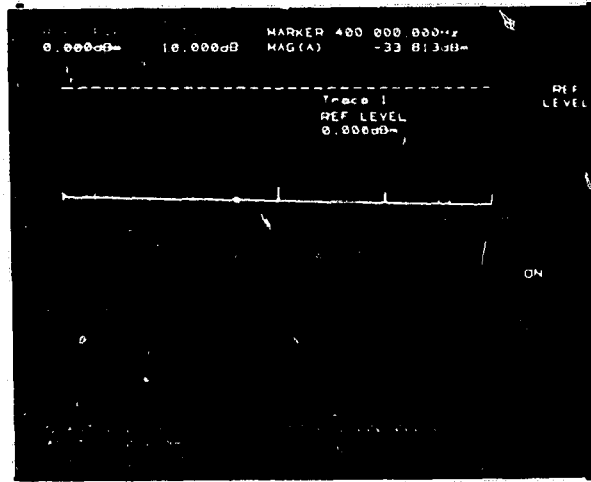


Figure 4.7: Typical amplitude response of the CMCCD.

Clock frequency (MHz)	1	2.5	10	50	100
CTE	0.9982	0.9990	0.9990	0.9955	0.9810

Table 4.4: Measured CTEs at different clock frequencies.

tributed to the dark current which is due to thermally generated electron-hole pairs in the depleted CMCCD channel. The dark current limits the device performance

at low frequency [25]. At very low clock frequency the dark current begins to fill up the potential well and causes a decrease in the amplitude of the output signal. This therefore reduces the CTE.

The effect of the clock waveforms on the CTE was also tested by measuring the amplitude responses of the CMCCD for the different leading edges of the clock pulses. The CTEs were then calculated using Eq. 4.2. The measured and simulated CTEs versus the normalized leading edges of the clock pulses are plotted in Fig. 4.8. The simulation result was obtained by running a SPICE3e2 program incorporated with the equivalent circuit model for the uni-phase CMCCD. The difference between the measurement and the simulation, as observed, is quite small.

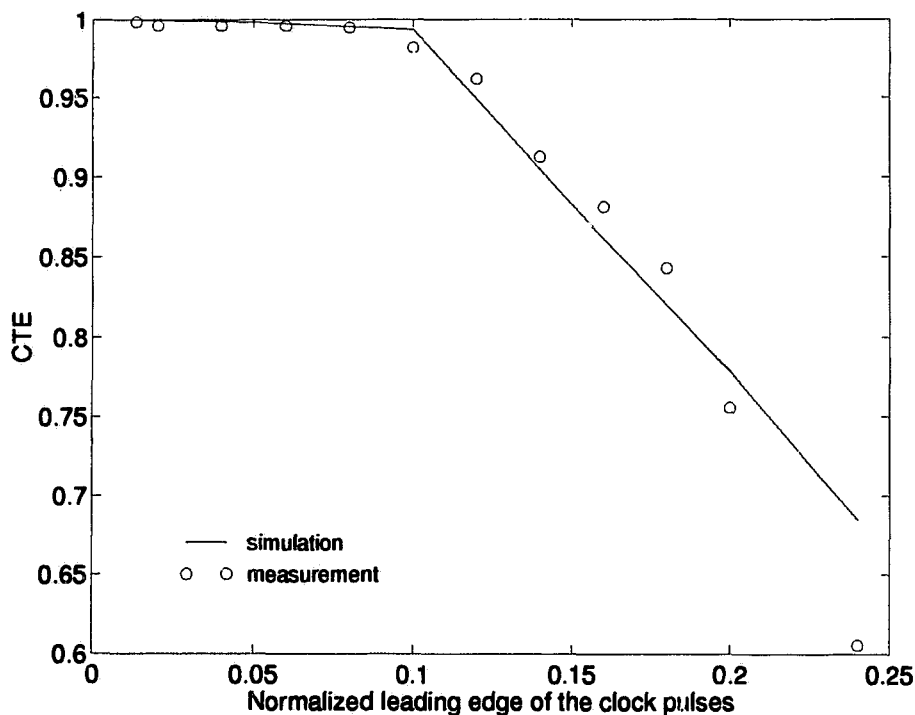


Figure 4.8: Measurement of the CTE with respect to the normalized leading edge of the clock pulses.

Chapter 5

Hybrid Architecture for ANN Implementation

5.1 Introduction

Artificial neural networks (ANNs) attempt to mimic, at least partially, the structure and the function of the human brain and the nervous system. Advances in VLSI technology and development of the simplified neural model originated by Hopfield [38] in 1982 make it possible to realize electronic neural networks in large-scale. In this chapter, the concept of the ANN is introduced, and the mathematical model of the Hopfield ANN is described. Based on the Hopfield model a hybrid architecture for ANN implementation is developed to overcome the problems suffered by the conventional parallel architecture. The overview of the hybrid architecture is also given.

5.2 Concept of ANNs

The biological neuron is a complex analog computing device. Fig. 5.1 shows the models of a biological neuron and an artificial neuron. The neuron consists of a cell body, or *soma*, branching extensions that serve as inputs, or *dendrites*, and the output channel of the cell, or *axon*. The axon carries the electrical signal to the other

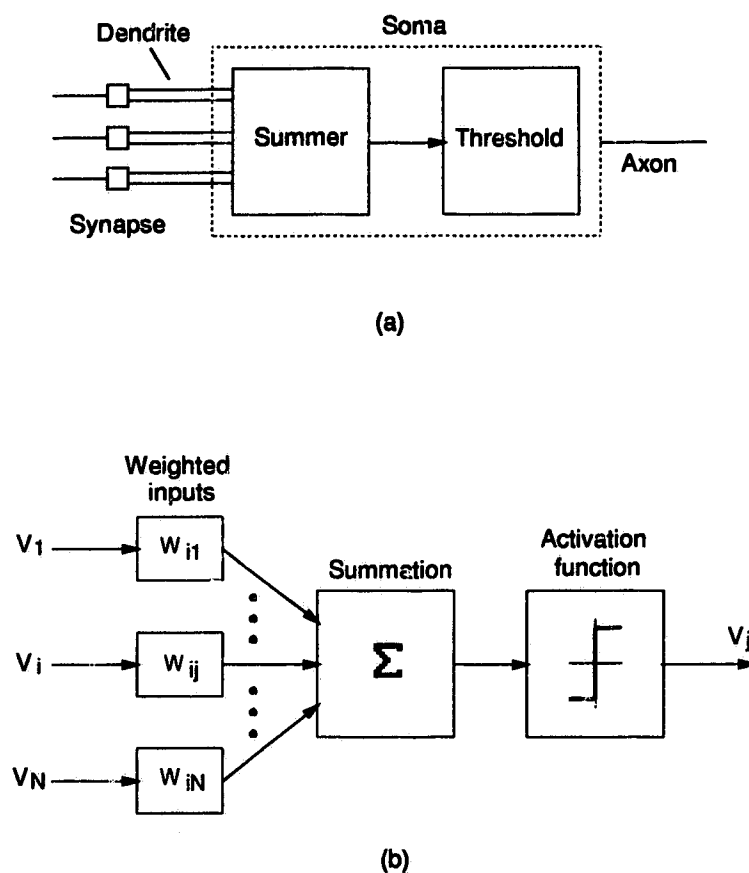


Figure 5.1: Models of (a) a biological neuron and (b) an artificial neuron.

neurons. The connections between the neurons are through specialized contacts called *synapses* that can change either positively or negatively the potential of the axon. The function of the synapses is therefore to weight the the neuron inputs which

are the outputs of the other neurons. In general, the biological neuron performs the simple operation of summing the weighted input signals, i.e., inner product of two vectors, one for the neuron outputs and other for the synaptic weights. The total weighted sum changes the level of activity of the neuron. If the result exceeds a certain threshold, a different signal is generated from the neuron. An artificial neuron can be modeled according to the simplified biological neuron model. As shown in Fig. 5.1(b), an artificial neuron consists of the weighted inputs, the summation function, and the activation function (also called the threshold function) that generates a binary output.

The operation performed by the neurons can be used in pattern recognition where the neural network is also known as associative memory. Unlike conventional digital memory in which each stored word is retrieved by providing its address, the associative memory has no address, a memory word (or pattern) is retrieved by identifying the best match between an input pattern and a stored pattern. If the input pattern is the nearest match to one of the stored patterns, that stored pattern will appear at the memory output. This is similar to the way the human brain seems to behave, i.e., some of a memory input can evoke other associated data.

The Hopfield ANN can perform the function of an associative memory by coding the patterns to be stored in the synaptic weights [39, 48]. Fig. 5.2 shows the Hopfield ANN with N neurons. The output of each neuron is fed back to the other neurons through the synapses, and each synapse computes a new output according to the outputs of the neurons and the weights stored in the synapses. The outputs of all the synapses connected to the same neuron are then summed and this summation updates the neuron state (output) through its activation function. If an unknown N -bit binary input pattern is applied at time zero, the network will iterate until convergence when the neuron outputs reach the stable states most similar to the initial inputs. Because the Hopfield ANN can employ an enormous number of interconnections among the neurons to perform parallel processing, they have the potential benefits of a high

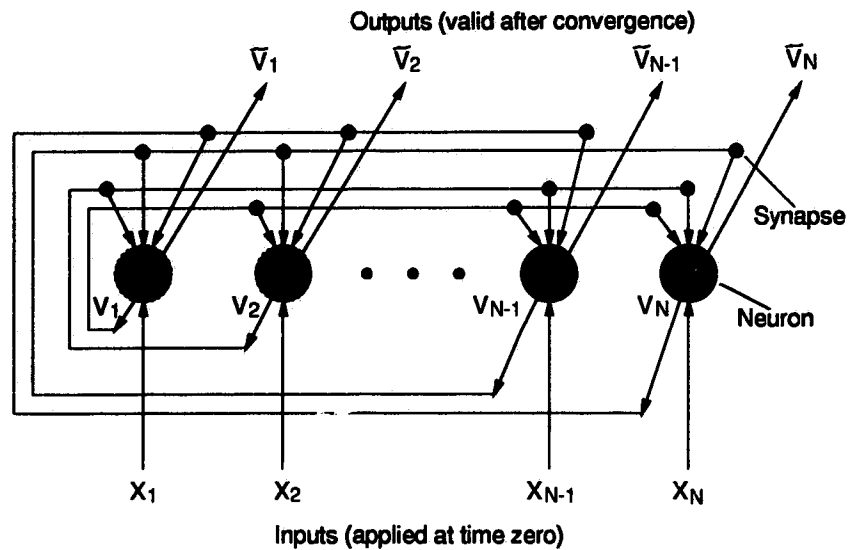


Figure 5.2: Hopfield ANN.

computation rate as provided by a massive parallelism and a robustness or fault-tolerance due to the fact that a few degraded or nonfunctional neurons will not greatly affect the overall operation of the network. The speed and the robustness of the neural networks make them very attractive for a variety of applications, such as pattern recognition, robotic control, and combinational optimization.

5.3 Mathematical Description of the Hopfield ANN

First, we define

V_j *Neuron State* : The neuron output after the activation function of the neuron has been performed.

u_i *Neuron Activity* : The input to the activation function, i.e., the total level of activity in neuron i mediated by the input stimuli and the interneural connections.

z_{ij} *Synaptic Weight Function* : The strength of connection between neurons i and j . Learning can change this term

A_i *Self-Term* : This term represents the passive decay of neural activity in the absence of both synaptic input and external input.

x_i *Initial Input* : This term can force an initial state on neuron i or can be switched off completely, to allow the network to settle.

B_{ij} *Forgetting Term* : This term represents the passive decay of the synaptic weight if B_{ij} is a constant. Memory loss is modulated if B_{ij} is variable.

V'_j *Neuron Learning Signal*: A different definition of the neuron state for learning purpose.

M The number of patterns (or vectors) to be learned by the network.

D_{ij} *Learning Strength* : This term allows learning to be modulated for each synaptic link.

f_θ *Hard-limiting Activation Function* : $f_\theta = 0$ if $u_i \leq \theta$, and $f_\theta = 1$ if $u_i \geq \theta$.

The dynamic behavior of the Hopfield ANN with N neurons is given by [50]:

$$\frac{\partial u_i}{\partial t} = -A_i u_i + \sum_{j=1}^N z_{ij} V_j + x_i \quad (5.1)$$

and

$$\frac{\partial z_{ij}}{\partial t} = -B_{ij}z_{ij} + D_{ij}V_j' f_{\theta}(u_i) \quad (5.2)$$

These two equations form a general description of the way the neuron states $\{V_i\}$ and the synaptic weights $\{z_{ij}\}$ develop in time.

The time evolutions of the state of the network may therefore be written as:

$$u_i(t + \Delta t) = u_i(t) + \Delta t \left[-A_i u_i(t) + \sum_{j=1}^N z_{ij} V_j(t) + x_i(t) \right] \quad (5.3)$$

If we choose $\Delta t = A_i^{-1}$, and let the synaptic weight $w_{ij} = z_{ij} A_i^{-1}$ and $x_i' = x_i(t) A_i^{-1}$, Eq. (5.3) becomes

$$u_i(t + \Delta t) = \sum_{j=1}^N w_{ij} V_j(t) + x_i'(t) \quad (5.4)$$

Hopfield applied the following restrictions on the above equation: $w_{ij} = w_{ji}$, $w_{ii} = 0$, and V_i is a binary value with no intermediate values. Hopfield also applied a hard-limiting activation function, f_{θ} with $\theta = 0$, to give a scheme for updating the neuron states $\{V_i\}$. The new neuron states $\{\tilde{V}_i\}$ is given by:

$$\tilde{V}_i = f_0 \left(\sum_{j=1}^N w_{ij} V_j + x_i'(t) \right) \quad (5.5)$$

Consider $\{x_i'\}$ to be the initial inputs to the network, i.e., $V_i(0) = x_i'$, the neuron states updating with respect to time is then given as:

$$V_i(t + 1) = f_0 \left(\sum_{j=1}^N w_{ij} V_j(t) \right) \quad (5.6)$$

If we expand Eq. (5.2) in the same way as we expand Eq. (5.1) to obtain Eq. (5.3), and consider f_{θ} is a hard-limiting activation function relating $\{V_i\}$ and $\{V_i'\}$ to the activities $\{u_i(t)\}$, we can obtain an expression for iterating w_{ij} ($= z_{ij} A_i^{-1}$)

$$w_{ij}(t + \Delta t) = w_{ij}(t) + \Delta t \left[-B_i w_{ij}(t) + E_{ij} V_i'(t) V_j'(t) \right] \quad (5.7)$$

where $E_{ij} = A_i D_{ij}$. Hopfield's learning strategy, or "storage prescription" therefore adjusts the synaptic weights $\{w_{ij}\}$ to store a set of M N -bit input vectors, or patterns, $\{V'_i(n)\}$. In other words, $\{V'_i(n)\}$ will become the stable states of the network. For the network to learn the states through Eq. (5.7), it must be held in each of the vectors $\{V'_i(n)\}$ for a fixed time Δt , while the synaptic weights evolve according to Eq. (5.7). If the synaptic weights are initially $\{w_{ij}(0)\}$, and are $\{w_{ij}(n)\}$ after exposure to the n th input vector, $V'_i(n)$, then Eq. (5.7) gives:

$$w_{ij}(n) = w_{ij}(n-1) + \Delta t \left[-B_{ij} w_{ij}(n-1) + E_{ij} V'_i(n-1) V'_j(n-1) \right] \quad (5.8)$$

If we define $\delta_{ij} = B_{ij} \Delta t$ and $\epsilon_{ij} = E_{ij} \Delta t$, this updating scheme gives a final form (once all M vectors $\{V'_i(n)\}$ have been presented to the network) as:

$$w_{ij}(M) = (1 - \delta_{ij})^M w_{ij}(0) + \sum_{n=1}^M V'_i(n) V'_j(n) \times \epsilon_{ij} (1 - \delta_{ij})^{M-n} \quad (5.9)$$

If we start with a set of null weights, $\{w_{ij}(0) = 0\}$, and there is no forgetting term in the learning equation, i.e., $B_{ij} = 0$, then for $\epsilon_{ij} = 1$ ($E_{ij}^{-1} = \Delta t$), Eq. (5.9) becomes

$$w_{ij}(M) = \sum_{n=1}^M V'_i(n) V'_j(n) \quad (5.10)$$

Note that Eqs. (5.6) and (5.10) define the neuron state updating and the synaptic weight learning in the Hopfield ANN, respectively.

5.4 Development of A Hybrid ANN Architecture

As can be seen from the preceding section, the Hopfield ANN is presented with a set of inputs $\mathbf{X} = \{x_i\}$, for which it generates a set of outputs $\mathbf{V} = \{V_i\}$. Assume the network consists of N binary neurons (two-state neurons) which are connected to each other by the synapses. The neuron states are then updated according to Eq. (5.6). The synaptic weights are given by an array \mathbf{W} , where w_{ij} is the analog

weight between the i th neuron and the j th neuron. If the input \mathbf{X} to the network is not one of its stored patterns, the network will be attracted by the dynamics to the stable state most similar to the input \mathbf{X} .

Most Hopfield ANN implementations have used a parallel architecture [49], as shown in Fig. 5.3. The neuron activation function is realized by an activation circuit which is usually a high-gain amplifier. Each neuron is connected to all other neurons by the multipliers acting as the synapses. The synaptic weights can be stored in

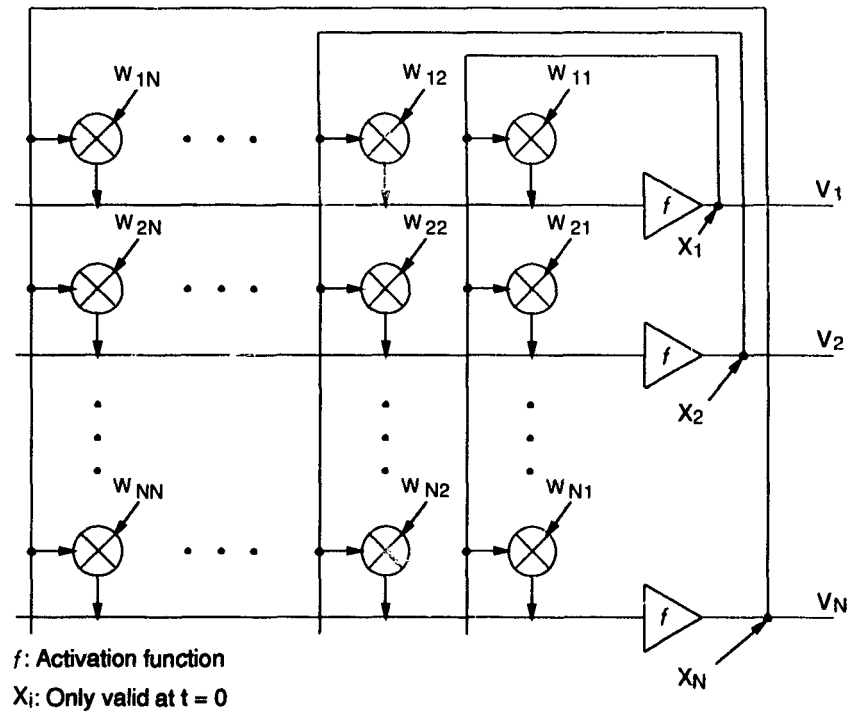


Figure 5.3: Parallel architecture of the ANN.

either digital or analog form, depending on the method of the realization of the synapses. All the neurons in the network update their states simultaneously. This architecture is most similar to the biological neural network. It should be pointed out, however, that the parallel approach consumes a large chip area not only for the activation circuits and the multipliers but also for the massive interconnections. With

the increase in the number of the neurons, it becomes more and more difficult to place the electronic neurons and route the interconnections two-dimensionally on a wafer. The massive interconnections in fact represent a major bottleneck for the parallel implementation. In addition, this approach suffers from the problem of significant power dissipation because in a network with N neurons there are N activation circuits and N^2 multipliers.

The other end of the spectrum is the serial architecture that only employs one multiplier, one adder, one activation circuit, and one shift-register, as shown in Fig. 5.4. The weight array is flattened into a N^2 long vector and stored in an analog shift-register. This approach is most economical in terms of power and area, but takes a longer computation time. The computation speed of the serial architecture is N^2 times slower than that of the parallel architecture.

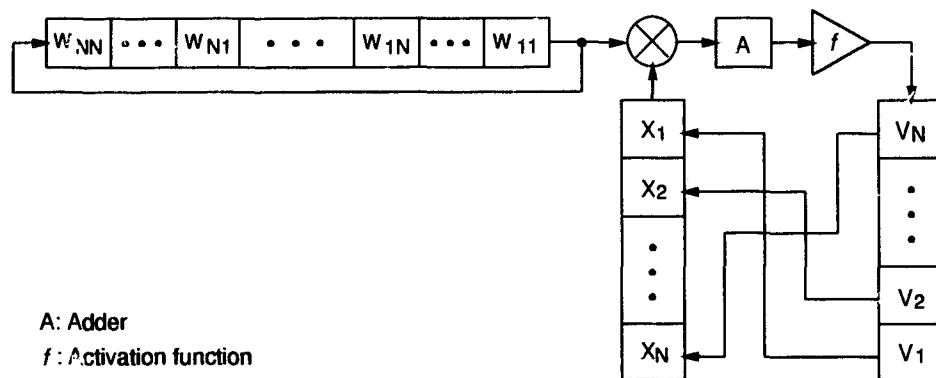


Figure 5.4: Serial architecture of the ANN.

By merging the parallel and serial architectures, we create a hybrid architecture, as shown in Fig. 5.5. This approach uses N multipliers, one adder, and one activation circuit which is multiplexed over all neurons. Therefore, the problems such as large power dissipation and chip area are obviated as in the serial approach, whereas the computation speed is reduced by N compared to N^2 in the serial approach. The

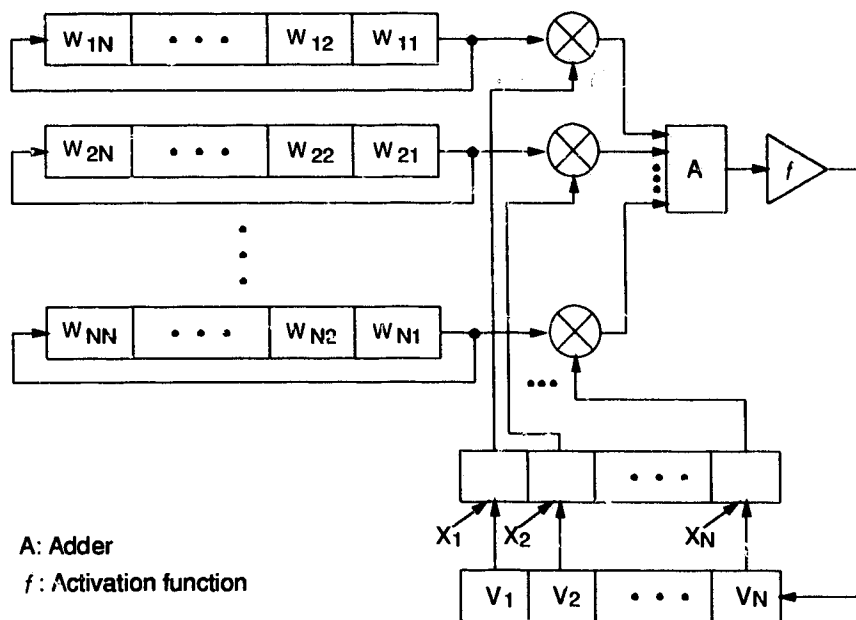


Figure 5.5: Hybrid architecture of the ANN.

weight-summation operations are completed in parallel for the individual neurons and in series for the other parts of the network. It can be seen that the hybrid architecture ranks first in power dissipation and chip area and second in computation speed among the three architectures.

5.5 Overview of the Hybrid ANN Architecture

Fig. 5.6 shows the block diagram of the proposed hybrid ANN architecture. In order to use the ANN to solve a particular function, we must first obtain the synaptic weights. The weights can be calculated by either on-line learning or off-line learning [52]. In on-line learning, the ANN is connected to a computer. The computer trains the ANN by introducing known input vectors to be stored in the network, evaluates the outputs and changes the synaptic weights. Eq. (5.7) describes this learning process.

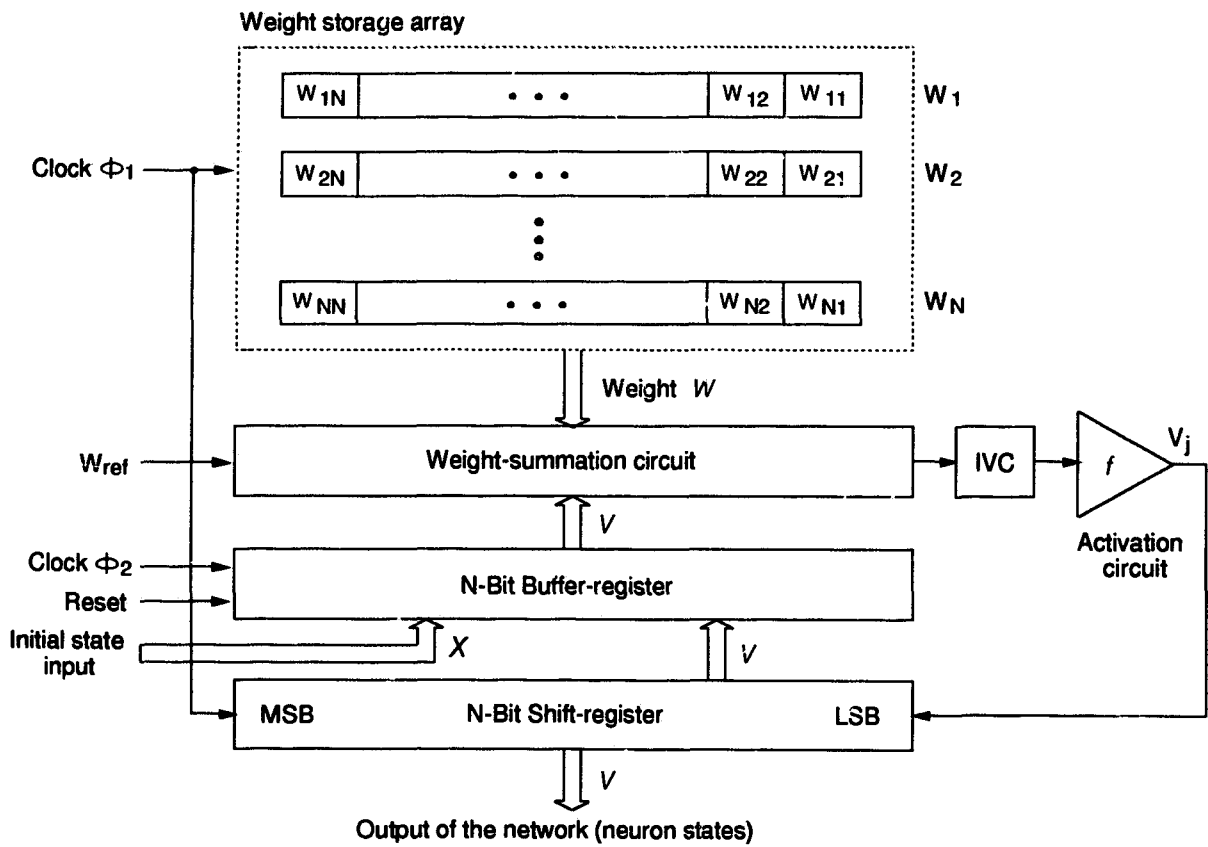


Figure 5.6: Block diagram of the proposed hybrid ANN architecture.

In off-line learning, the weights are computed directly using Eq. (5.10). After the weights are obtained, they are loaded into the network for operation. In general, on-line learning produces much better results than off-line learning, but need extra hardware and software for the learning process. On the other hand, off-line learning allows the designer to quickly change and modify the synaptic weights in the test of the ANN. Since our work is focused on the development of a hybrid ANN architecture rather than the ANN learning method, we use off-line learning to produce the synaptic weights in this work.

The synaptic weights are stored in the analog shift-register array which is composed of N self-connected N -bit shift-registers. It has been suggested that CCDs would be the best candidate for the analog shift-register and hence for the storage of the synaptic weights [37]. The weight w_{ij} may vary positively or negatively with respect to a reference w_{ref} . Note that in the hybrid ANN architecture the weight array is represented by the weight vector $\mathbf{W} = \{ W_1, W_2, \dots, W_i, \dots, W_N \}$, where W_i is $\{ w_{i1}, w_{i2}, \dots, w_{iN} \}$, and is electrically represented by a string of pulses.

The multiplication of the weight w_{ij} with the input V_i is realized by a transconductance amplifier. The N current components resulted from the multiplication are summed together concurrently. The total current is then converted to a voltage which is the input of the activation circuit. A N -bit binary shift-register is used to store and serially shift the updated neuron states (\mathbf{V}) generated from the activation circuit. A N -bit binary buffer-register is used to either set the initial neuron states (\mathbf{X}) or store the updated neuron states (\mathbf{V}) delivered in parallel from the N -bit shift-register.

As far as timing is concerned, the network is controlled by two groups of clocks. Clock ϕ_1 and its complement $\overline{\phi_1}$ control the weight storage array and the N -bit shift-register, whereas ϕ_2 controls the N -bit buffer-register. Fig. 5.7 shows the timing diagram. In the beginning of the operation, $\phi_1 = 0$ and $\phi_2 = 1$, the input of the initial neuron states are loaded into the buffer-register by the pulse *Reset*. During

the first clock pulse of ϕ_1 , the multiplication and summation is done, and the state of the first neuron is updated through the activation circuit and then stored in the shift-register. Because ϕ_2 is low, the outputs of the buffer-register are unchanged

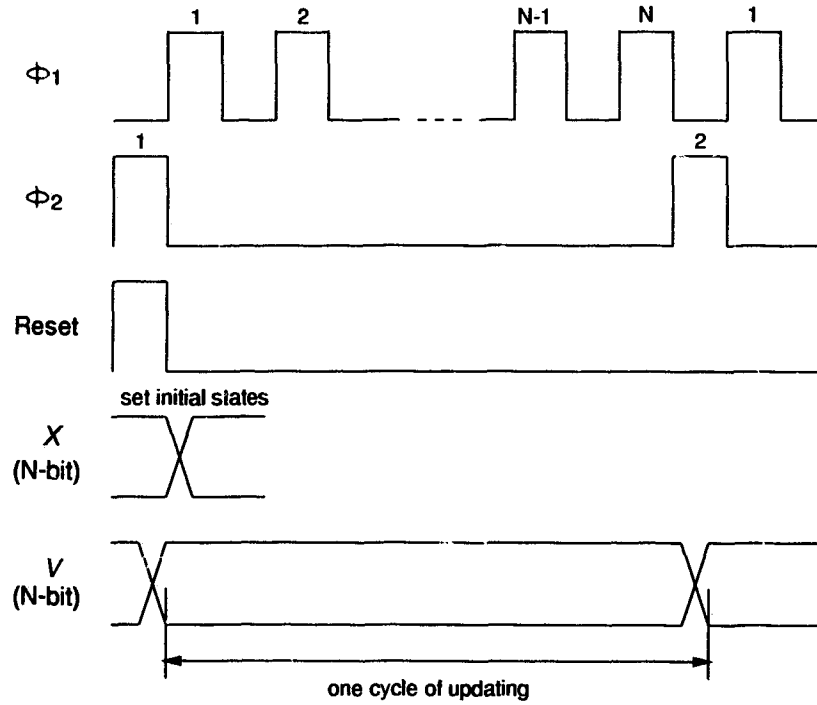


Figure 5.7: Timing diagram of the neuron states updating.

during this time. This process of computation controlled by ϕ_1 and $\bar{\phi}_1$ is repeated till the state of the N th neuron has been updated and shifted into the LSB (bit N) of the shift-register. After N pulses of ϕ_1 , ϕ_2 becomes high, and the updated neuron states are loaded in parallel into the buffer-register for the next cycle of computation and updating. It can be seen from Fig. 5.7 that the computation time is given by the inverse of the frequency of ϕ_2 which is $1/N$ times of ϕ_1 , i.e.,

$$\tau = \frac{N}{f_1} = \frac{1}{f_2} \quad (5.11)$$

Assuming the frequency of ϕ_1 is 100 MHz which is not difficult to achieve for a GaAs

CMCCD and 1000 neurons are integrated in the network, the speed of one cycle of computation as determined by ϕ_2 is 10^5 times per second.

Chapter 6

GaAs Analog Circuit Design of Hybrid ANN Architecture

6.1 Introduction

The hybrid ANN architecture basically consists of a weight storage array formed by N self-connected N -bit serial analog memories, a weight-summation circuit composed of N transconductance amplifiers, an activation circuit with a threshold function, a N -bit shift-register, and a N -bit buffer-register. This chapter describes the GaAs circuit design of the analog part of the hybrid ANN architecture, i.e., (1) the CMCCD analog memory, (2) the weight-summation circuit, and (3) the activation circuit. The experimental results of the fabricated circuits are also described for each design.

6.2 CMCCD Analog Memory

6.2.1 Structure and Operation

The basic features of the CMCCD analog memory loop are illustrated in Fig. 6.1. The analog signal is represented by the charge packets in the potential wells in the CMCCDs. The full-well determines the dynamic range of the signal. The analog signal detected at the CCD output may be at a different dc level compared to the input signal. Thus, a level-shifting circuit is required. A 2-to-1 multiplexer is incorporated in the loop to control the write and read operations.

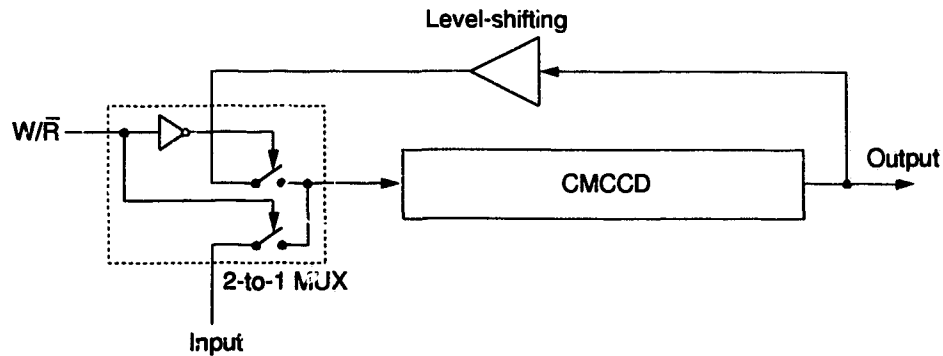


Figure 6.1: CMCCD analog memory.

Fig. 6.2 shows the timing diagram of the CMCCD analog memory. During the write period T_w , W/\bar{R} is high, and the analog signal is transferred into the CMCCD serially. When W/\bar{R} is low, the stored data is read out at the output of the CCD and fed back to the input of the CMCCD after its level-shifting. In order for the CMCCD memory to store one bit data per pixel, the write time T_w is chosen to be T_{tr} , equal to the one cycle time for the charge packet to transfer from the input to the output of the CMCCD. Note T_{tr} is the same as delay time (t_{delay}) given in Chapter 4. If we assume that both charge injection and detection of the CMCCD take place in the negative cycle of the clock (as described in Chapter 2), the clock feedthrough

at the output of the CMCCD memory will not affect the operation of the CMCCD memory. On the other hand, only the speed of the level-shifting circuit and the 2-to-1 multiplexer will affect the speed of the memory system.

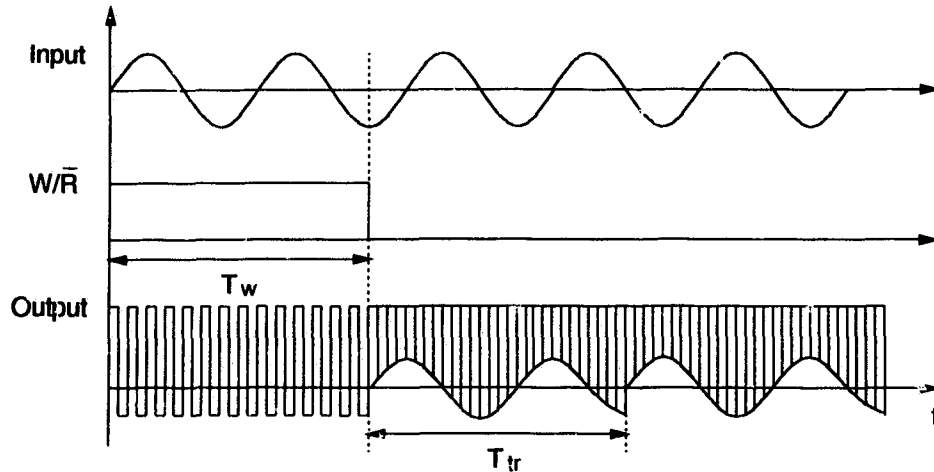


Figure 6.2: Timing diagram of the CMCCD analog memory.

6.2.2 Performance Evaluation

Intrinsically a GaAs CMCCD has very low noise and almost perfect signal linearity. In many cases, the noise and the linearity of the GaAs CMCCD analog memory are determined by the input and the output circuitry. Although the level-shifting circuit and the 2-to-1 multiplexer may also contribute noise and nonlinearity to the memory system, the following discussion only focuses on the noise and the linearity of the CMCCD itself.

Noise in a CMCCD can be separated into four components, i.e., (1) thermal generation noise (or dark current noise), (2) bulk trapping noise (or fast interface state noise), (3) input noise and (4) output noise (or *KTC* noise). The first two types of noise are associated with charge transfer and the last two are associated with charge

injection and detection. Appendix D gives the detailed analysis on the CMCCD noise. Using the parameters obtained for the two-phase and the uni-phase CMCCDs, the calculated noise for the two-phase and the uni-phase CMCCDs at 100 MHz are 273 electrons and 389 electrons, respectively. The maximum charge handling capabilities of the two-phase and the uni-phase CMCCDs (from Appendix D) are 8.6×10^5 electrons and 6.0×10^5 electrons, respectively. If we define the signal-to-noise ratio of the CMCCD as:

$$S/N = 20 \log(Q_{max}/Q_{noise}) \quad (6.1)$$

then the S/Ns are 67 dB and 63 dB for the two-phase and the uni-phase CMCCDs, respectively.

Linearity is another factor that affects the CMCCD as an analog memory. The major source of nonlinearity in a CMCCD is the variation in the depletion capacitances associated with the input gate. For the charge injection and detection methods described of Chapter 2, the voltage-to-charge and charge-to-voltage conversions are given by:

$$Q_{SIG} = C_{inp}(V_{ref} - V_{in}) \quad (6.2)$$

$$V_{out} = V_P + \frac{Q_{SIG}}{C_{op}} \quad (6.3)$$

where V_P is the precharged voltage at the output ohmic contact, and C_{op} is the parasitic capacitance at the output ohmic contact. C_{op} is given by [53]:

$$C_{op} = C_{GD} + (1 - G_v)C_{GS} \quad (6.4)$$

where G_v is the voltage gain of the output buffer of the CMCCD. C_{inp} is the depletion capacitance of a Schottky barrier (gate G_1) and it depends on the applied voltage. The relation between V_{in} and C_{inp} is given by [65]:

$$C_{inp} = A_{in} \sqrt{\frac{qN_d\epsilon_s}{2(V_b - V_{in})}} \quad (6.5)$$

where A_{in} is the area of the input gate, V_b is the built-in voltage of GaAs, and N_d is the donor density of the active layer of the CMCCD channel. Assuming V_{ref} is zero and using Eqs. (6.5) and (D.8) we can rewrite Eqs. (6.2) and (6.3) as:

$$Q_{SIG} = -V_{in} A_{in} \sqrt{\frac{q N_d \epsilon_s}{2(V_b - V_{in})}} \quad (6.6)$$

$$V_{out} = V_P + \frac{Q_{SIG}}{C_{GD} + (1 - G_v) C_{GS}} \quad (6.7)$$

The relationship between V_{in} and V_{out} can be obtained by solving Eqs. (6.6) and (6.7), and is plotted in Fig. 6.3 for different width of MESFET in the output buffer. The parameters used in the computation are listed in Table 6.1. It can be seen that by

Parameter	Unit	Value
A_{in}	μm^2	500
q	C	1.6×10^{-19}
ϵ_s	F/cm	116×10^{-14}
V_b	V	0.7
V_P	V	3
N_d	/cm ³	5×10^{16}
C_{GD}	fF/ μm	1.5
C_{GS}	fF/ μm	1.5
G_v		0.85

Table 6.1: Parameters used in computing CMCCD linearity.

changing the width of the buffer MESFET so that C_{inp} equals C_{op} , the nonlinearity can be reduced to minimum. In this design, when the width of the buffer MESFET is around 50 μm , the best linearity can be achieved.

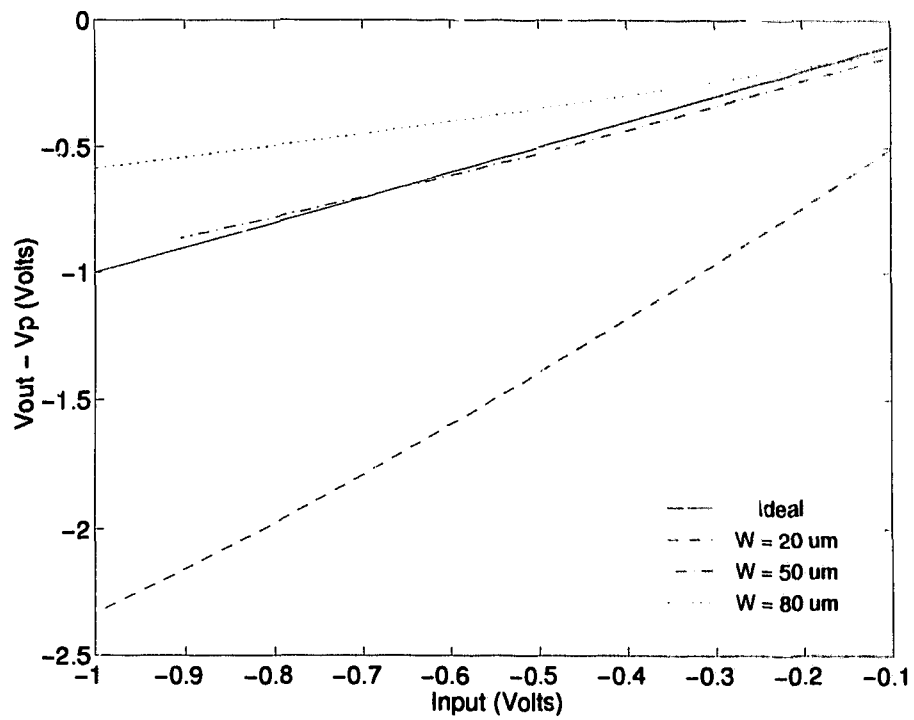


Figure 6.3: V_{out} v.s. V_{in} of the CMCCD for different width of MESFET in the output buffer.

6.2.3 Experimental Results

The operation of the CMCCD analog memory was verified by measuring the uni-phase CMCCD in a close-loop form. Fig. 6.4 shows the test setup. A IH5352 quad switch IC chip was used to realise the 2-to-1 multiplexer, and a NE592 IC chip was used in the level-shifting circuit. Fig. 6.5 displays the measured waveforms which show the proper operation of the analog memory. High speed testing (beyond 10 MHz) was limited by the performance of the NE592 and IH5352 ICs at high frequency.

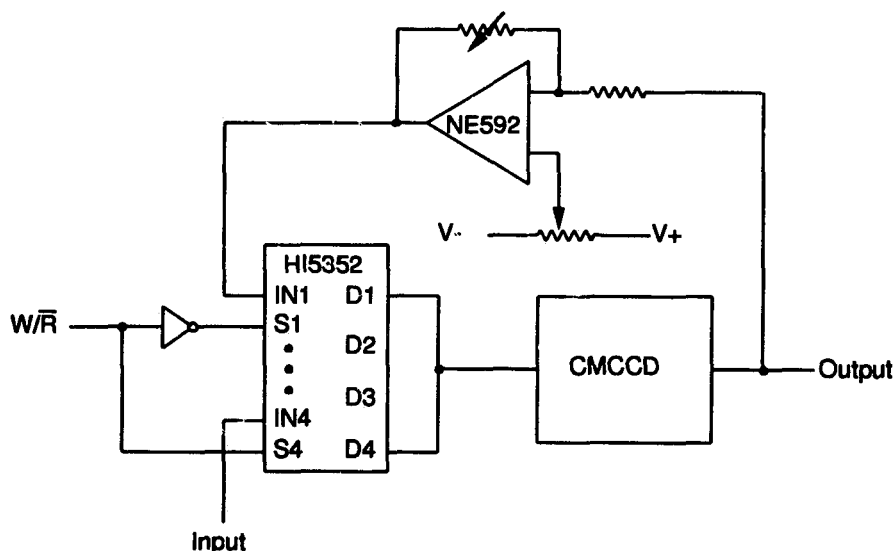


Figure 6.4: Test setup of the CMCCD analog memory.

The signal-to-noise ratio of the uni-phase CMCCD was measured using the tangential method [67] for different clock frequencies. Fig. 6.6 shows the signal-to-noise ratio in dB with respect to the clock frequency when the amplitude of the input signal is 0.5 V. As expected, the difference in the signal-to-noise ratio between the measurement and the theoretical calculation is very small.

The overall linearity between the input and the output of the CMCCD was also measured. The input of the CMCCD is a square waveform with one-half the clock fre-

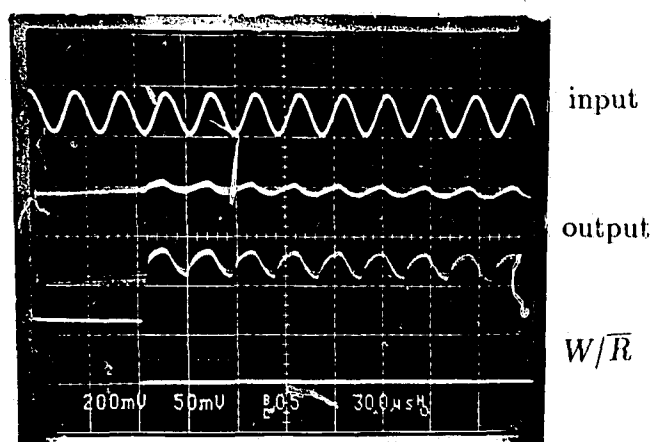


Figure 6.5: Oscillograph of the input, control and output waveforms of the CMCCD analog memory.

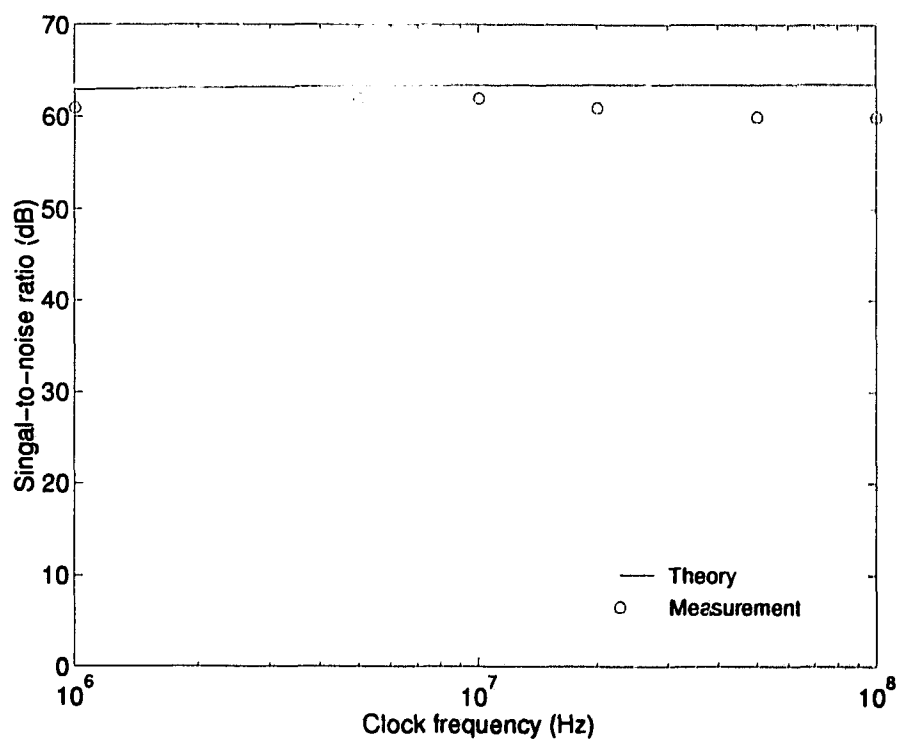


Figure 6.6: Signal-to-noise ratio of the CMCCD with respect to the clock frequency obtained from the measurement and the calculation.

quency. Fig. 6.7 shows the oscillograph of the CMCCD for the linearity measurement, where the upper and lower traces display the input and output waveforms, respectively. By changing the amplitude of the input square waveform, we measured the amplitude of the output waveform in the negative cycle of the clock. The measured input-output relationship for the different clock frequency are plotted in Fig. 6.8. It can be seen from Fig. 6.8 that the linearity of the CMCCD is satisfactory when the amplitude of the input is within 0.7 V and the clock frequency is within the range from 1 MHz to 50 MHz. Since the minimum amplitude of the output voltage of the HP-8112A pulse generator is 10 mV, the bottom line of the CMCCD output that can be measured is therefore around 10 mV. This gives a dynamic range of the CMCCD of at least 38 dB.

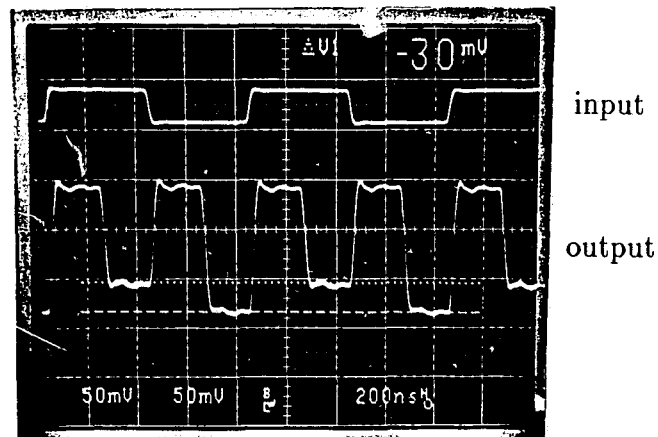


Figure 6.7: Oscillograph of the CMCCD for the linearity measurement.

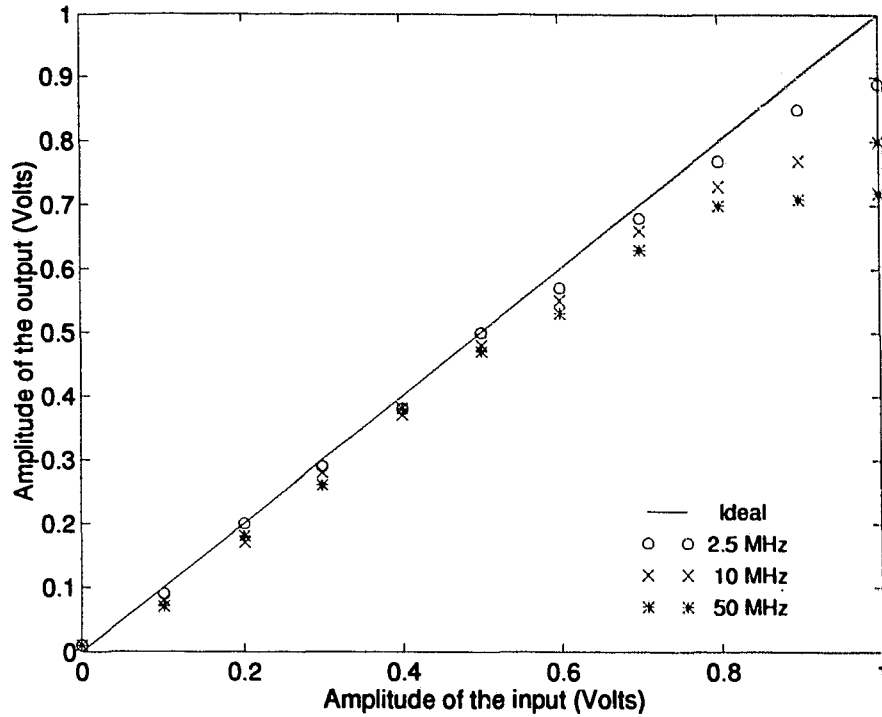


Figure 6.8: Measured linearity between the input and output of the CMCCD.

It should be pointed out, however, that the CCD analog memory shown in Fig. 6.1 may have poor performance limited by the CTE of the CCD, the linearity of the level-shifting circuit, and the performance of the switches in the multiplexer. This can be observed from the output waveform of Fig. 6.4, where the output signal has a distortion compared to the input signal. Therefore, only multiple-level signals (for example, 4 levels) could be stored and transferred in the CCD analog memory and a regeneration circuit may be required [25].

6.3 Weight-summation Circuit

6.3.1 Circuit Design

The conventional design of the weight-summation circuits (WSCs) is to source or sink the currents of $W_i V_i$ to the single input of the activation function [29]. This kind of design, however, suffers from the problem of summing the noise from each multiplier in the WSC and hence reducing the signal-to-noise ratio of the output of the WSC. To overcome this drawback, we developed a WSC using two distinct lines to sum the currents: one for excitatory current and the other for inhibitory current. The noise generated from these current lines can be cancelled at the differential input of the activation circuit.

Fig. 6.9 shows the block diagram of the WSC which is composed of N transcon-

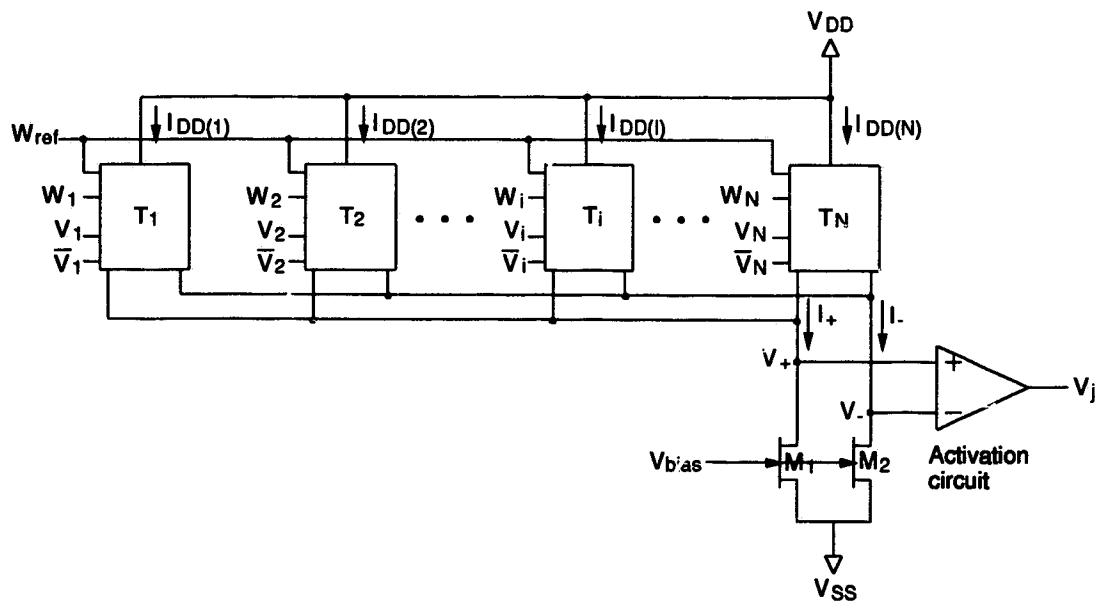


Figure 6.9: Block diagram of the weight-summation circuit.

ductance amplifiers acting as the multipliers. The result of the multiplication of

the neuron state V_i and the synaptic weight W_i is in the form of current. Each transconductance amplifier generates two distinct currents (i.e., excitatory current and inhibitory current). These two currents are summed separately. The summed currents, I_+ and I_- , are then converted into the corresponding voltages V_+ and V_- across the MESFETs M1 and M2. These voltages are compared at the differential inputs of the activation circuit.

Fig. 6.10 shows the schematic of a 6-transistor transconductance amplifier where all MESFETs are assumed to be of the same size except for the active load M_L . M1 and M2 are controlled by the analog synaptic weight W_i and weight reference w_{ref} . M3 ~ M6 are controlled by the binary neuron states V_i and its complement \bar{V}_i . To make the circuit work properly, M1 and M2 must operate in the saturation region and M3 ~ M6 are either in the linear region or in the cutoff region. Assume $V_i =$

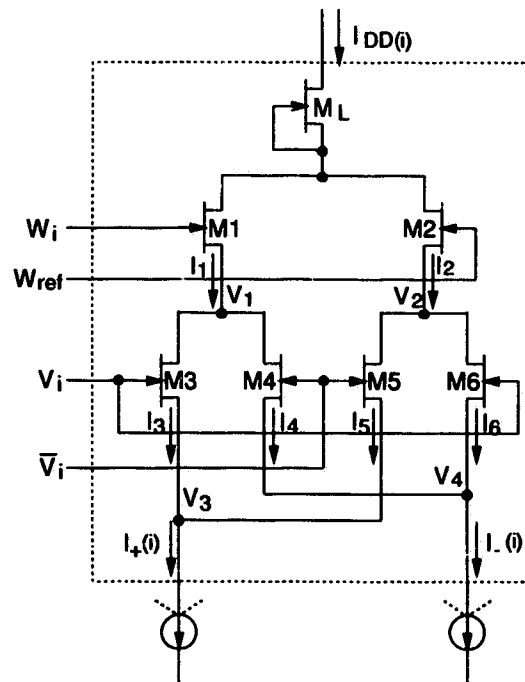


Figure 6.10: Schematic of the transconductance amplifier.

1, M3 and M6 are in the linear region and M4 and M5 are in the cutoff region. The branch currents are given by:

$$I_1 = \frac{\beta}{2}(W_i - V_1 - V_T)^2 \quad (6.8)$$

$$I_2 = \frac{\beta}{2}(w_{ref} - V_2 - V_T)^2 \quad (6.9)$$

$$I_3 = \beta(V_i - V_3)(V_1 - V_3) \quad (6.10)$$

$$I_6 = \beta(V_i - V_4)(V_2 - V_4) \quad (6.11)$$

$$I_+(i) = I_3 = I_1 \quad (6.12)$$

$$I_-(i) = I_6 = I_2 \quad (6.13)$$

where β is a transconductance parameter. The differential current $I_+(i) - I_-(i)$ can be derived from the above equations and is given approximately by:

$$I_+(i) - I_-(i) \approx \frac{3\beta}{4}(W_i - w_{ref})\sqrt{\frac{4I_{DD}(i)}{\beta} - \left[\frac{3}{2}(W_i - w_{ref})\right]^2} \quad (6.14)$$

for $V_i = 1$. Similarly, we can obtain

$$I_+(i) - I_-(i) \approx -\frac{3\beta}{4}(W_i - w_{ref})\sqrt{\frac{4I_{DD}(i)}{\beta} - \left[\frac{3}{2}(W_i - w_{ref})\right]^2} \quad (6.15)$$

for $V_i = -1$. It is observed from Eqs. (6.14) and 6.15 that if $(W_i - w_{ref})^2 \ll 4I_{DD}(i)/\beta$, then the transconductance amplifier will perform four-quadrant multiplication. Fig. 6.11 shows a plot of the I - V characteristics of the transconductance amplifier resulting from Eqs. (6.14) and (6.15) and from HSPICE simulation where $I_{DD}(i)$ was $100 \mu\text{A}$. Note that to make the WSC work properly and accurately, the currents flowing into the transconductance amplifiers should be the same. This can be achieved by keeping the sizes of the load MESFETs the same.

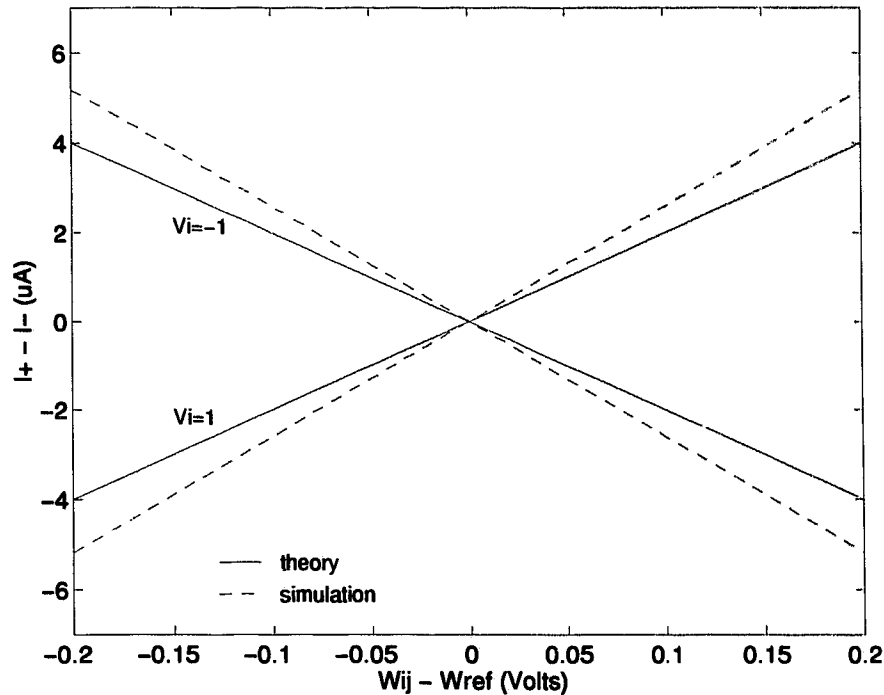


Figure 6.11: I - V characteristics of the transconductance amplifier.

6.3.2 Performance Evaluation

A WSC with 16 transconductance amplifiers was simulated using HSPICE incorporated with the device models of the NT/BNR (Northern Telecom/Bell Northern Research) $0.8 \mu\text{m}$ GaAs depletion-mode MESFET process. The analog synaptic weights change between 0.2 V and 0.8 V and w_{ref} is 0.5 V . The voltage level of the neuron state V_i is either 0 V or -4 V to ensure that the MESFETs connected to V_i and \bar{V}_i are either cutoff or in the linear region. The widths of the MESFETs as determined by HSPICE simulation are $8 \mu\text{m}$ for the MESFETs in the 6-transistor transconductance amplifier and $6 \mu\text{m}$ for each of the load MESFETs. Fig. 6.12 shows one group of the waveforms of W_i and V_i used in the simulation. The simulation waveforms of the output currents of the WSC are shown in Fig. 6.13 where the solid-line represents I_+ and the dashed-line represents I_- . As can be seen from Fig. 6.13, the linearity of the

inner product computation is quite good. The power dissipation of the WSC is 21 mW according to the simulation when V_{DD} and V_{SS} are 5 V and -2 V, respectively.

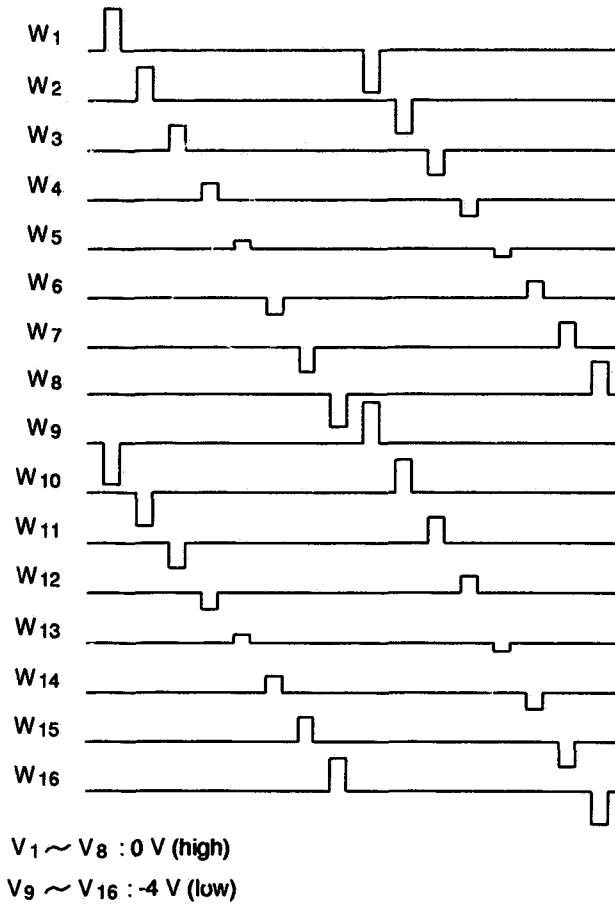


Figure 6.12: Waveforms of W_i and V_i used in the WSC simulation.

The design and simulation of the WSC were based on the assumption that all the transconductance amplifiers are perfectly matched and the load MESFETs are of the same size. It should be pointed out, however, the process spread will cause some variations in size even for a set of identical MESFETs. In order to investigate this effect, we simulated the WSC by randomly changing the sizes of the MESFETs — positively or negatively — $1 \mu\text{m}$ from their standard sizes. It was found that

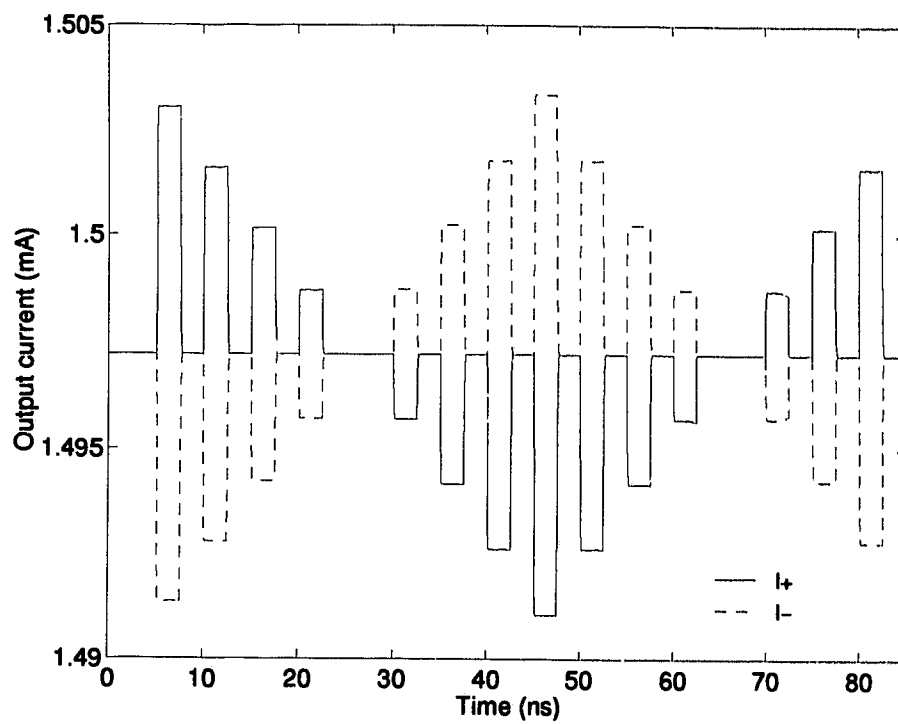


Figure 6.13: Simulation waveforms of the output currents of the WSC.

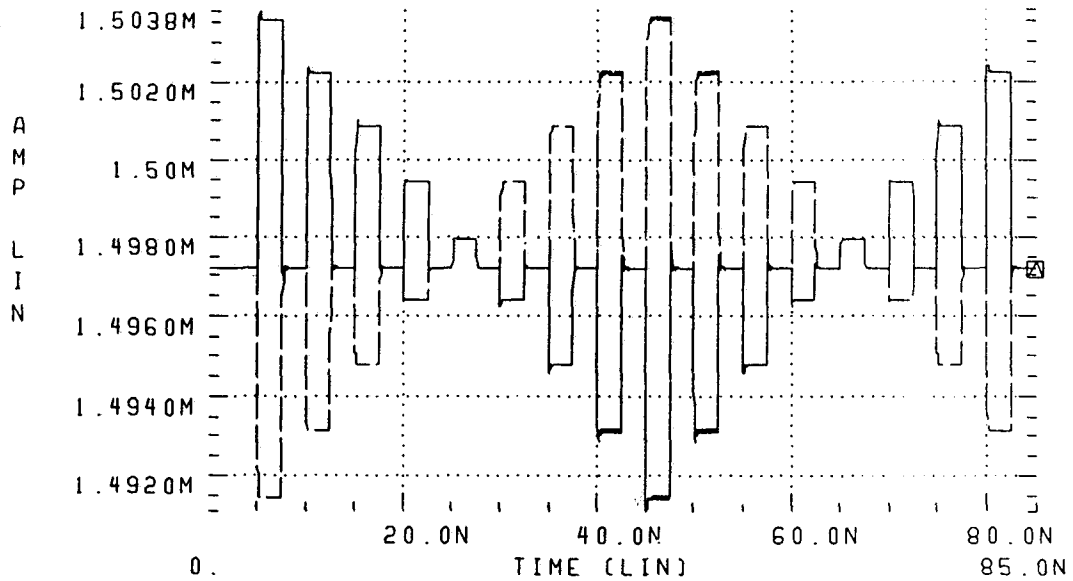


Figure 6.14: Simulation waveforms of the output currents of the WSC when the load MESFETs are not the same.

the linearity of the computation for the WSC were degraded by an average of 7%. Fig. 6.14 shows the simulation waveforms of the output currents I_+ and I_- of the WSC when the load MESFETs are not the same.

It has been noted that for a given V_{DD} and V_{SS} , increasing the number of transconductance amplifiers will cause an increase in the currents I_{DD} , and hence I_+ and I_- . The increase in I_+ and I_- will in turn increase the voltage drops V_+ and V_- across the MESFETs M1 and M2 in Fig. 6.9 because M1 and M2 are not the ideal current sources. These increases in V_+ and V_- will reduce the output dynamic range of the activation circuit due to the common-mode voltage of the activation circuit (will be described in the next section). In order to determine how the number of transconductance amplifiers affect the output voltage of the activation circuit, a HSPICE simulation was performed under the worst-case condition, i.e., when $N/2 - 1$ multipliers are connected to the “on”-neurons and $N/2 + 1$ to the “off”-neurons (i.e.,

$I_- > I_+$), or the reverse (i.e., $I_+ > I_-$), while the weights are kept constant. Fig. 6.15 shows the simulation result. When the number of transconductance amplifiers increases to 27, the dynamic output range of the activation circuit decreases to about 1 V, which is still acceptable for the input of the shift-register. However, if the number of transconductance amplifiers increases to 30, the dynamic range shrinks to less than 0.5 V. To make the WSC function properly for a large number of transconductance amplifiers (more than 28), the MESFETs have to be resized.

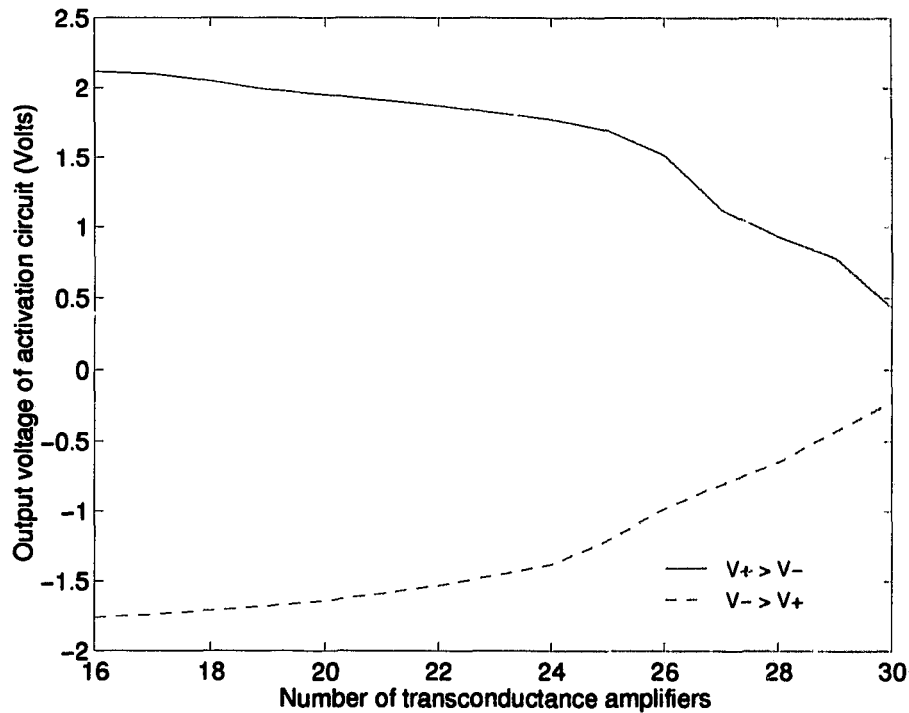


Figure 6.15: Output voltage of the activation circuit with respect to the number of the transconductance amplifiers.

6.3.3 Experimental Results

We fabricated a WSC with 16 multipliers using the NT/BNR 0.8 μm GaAs depletion-mode MESFET process. Fig. 6.16 shows the microphotograph of one of the transconductance amplifiers. The performance of the WSC was measured in terms of the accuracy of the inner product computation. In the measurement, the synaptic weights were binary (i.e., $W_i - w_{ref} = 0$ or 1). This is sufficient to show the function of the

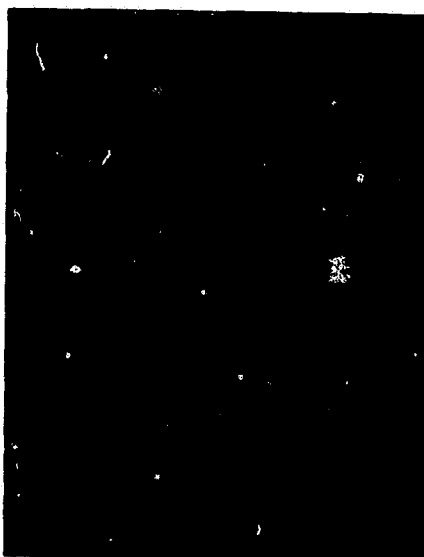
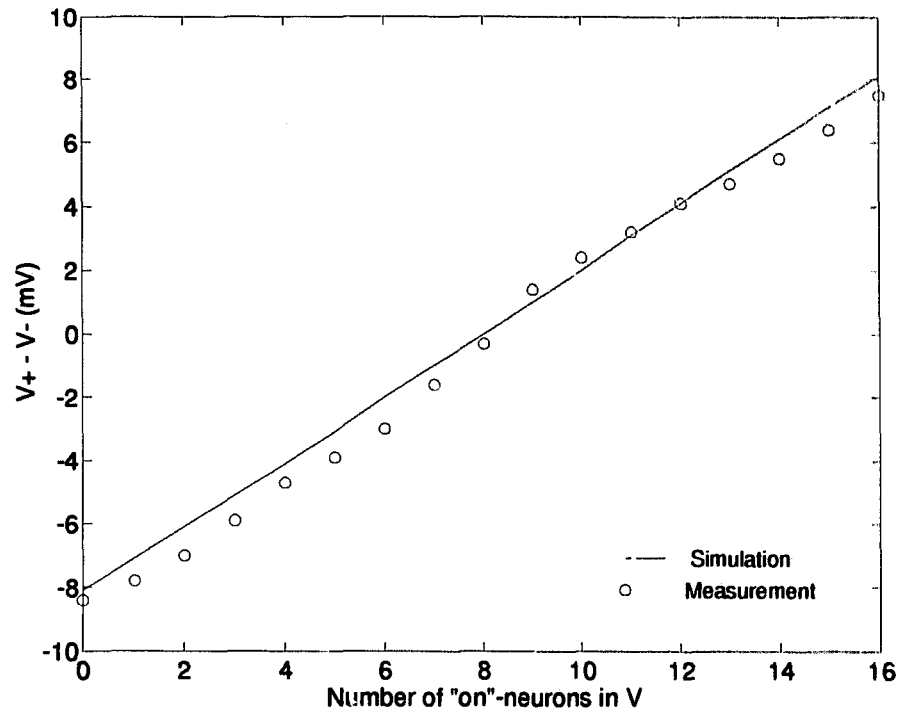


Figure 6.16: Microphotograph of one transconductance amplifiers in the WSC.

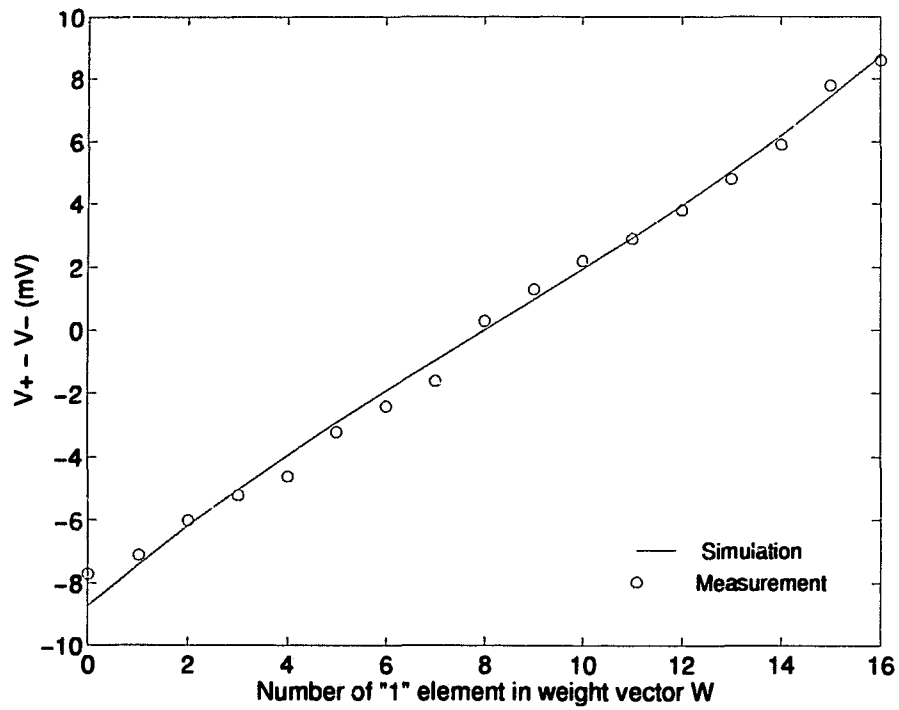
WSC although in general the weights can also be analog, as we shall see in Chapter 8. In measuring the linearity of the weight-summation operation, we kept one of the vectors, \mathbf{V} or \mathbf{W} , constant, and changed the elements in the other vector to a high voltage level one at a time until all the elements are high. Fig. 6.17(a) shows the result of a constant vector of weights \mathbf{W} multiplied by \mathbf{V} . The output of the WSC ($\Delta V = V_+ - V_-$) is a function of the number of “on”-neurons in \mathbf{V} . Fig. 6.17(b) shows the result of a constant neuron output \mathbf{V} multiplied by a vector of weights \mathbf{W} . The output of the WSC is a function of the number of 1’s in \mathbf{W} while \mathbf{V} is

constant. Fig. 6.17 shows that the measured linearity of the WSC has some random divergences. The divergences are due to the variations in the sizes of the devices in the fabrication process. They can also be observed in the oscillographs we obtained from the ANN chip, which will be described in Chapter 8.



(a)

Figure 6.17: Measured and simulated linearity of the WSC for binary weights.



(b)

Figure 6.17: Measured and simulated linearity of the WSC for binary weights (continued).

6.4 Activation Circuit

In general, there are three common activation functions for the artificial neurons, namely, the hard-limiting function, the linear function, and the sigmoid function. Gain-adjustable function has also been proposed to apply the ANN to the optimization problems [51]. When the Hopfield ANN was used as an associative memory, the activation function is usually a hard-limiting function [38]. In this work, therefore, we assumed the neurons have hard-limiting activation function.

6.4.1 Circuit Design

The activation circuit required to implement a hard-limiting function should ideally have an infinite gain. For the GaAs MESFET technology, however, a high-gain amplifier is generally difficult to achieve because of the inherently low transconductance and the low output resistance of the MESFET. To obtain high-gain within a wide frequency range, we used an operational amplifier configuration for the design of the activation circuit. Fig. 6.18 shows the schematic of the activation circuit which has an input stage, a gain stage and an output stage. HSPICE simulation was used to determine the sizes of the MESFETs. Frequency compensation was achieved by addition of an on-chip capacitor C_c ($= 0.6$ pF) around the gain stage (see Appendix E).

6.4.1.1 Input Stage

In order to compare the currents I_+ and I_- from the WSC, a differential amplifier is used as the input stage of the activation circuit. For a basic common-source differential amplifier with small MESFETs as active loads, the differential-mode voltage gain is approximately given as:

$$G_d \approx -\frac{1}{2}g_m r_{ds} \quad (6.16)$$

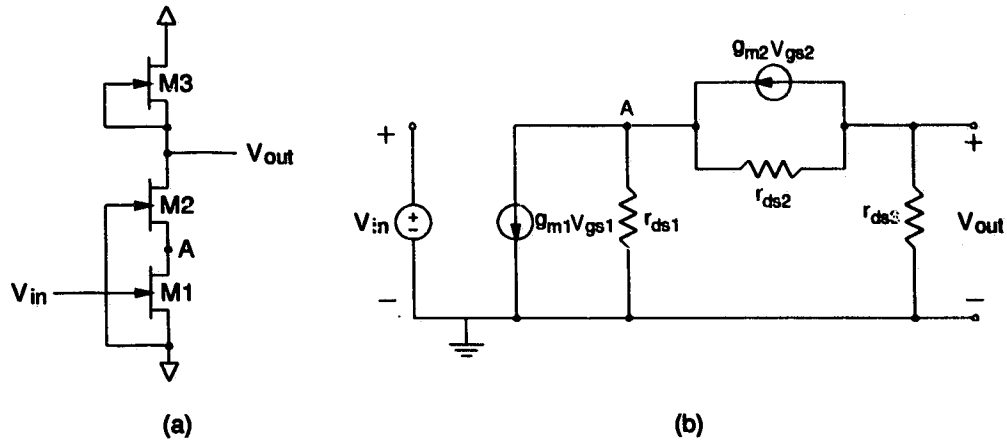


Figure 6.19: (a) MESFET cascode circuit and (b) its small-signal equivalent circuit.

given by:

$$G_v = \frac{v_{out}}{v_{in}} = -\frac{r_{ds3}(g_{m1}r_{ds1} + g_{m1}g_{m2}r_{ds1}r_{ds2})}{r_{ds1} + r_{ds2} + r_{ds3} + g_{m2}r_{ds1}r_{ds2}} \quad (6.17)$$

Note that the voltage gain of the differential amplifier formed by the cascode configuration is given by $G_v/2$. To minimize the power dissipation, we restrict the current of the power supply by using a narrow MESFET M3 as the active load. The output resistance of M3 is then much greater than those of M1 and M2, and the differential gain of the cascode configuration can be approximated by:

$$G_d \approx -\frac{1}{2}g_{m1}r_{ds1}(1 + g_{m1}r_{ds1}) \quad (6.18)$$

where M1 and M2 are assumed to be of the same size. It can be seen from Eq. (6.18) that the differential gain of the cascode configuration is significantly increased compared to that of the common-source configuration. In addition to the increase of the differential gain, the cascode configuration offers a good high-frequency performance by lowering the Miller capacitance multiplication effect.

The differential input stage shown in Fig. 6.19 uses the cascode MESFETs M1 ~ M4 as the differential pair, M5 and M6 as the active loads, and a source follower (M10

and M11) as a level-shifting circuit to drive the gain stage consisting of normally-on MESFETs. Local feedback through M8 and M9 keeps the amplifier biased in the high-gain region by dynamically adjusting the current source M7 to compensate for the process variation and it also provides differential to single-ended conversion [72]. The differential gain of the input stage is given by Eq. (6.18).

6.4.1.2 Gain Stage

A current bleeder with positive feedback as proposed in [73] was used in the gain stage, as shown in Fig. 6.19. The voltage gain of this stage is given by:

$$G_g = \frac{v''}{v'} \approx \frac{g_{m12}r_{ds14}}{1 - r_{ds14}/r_{ds16}} \quad (6.19)$$

where v'' and v' are labelled in Fig. 6.19. It is theoretically possible to make M14 and M16 of the same size and obtain infinite gain. However, positive feedback may cause self-oscillation and increase the output impedance of the gain stage.

6.4.1.3 Output Stage

A source follower with serially-connected Schottky diodes was used as the output stage of the activation circuit to provide a low output impedance and a level-shifting function. To increase the output voltage swing of the activation circuit and reduce the offset voltage, V_{SS1} , a negative power supply which is more negative than V_{SS} was used in the output stage. The output resistance and the voltage gain of the output stage are approximately given by:

$$r_{out} \approx \frac{1}{g_{m17}} \quad (6.20)$$

$$G_o = \frac{v_{out}}{v''} \approx \frac{1}{1 + (1/g_{m17}r_{ds17}) + (1/g_{m17}r_{ds18})} \quad (6.21)$$

It can be seen from Eq. (6.21) that if $g_{m17} \gg 1/r_{ds17} + 1/r_{ds18}$, unity gain will be achieved. For a given quiescent point, however, if g_m is increased by increasing the

width of the MESFET, r_{ds} will be decreased simultaneously. In this design, the gain of the output stage is about 0.8.

6.4.2 Experimental Results

The activation circuit was simulated by HSPICE and fabricated using the NT/BNR 0.8 μm GaAs depletion-mode MESFET process. Fig. 6.20 shows the microphotograph of the activation circuit. Fig. 6.21 shows the oscillographs of the input and output waveforms of the activation circuit operating at 20 MHz and 100 MHz. The dc

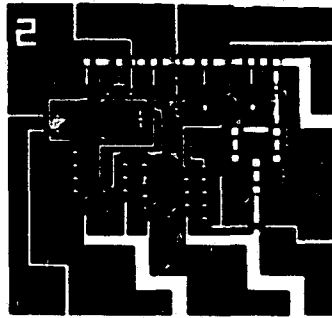


Figure 6.20: Microphotograph of the activation circuit.

transfer characteristics of the activation circuit obtained from the measurement and the simulation are shown in Fig. 6.22. It can be seen that the output swing is larger than ± 1.7 V when V_{DD} , V_{SS} and V_{SS1} are +5 V, -2 V and -3 V. The input offset voltage is about 50 mV. Fig. 6.23 shows gain responses obtained from the measurement and the simulation. The gain is 50 dB at low frequency and drops to 40 dB at 5 MHz. The unity-gain frequency is beyond 500 MHz. Table 6.2 summarizes the performance of the activation circuit.

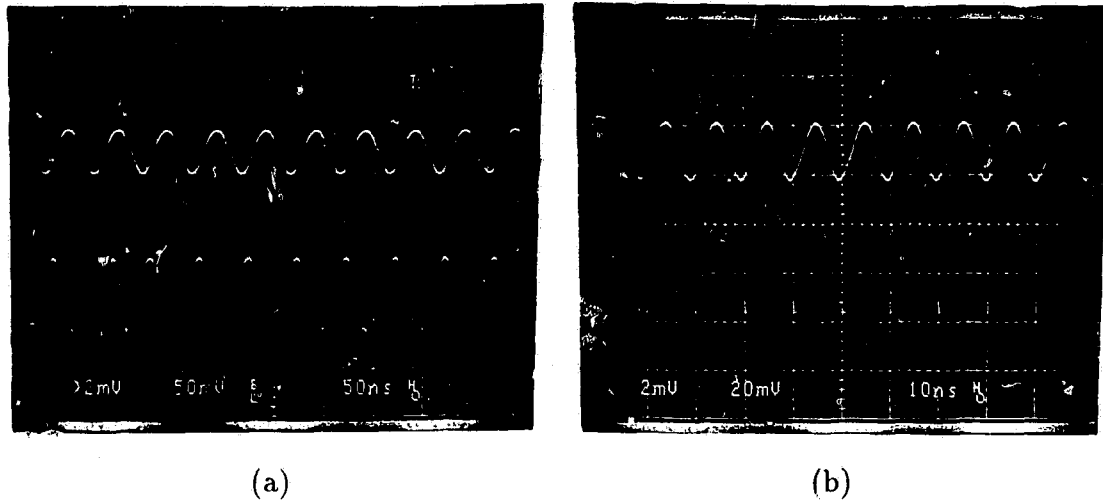


Figure 6.21: Oscillographs of the input and the output waveforms of the activation circuit at (a) 20 MHz and (b) 100 MHz.

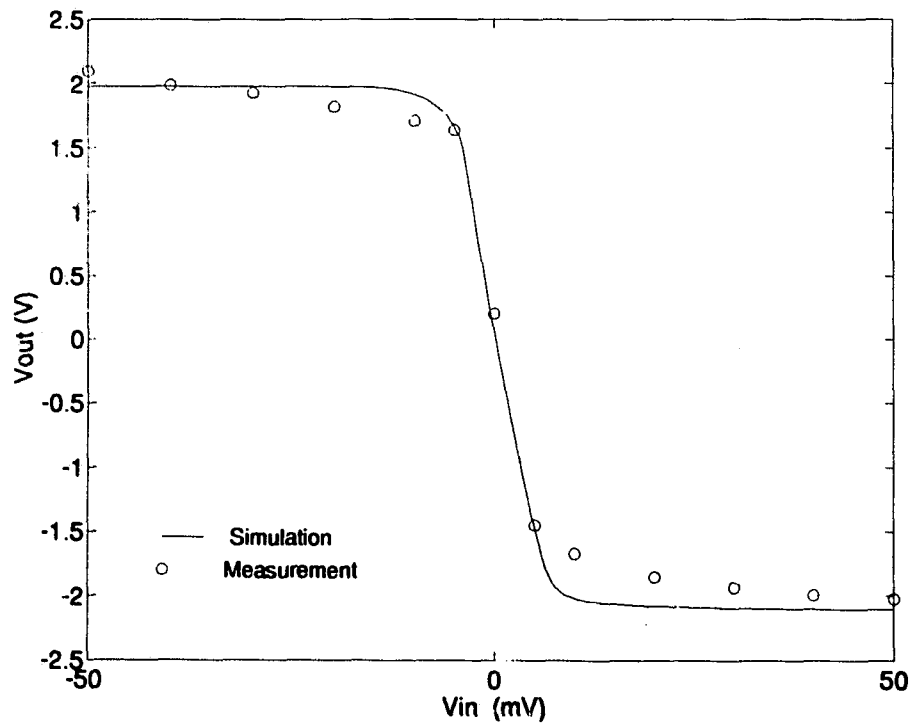


Figure 6.22: Transfer characteristics of the activation circuit obtained from the measurement and the simulation.

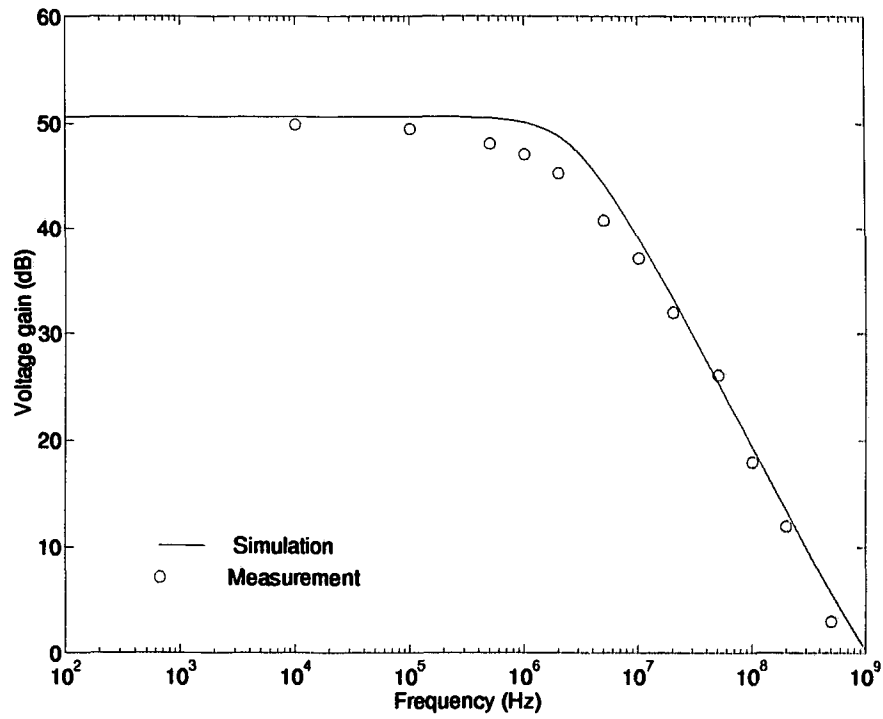


Figure 6.23: Amplitude response of the activation circuit obtained from the measurement and the simulation .

Parameter	Value
Voltage gain	50 dB
Unity-gain frequency	> 500 MHz ($C_L = 11$ pF)
V_{offset}	0.5 mV
Output swing	± 1.7 V
Settling time	50 ns ($C_L = 11$ pF)
Input resistance	120 K Ω (at 100 MHz)
Output resistance	60 Ω (at 100 MHz)
Phase margin	115 $^\circ$ (at 100 MHz)
CMRR	63.5 dB
Equivalent input noise*	5 nV/ \sqrt{Hz}
Power dissipation	150 mW

*: obtained from simulation.

Table 6.2: Performance of the activation circuit.

Chapter 7

GaAs Digital Circuit Design of Hybrid ANN Architecture

7.1 Introduction

The digital part in the hybrid ANN architecture includes a N -bit shift-register, a N -bit buffer-register, and other auxiliary circuits. This chapter describes the design of the shift-registers and the buffer-registers using the MESFET and the CMCCD technologies. Performance of the circuits are also discussed.

7.2 MESFET Shift- and Buffer-registers

7.2.1 Dual-inverter

It is required that both the shift-register and the buffer-register should provide the binary outputs as well as their complements since the inputs of the WSC require both the neuron state V_i and its complement \bar{V}_i . A dual-inverter based on BFL (Buffered FET Logic) [75] was designed as the core element in the both shift-register and the

buffer-register. Fig. 7.1 shows the schematic of the dual-inverter with the size of each MESFET as labelled. Fig. 7.2 shows the microphotograph of the dual-inverter.

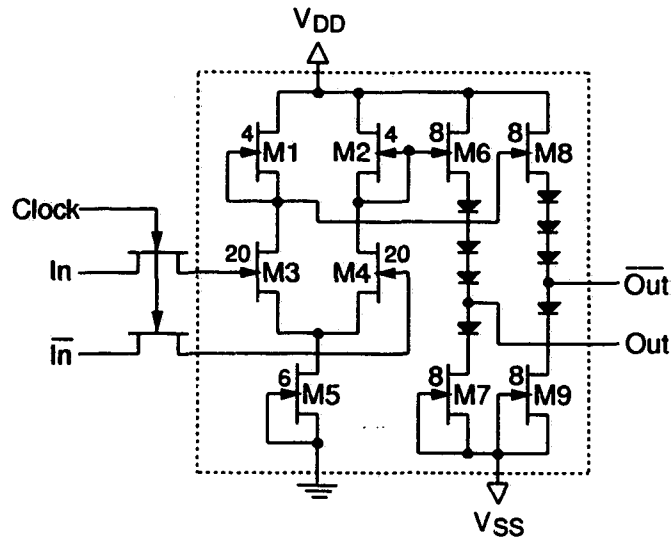


Figure 7.1: Schematic diagram of the dual-inverter.

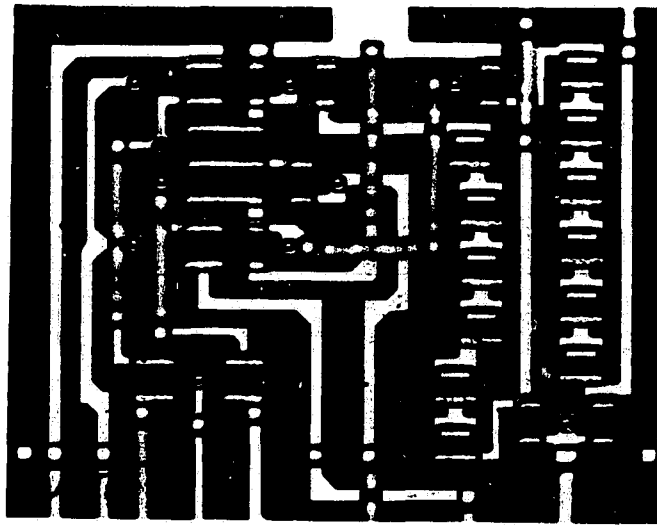


Figure 7.2: Microphotograph of the dual-inverter.

The dc performance of a logic gate is usually defined in terms of the dynamic

range of the output voltage and the noise margin. The dynamic range of the output voltage is given by:

$$\Delta V = V_{OH} - V_{OL} \quad (7.1)$$

where V_{OH} is the high output voltage and V_{OL} is the low output voltage. The noise margin is defined as:

$$V_{NH} = V_{OH} - V_{IH} \quad (7.2)$$

$$V_{NL} = V_{IL} - V_{OL} \quad (7.3)$$

where V_{IH} and V_{IL} are the minimum-high and maximum-low input voltages, respectively. The MESFETs were sized to obtain a better noise margin and dynamic range. The transfer characteristics and the dc parameters of the dual-inverter were obtained by HSPICE simulation for an inverter chain [76], as shown in Fig. 7.3. The device models of the NT/BNR GaAs process were used in the simulation. The simulated transfer characteristics of the dual-inverter are shown in Fig. 7.4. To calculate the noise margin, we need to obtain V_{IL} and V_{IH} . In this work V_{IL} and V_{IH} are defined as the input voltages in the transfer curve where the slopes are -1 [76]. Table 7.1 summarises the dc performance of the dual-inverter obtained using HSPICE simulation when V_{DD} and V_{SS} are 3.5 V and -2 V, respectively.

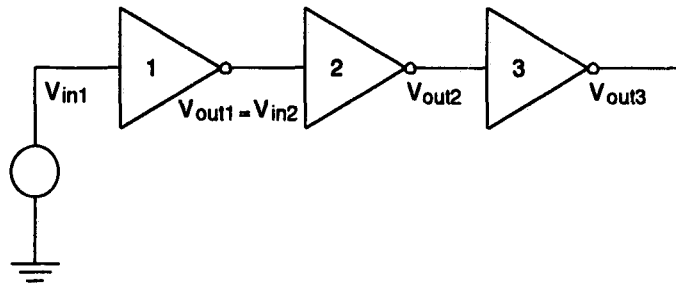


Figure 7.3: Inverter chain used for DC analysis of the inverter.

One of the transient properties of an inverter is the propagation delay. The propagation delay τ_{TP} is defined as the time interval between the 50% value of the input and

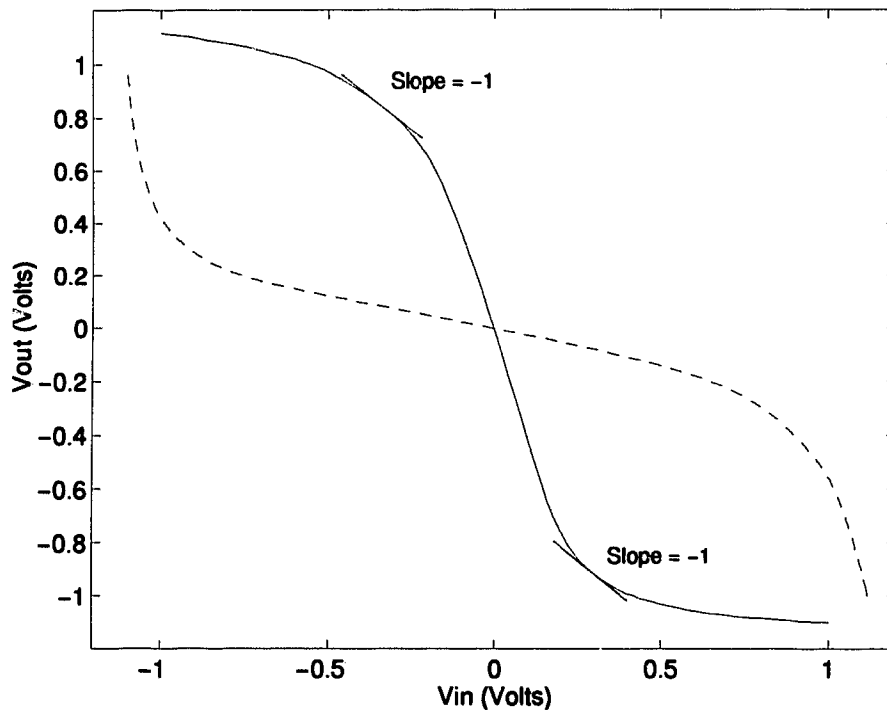


Figure 7.4: Transfer characteristic of the dual-inverter.

Parameter	Value
V_{OH}	1.12 V
V_{OL}	-1.08V
V_{IL}	-0.25 V
V_{IH}	0.30 V
V_{NL}	0.83 V
V_{NH}	0.82 V
R_{in}	60 K Ω
R_{out}	450 Ω
Power dissipation	12.05 mW

Table 7.1: DC performance of the dual-inverter.

the 50% value of the output. In general, the propagation delay for high-to-low transition, τ_{PHL} , is different from that for low-to-high transition, τ_{PLH} . The propagation delay is then determined by the average value of τ_{PHL} and τ_{PLH} [76], i.e.,

$$\tau_{PD} = \frac{1}{2}(\tau_{PHL} + \tau_{PLH}) \quad (7.4)$$

The propagation delay of the dual-inverter was calculated by a transient analysis of the dual-inverter chain. A square wave was used as the excitation source and the response of the first and second inverters were measured. Fig. 7.5 shows the waveforms of the input, the first and second outputs of the dual-inverter chain. It was measured (from Fig. 7.5) that $\tau_{PLH} = 85$ ps, $\tau_{PHL} = 27$ ps, and $\tau_{PD} = 56$ ps. The oscillograph of the input and the output waveforms of the dual-inverter are displayed in Fig. 7.6.

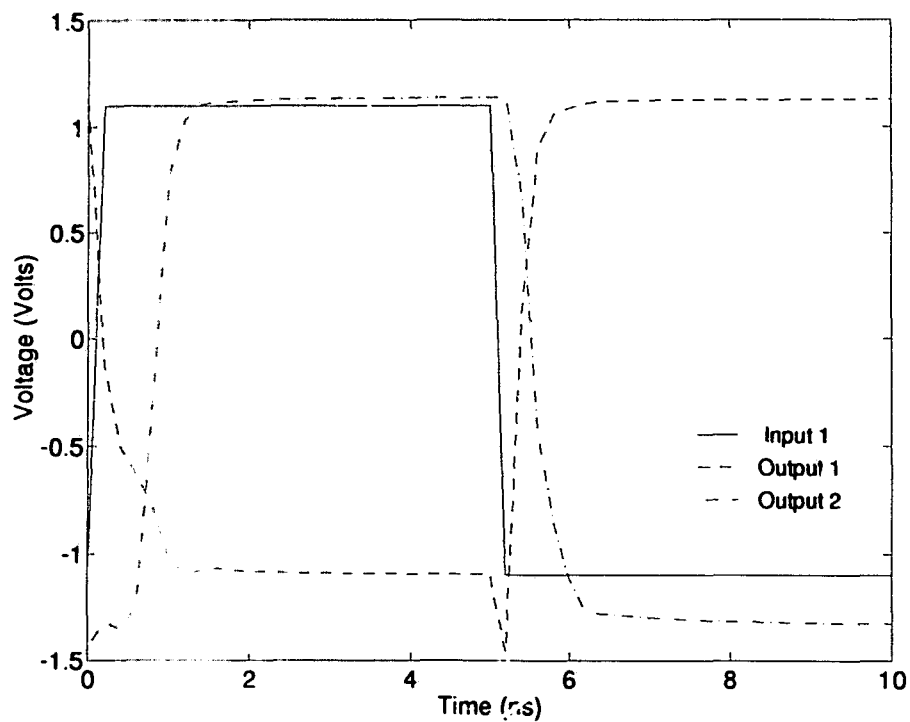


Figure 7.5: Transient response of the dual-inverter.

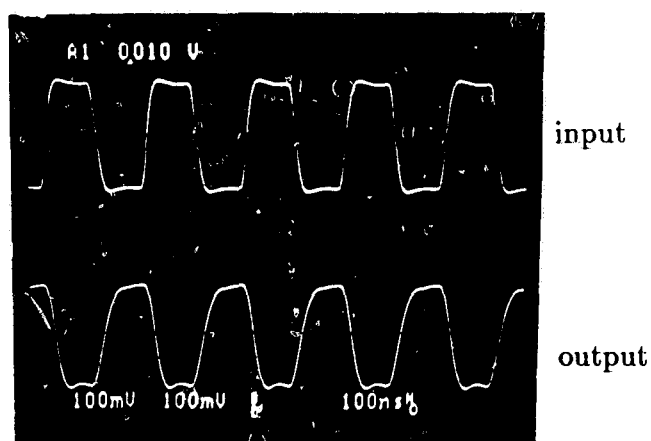


Figure 7.6: Oscilloscope of the input and the output waveforms of the dual-inverter.

7.2.2 MESFET Shift-register

The MESFET shift-register was designed using the dynamic flip-flops [76] configured with the pass-gates and the dual-inverters clocked by ϕ_1 and $\bar{\phi}_1$, as shown in Fig. 7.7. As can be seen from Fig. 7.7, each stage of the shift-register consists of two dual-inverters and four pass-gates. Fig. 7.8 shows the oscillograph of a 16-bit shift-register.

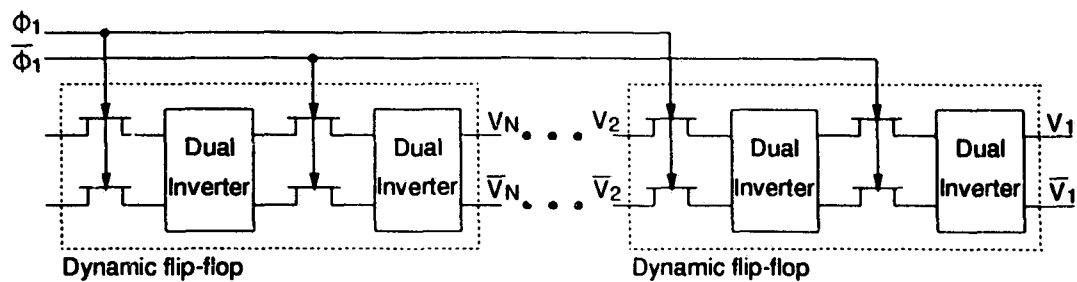


Figure 7.7: Diagram of the shift-register.

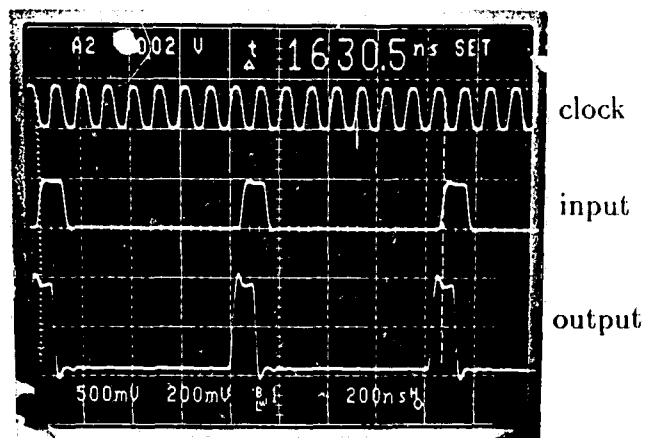


Figure 7.8: Oscillograph of the input and the output waveforms of the shift-register.

Clock ϕ_1 , the input pulse and the delayed pulse at the output of the MSB (bit 1) are displayed in the upper, middle and lower traces, respectively. Note that the outputs of the shift-registers are only valid during the negative cycle of ϕ_1 . It is observed that

the input signal is delayed by $1.6 \mu\text{s}$ after 16 clock pulses when the circuit operates at 10 MHz. It is also observed that there exist voltage spikes at both the positive and the negative transitions of the output. This is the result of clock feedthrough in the pass-gates of the dual-inverter.

7.2.3 MESFET Buffer-register

The buffer-register was designed using the dual-inverters and the pass gates, as shown in Fig. 7.9. There is one dual-inverter and four pass-gates at each stage of the buffer-register. When the initial neuron states are loaded into the buffer-register, all pass-gates connected to X_i and \bar{X}_i are on due to the high voltage of *Reset* and those connected to V_i and \bar{V}_i are off due to the low output voltage of the XOR gate. After the completion of one cycle of the computation, all pass-gates connected to V_i and \bar{V}_i are on due to the high voltage of ϕ_2 and those connected to X_i and \bar{X}_i are off due to the low voltage of *Reset*. During the clock pulse of ϕ_2 , V_i and \bar{V}_i are transferred to the outputs of the buffer-register and ready for the next cycle of computation.

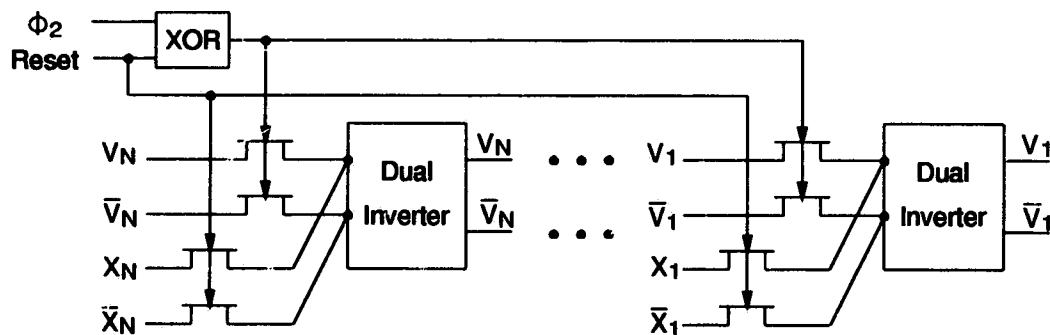


Figure 7.9: Diagram of the buffer-register.

Fig. 7.10 shows the schematic diagram of the XOR gate. Table 7.2 lists its logic

function and Fig. 7.11 shows the input and the output waveforms of the XOR gate obtained from HSPICE simulation.

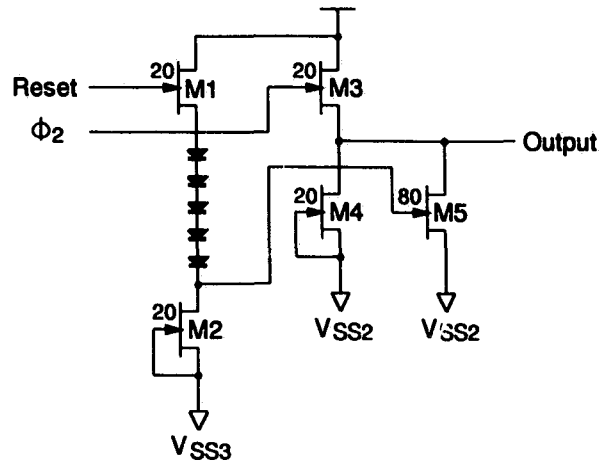


Figure 7.10: Schematic diagram of the XOR gate.

ϕ_2	Reset	Output of XOR
Low	Low	Low
Low	High	X
High	Low	High
High	High	Low

Table 7.2: Logic function of the XOR gate.

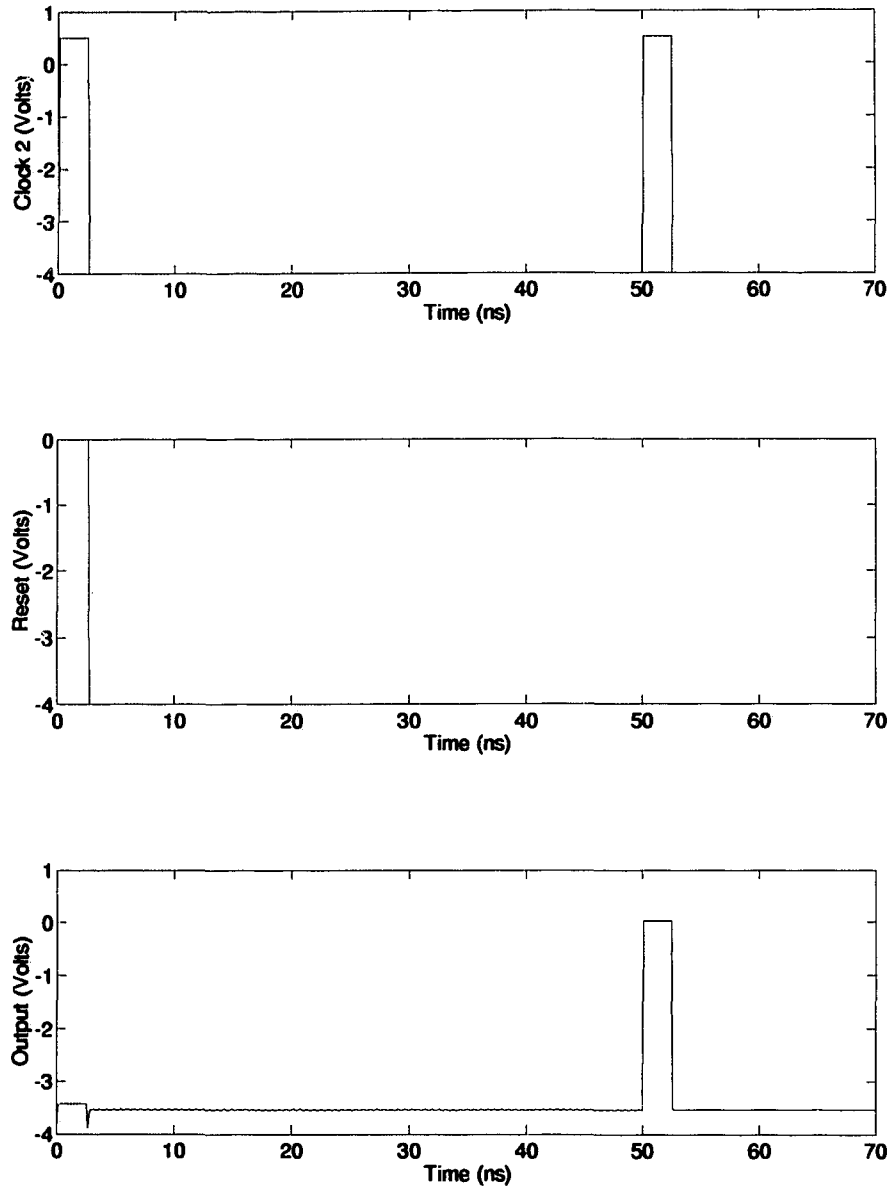


Figure 7.11: Simulation waveforms of the input and the output of the XOR gate.

7.3 CMCCD Shift- and Buffer-registers

Alternatively, the shift- and buffer-registers can be implemented using the CMCCD technology. CMCCD, by its construction, is a shift-register. As we discussed in the preceding chapter, the CMCCD has low noise and a large dynamic range, and it is therefore quite suitable for use in our design of the shift- and buffer-registers. Moreover, compared with the MESFET counterpart, a CMCCD shift-register should also be superior in bit density and power consumption, as we shall demonstrate in the end of this section.

7.3.1 Circuit Structure

The diagram of the N -bit CMCCD shift- and buffer-registers using the uni-phase clocking scheme is shown in Fig. 7.12. The main part of the circuit is a N -pixel CMCCD with four electrodes per pixel. The input structure of the shift-register is the same as that described in Chapter 2. The outputs of the shift-register appear vertically at the upper side of the $Bias2$ electrode in each pixel. Similar to the output structure described in Chapter 2, each output of the shift-register consists of an output gate controlled by ϕ_2 and an ohmic contact connected to a reset MESFET. As can be seen, this structure allows the shift-register to function as a series-in parallel-out register.

The buffer-register is composed of N output buffers. Fig. 7.13 shows the schematic of the output buffer which uses a flip-flop configuration. When the charge packet arrives at the output ohmic contact, the precharged voltage at the source of the reset MESFET changes, and the output buffer switches into one of the states due to regenerative action, depending on the relative value of the source voltage of the reset MESFET and V_{ref2} . Usually, V_{ref2} is chosen to lie midway between the source voltages of the reset MESFET representing logic 1 and 0. The output buffer generates

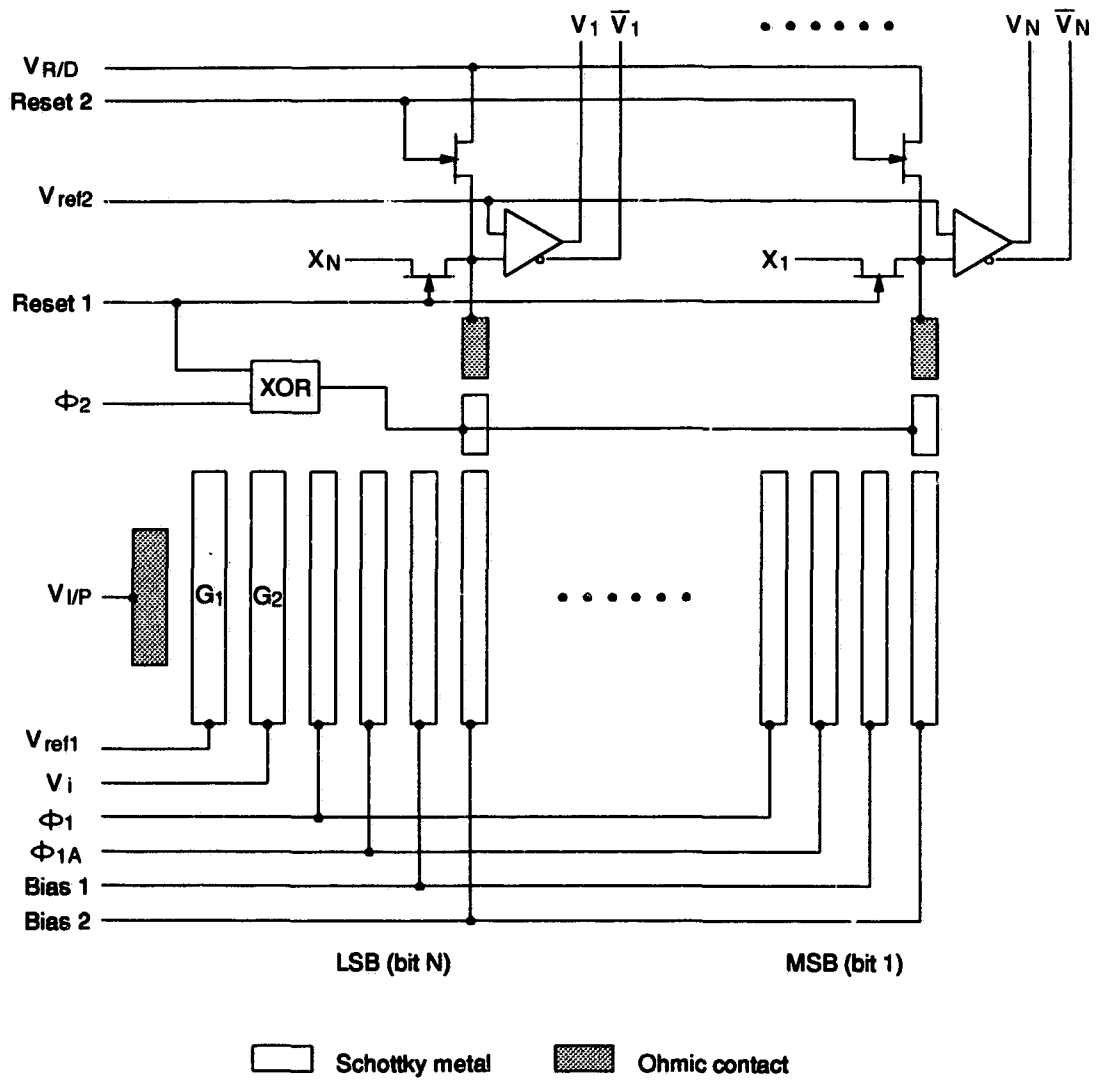


Figure 7.12: Diagram of the CMCCD shift- and buffer-registers.

an output signal and its complement, which are required in the weight-summation operation.

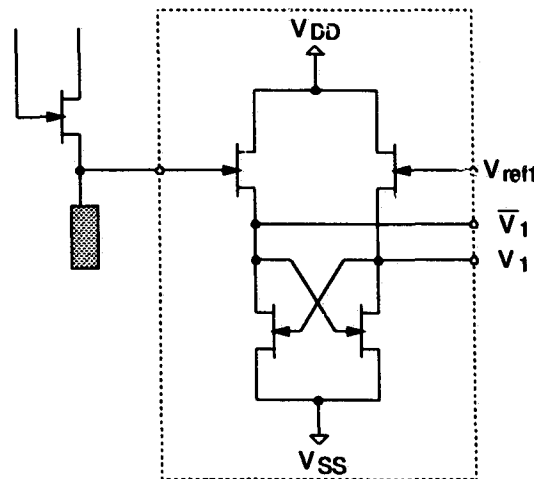


Figure 7.13: Schematic of the output buffer.

7.3.2 Operation of the CMCCD Shift- and Buffer-registers

The timing diagram of the CMCCD shift- and buffer-registers is given in Fig. 7.14. In the beginning of the computation, the initial neuron states X_1, X_2, \dots, X_N are loaded into the buffer-register through the pass-gates controlled by *Reset1*. Synchronized by ϕ_1 (ϕ_{1A}), the neuron states are updated and transferred serially into the CMCCD shift-register through the input gate G_2 . When the N th pulse of ϕ_1 (ϕ_{1A}) appears, the updated state of the N th neuron is loaded in the LSB (bit N) of the shift-register. Before the charge packets are detected, all the output ohmic contacts are precharged to a more positive voltage level than ϕ_2 . During the negative half cycle of ϕ_1 (ϕ_{1A}), all the charge packets representing the binary neuron states are under the electrodes connected to *Bias2* (-1 V), and no charge packets exist under the other electrodes. The charge packets move to the buffer register in parallel when ϕ_2 is high,

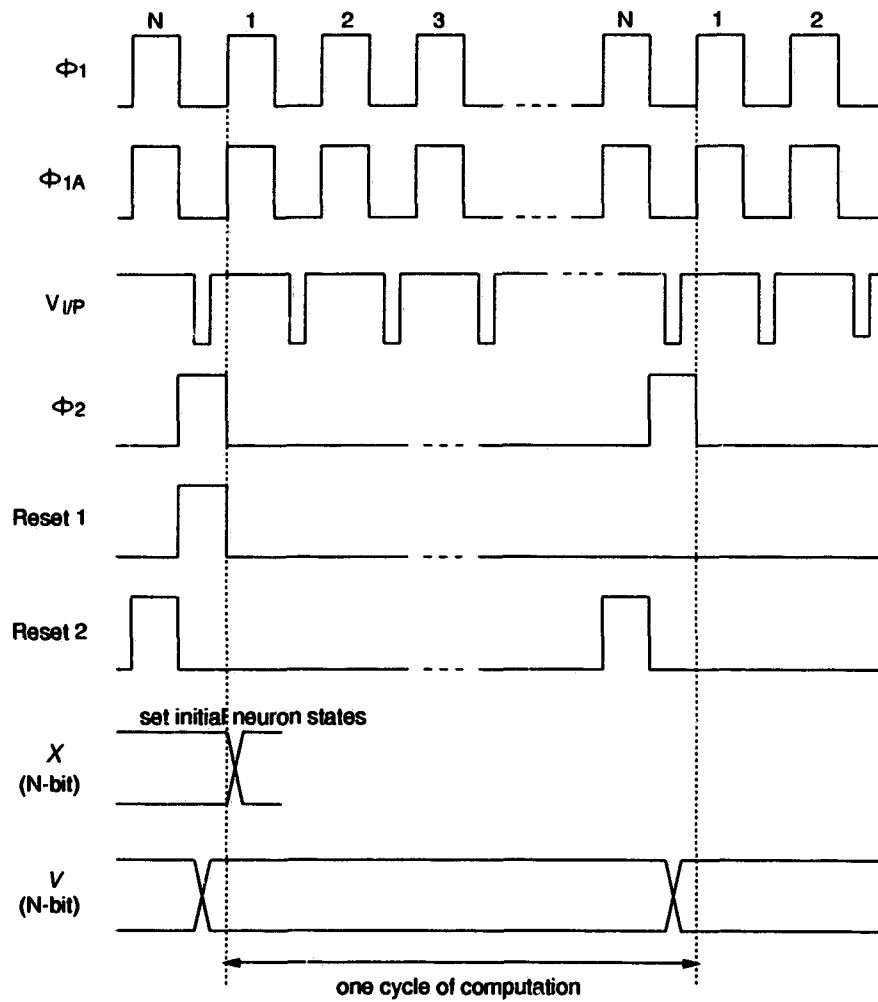


Figure 7.14: Timing diagram of the CMCCD shift- and buffer-registers.

and therefore, there is one charge packet in the channel of the CMCCD shift-register. The “reset” shift-register is then ready for the next cycle of computation. Note that when the shift-register is operating in series, no charge will be detected as ϕ_2 is low until the negative half cycle of the N th pulse of ϕ_1 (ϕ_{1A}). This means all the neuron states V_1, V_2, \dots, V_N are kept in the stable states in the buffer-register for N serial weight-summation operations.

7.4 Discussions on Power and Area Consumptions of the Shift- and Buffer-registers

The power consumed in a CMCCD shift-register is essentially that required to charge the gate capacitances. This power is dissipated mostly in the clock driver, not on the CCD chip itself. The power dissipated per pixel in the clock driver is given by:

$$P_{clock} = C_c V^2 f_c \quad (7.5)$$

where C_c is the total clock line capacitance associated with the electrodes connected to the clock within the pixel, V is the clock voltage, and f_c is the clock frequency. Using the parameters of the two-phase and the uni-phase CMCCDs, and assuming that clock frequency is 100 MHz and the amplitude of clock voltage is 5 V for the two-phase CMCCD and 4 V for the uni-phase CMCCD, we calculated the power dissipations of the clock drivers to be:

$$P_{clock} = 0.050 \text{ mW/pixel} \quad (7.6)$$

for the two-phase CMCCD, and

$$P_{clock} = 0.032 \text{ mW/pixel} \quad (7.7)$$

for the uni-phase CMCCD.

The power dissipation per pixel due to the current flowing and the charge packet lifting over a potential barrier is approximately given by [66]:

$$P_{CCD} \approx \frac{nq(f_c L)^2}{\mu} + 2nqf_c(\psi_{BH} - \frac{1}{2} \frac{nq}{C_e}) \quad (7.8)$$

where n is the number of electrons per pixel, L is the length of the pixel, μ is the electron mobility in GaAs. Fig. 7.15 shows the curves of the on-chip power dissipation with respect to the clock frequency for the two-phase and the uni-phase CMCCDs. From Eqs. (7.6) and (7.7) and Fig. 7.15, it can be shown that for both the two-phase and the uni-phase CMCCDs driven by a 100 MHz clock driver, the driver power dissipation per pixel is about 100 times greater than the on-chip power dissipation per pixel.

Since one stage of the MESFET shift-register consists of two dual-inverters, the power dissipation of the MESFET shift-register is, therefore, $2 \times 12.05 \text{ mW} = 24.1 \text{ mW}$. If we assume the CMCCD shift-register uses the uni-phase clocking scheme, the power consumed by one stage of the CMCCD shift-register is approximately 0.032 mW . This suggests that the CMCCD shift-register consumes only $1/80$ of the power consumed by the MESFET shift-register.

The active area of the dual-inverter layout shown in Fig. 7.2 is $135 \mu\text{m} \times 90 \mu\text{m}$, and therefore, the total active area of one stage of the MESFET shift-register is $270 \mu\text{m} \times 90 \mu\text{m} = 0.0243 \text{ mm}^2$. It should be mentioned here that the layout of the dual-inverter has not been compacted. If we use the geometric parameters obtained in Chapter 3 for the uni-phase CMCCD, the area occupied by one stage of the CMCCD shift-register is $24 \mu\text{m} \times 100 \mu\text{m} = 0.0024 \text{ mm}^2$. This means that CMCCD shift-register consumes only about $1/10$ of the area consumed by the MESFET shift-register.

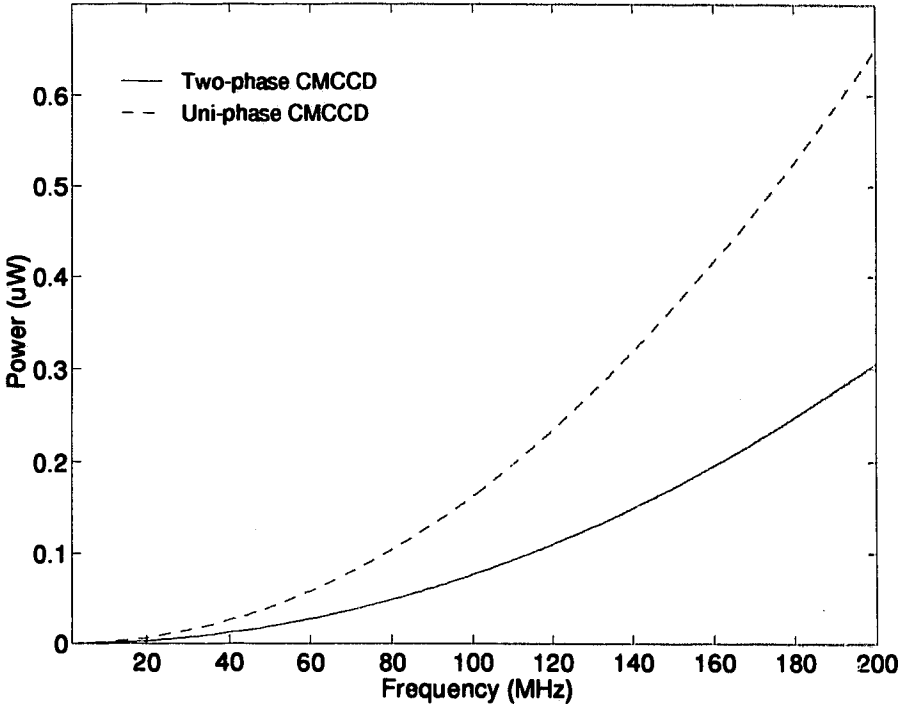


Figure 7.15: On-chip power dissipation versus clock frequency for the two-phase and the uni-phase CMCCDs.

Chapter 8

Fabrication and Measurement of ANN IC

8.1 Introduction

In order to verify the design of the hybrid ANN architecture described in Chapters 6 and 7, we constructed a Hopfield ANN circuit with 16 neurons and had it fabricated using the NT/BNR (Northern Telecom/Bell Northern Research) 0.8 μm GaAs MESFET process. Fig. 8.1 shows the schematic of the circuit which consists of 16 transconductance amplifiers used for the WSC, an activation circuit, a 16-bit shift-register, a 16-bit buffer-register, and other auxiliary circuits. Since the NT/BNR process is not compatible with the fabrication of GaAs CCD (we shall discuss this later), the circuit doesn't include the CMCCD weight storage array, and uses the MESFET shift- and buffer-registers. A multi-chip approach was taken for testing. The synaptic weights were generated externally using EPROMs. In Fig. 8.1, the symbols labelled by "T;" stand for the transconductance amplifiers, the symbol labelled by "AC" stands for the activation circuit, and the symbols labelled by "D-I"

stands for the dual-inverters. In this chapter, we describe briefly the simulation, the physical layout and the GaAs fabrication of the ANN chip. We then present the experimental results of the ANN IC and demonstrate its functionality as an associative memory. The problems related to the chip fabrication and the measurement are also discussed.

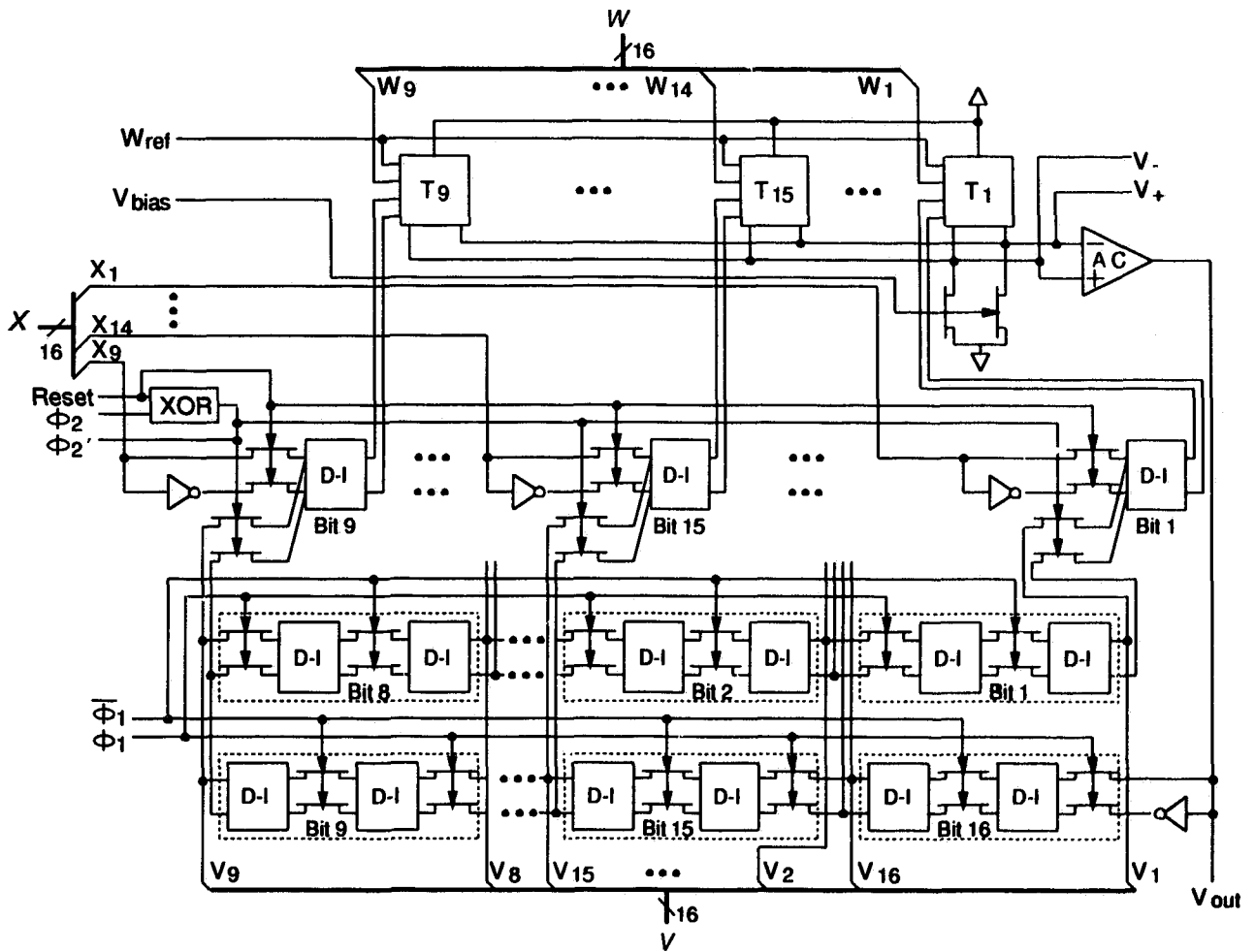


Figure 8.1: Schematic of the ANN chip.

8.2 Simulation, Physical Layout and Fabrication

8.2.1 Simulation

The circuit shown in Fig. 8.1 was simulated using HSPICE incorporated with the device models of the NT/BNR process prior to the physical layout and the fabrication. As can be seen from Fig. 8.1, the circuit has a set of inputs for the clocks and the reset signal, 16 inputs for the synaptic weights, 16 inputs for the initial neuron states and a group of dc power supplies. It has one output from the activation circuit, 16 outputs from the shift-register to give the updated neuron states, and two outputs, V_+ and V_- , of the WSC. All the inputs and the outputs of the activation circuit, the WSC, and each stage of the shift-register are accessible, and hence these subcircuits are testable. The frequency of ϕ_1 used in the simulation is 200 MHz and the frequency of ϕ_2 is 12.50 MHz. This means that each neuron will update its state every 2.5 ns and it takes 80 ns for the circuit to complete one full cycle of the computation.

Fig. 8.2 shows two 16-bit test patterns to be stored in the synaptic weight array. The number of patterns M is limited by the storage capacity of the Hopfield neural

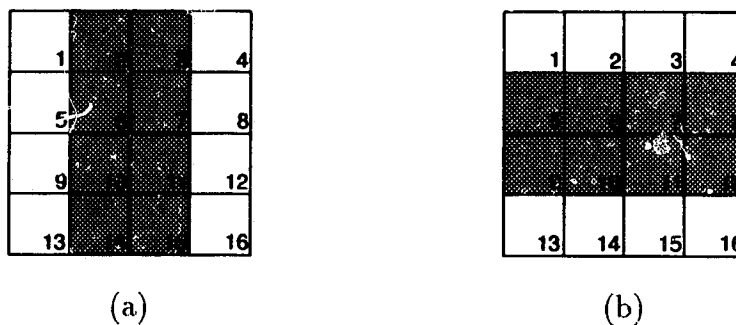


Figure 8.2: Two 16-bit test patterns stored in the synapses.

network and it is found to be less than $N/(2 \log N)$ [77]. For 16 neurons, $M = 2$. These two patterns are represented by the following vectors:

-1 +1 +1 -1 -1 +1 +1 -1 -1 +1 +1 -1 -1 +1 +1 -1
 -1 -1 -1 -1 +1 +1 +1 +1 +1 +1 +1 +1 -1 -1 -1 -1

where +1 stands for the black square and -1 stands for the white square. For these two patterns, the weight array was obtained from Eq. (5.10) and is given by:

16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	
0	0	0	2	0	-2	-2	0	0	-2	-2	0	2	0	0	2	W_1
0	0	2	0	-2	0	0	-2	-2	0	0	-2	0	2	2	0	W_2
0	2	0	0	-2	0	0	-2	-2	0	0	-2	0	2	2	0	W_3
2	0	0	0	0	-2	-2	0	0	-2	-2	0	2	0	0	2	W_4
0	-2	-2	0	0	0	0	2	2	0	0	2	0	-2	-2	0	W_5
-2	0	0	-2	0	0	2	0	0	2	2	0	-2	0	0	-2	W_6
-2	0	0	-2	0	2	0	0	0	2	2	0	-2	0	0	-2	W_7
0	-2	-2	0	2	0	0	0	2	0	0	2	0	-2	-2	0	W_8
0	-2	-2	0	2	0	0	2	0	0	0	2	0	-2	-2	0	W_9
-2	0	0	-2	0	2	2	0	0	0	2	0	-2	0	0	-2	W_{10}
-2	0	0	-2	0	2	2	0	0	2	0	0	-2	0	0	-2	W_{11}
0	-2	-2	0	2	0	0	2	2	0	0	0	0	-2	-2	0	W_{12}
2	0	0	2	0	-2	-2	0	0	-2	-2	0	0	0	0	2	W_{13}
0	2	2	0	-2	0	0	-2	-2	0	0	-2	0	0	2	0	W_{14}
0	2	2	0	-2	0	0	-2	-2	0	0	-2	0	2	0	0	W_{15}
2	0	0	2	0	-2	-2	0	0	-2	-2	0	2	0	0	0	W_{16}

Note that the weights have three levels and each row of the array represents an input signal to the W_i pin of the chip.

The functionality of the ANN IC as an associative memory was simulated by introducing an input vector to the circuit during the first pulse of ϕ_2 . Fig. 8.3 shows one of the patterns of the input vector which is represented by:

$$\begin{array}{cccccccccccccccc} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 & X_8 & X_9 & X_{10} & X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{16} \\ -1 & +1 & +1 & -1 & +1 & +1 & +1 & -1 & -1 & +1 & +1 & +1 & -1 & +1 & +1 & -1 \end{array}$$

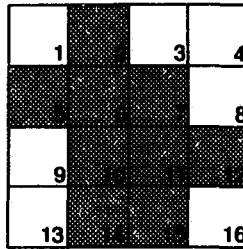


Figure 8.3: One of the input patterns applied to the circuit.

Fig. 8.4 shows the simulation waveforms of the 16-bit outputs of the ANN IC. It can be seen that after one cycle of the computation, i.e., when the second pulse of ϕ_2 comes at $t = 80$ ns, the outputs of the network reaches its stable states:

$$\begin{array}{cccccccccccccccc} V_1 & V_2 & V_3 & V_4 & V_5 & V_6 & V_7 & V_8 & V_9 & V_{10} & V_{11} & V_{12} & V_{13} & V_{14} & V_{15} & V_{16} \\ -1 & +1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 & -1 \end{array}$$

which is the first stored pattern and is the nearest match to the input pattern. Fig. 8.5 shows the simulation waveforms of the clocks ϕ_1 and ϕ_2 and the output of the activation circuit corresponding to the first stored pattern. The simulation waveforms corresponding to the second stored pattern are shown in Fig. 8.6.

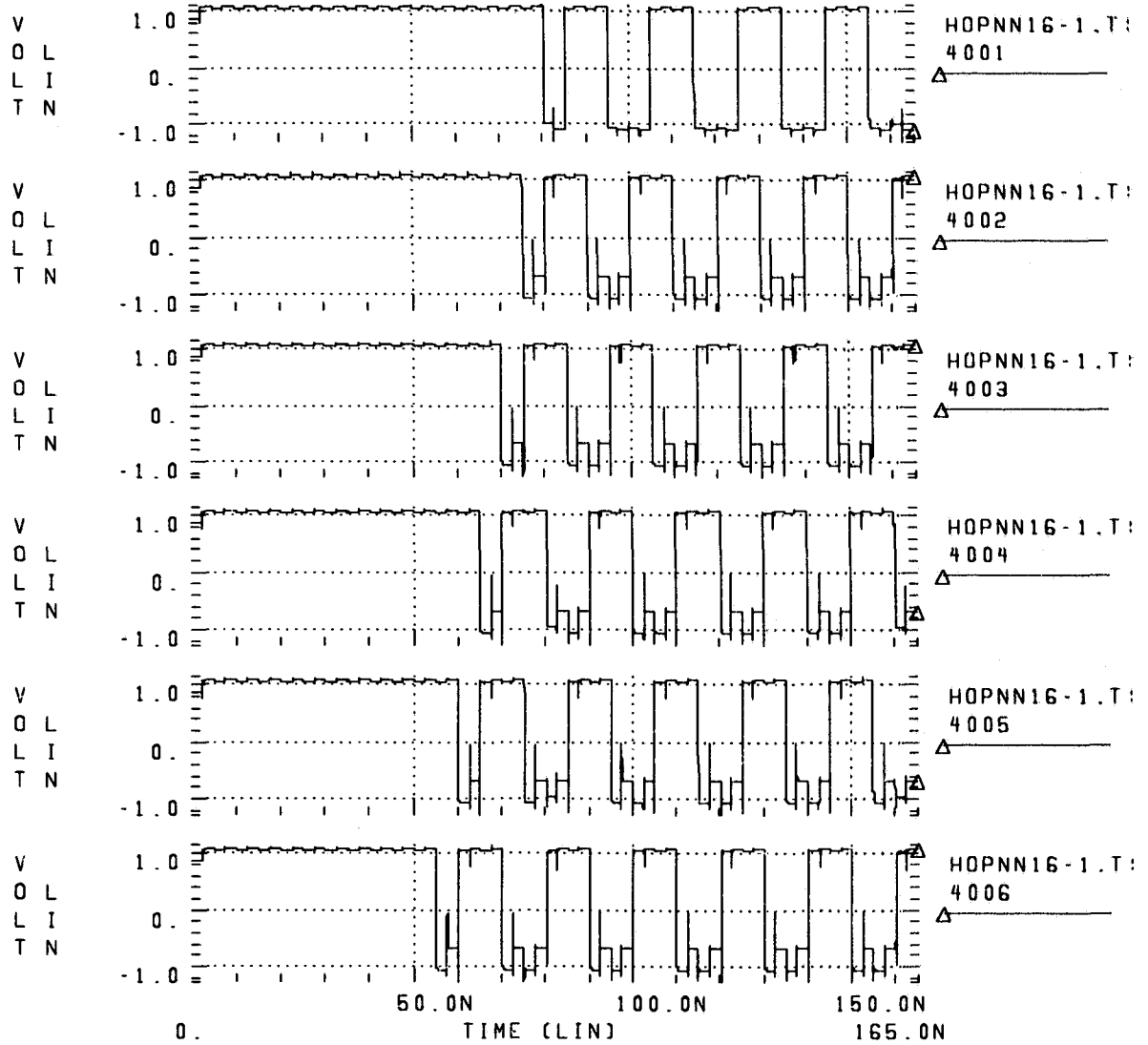


Figure 8.4: Simulation waveforms of the 16-bit outputs of the ANN IC.

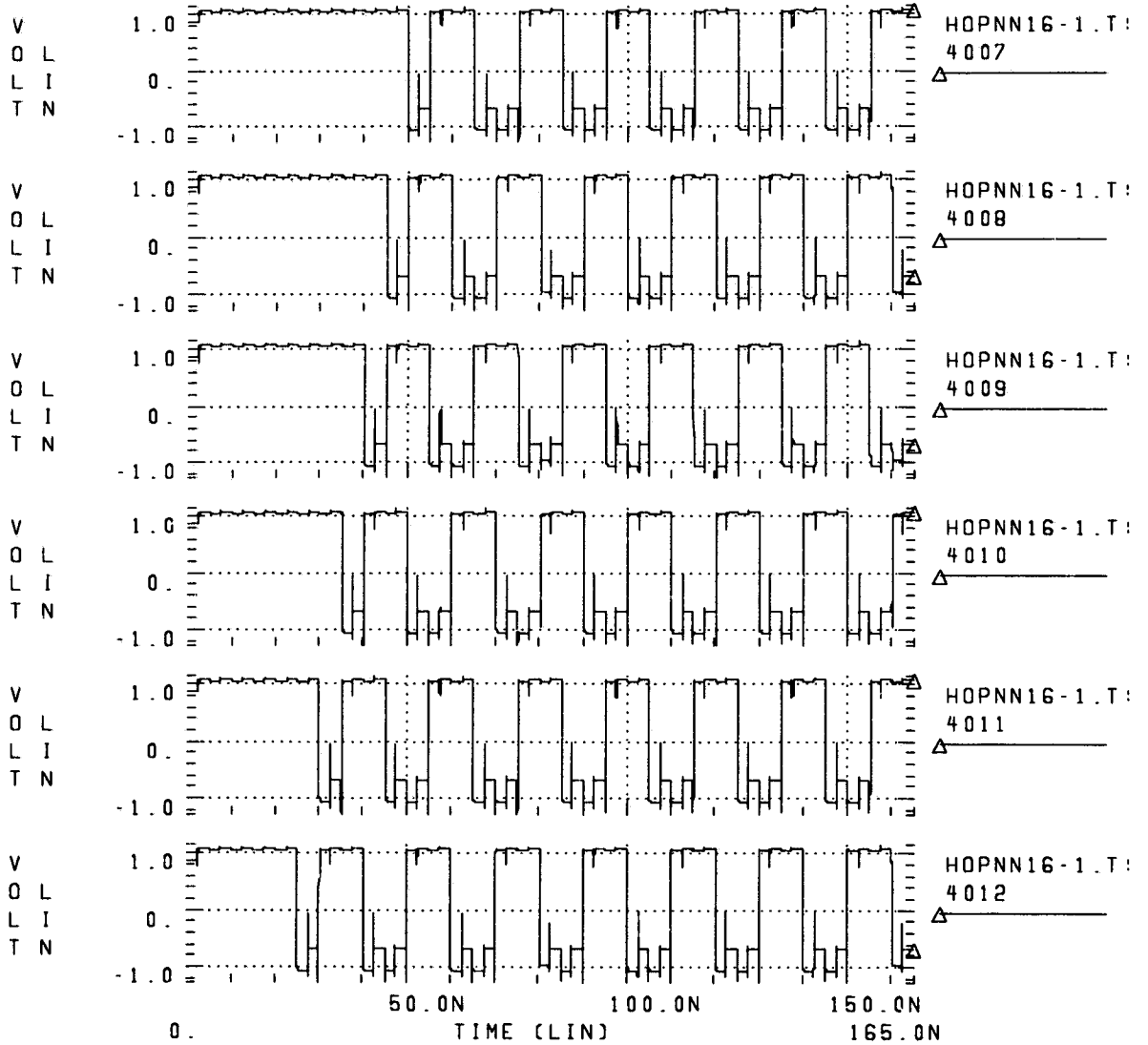
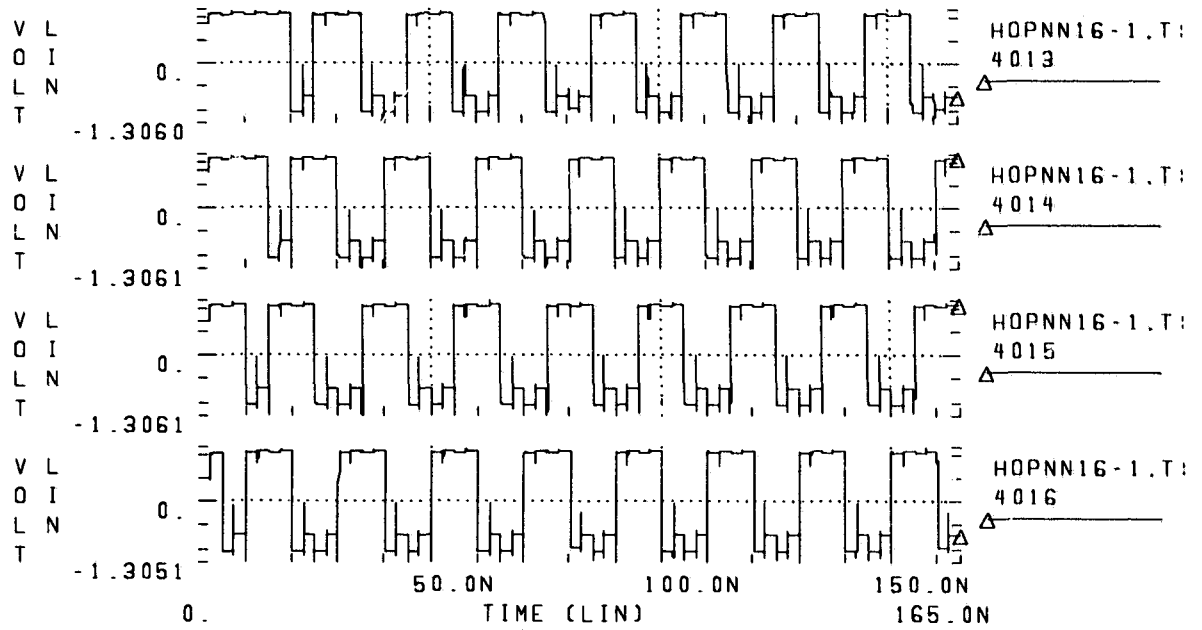


Figure 8.4: Simulation waveforms of the 16-bit outputs of the ANN IC (continued).



EMPTY PANEL 5

EMPTY PANEL 6

Figure 8.4: Simulation waveforms of the 16-bit outputs of the ANN IC (continued).

EMPTY PANEL 1

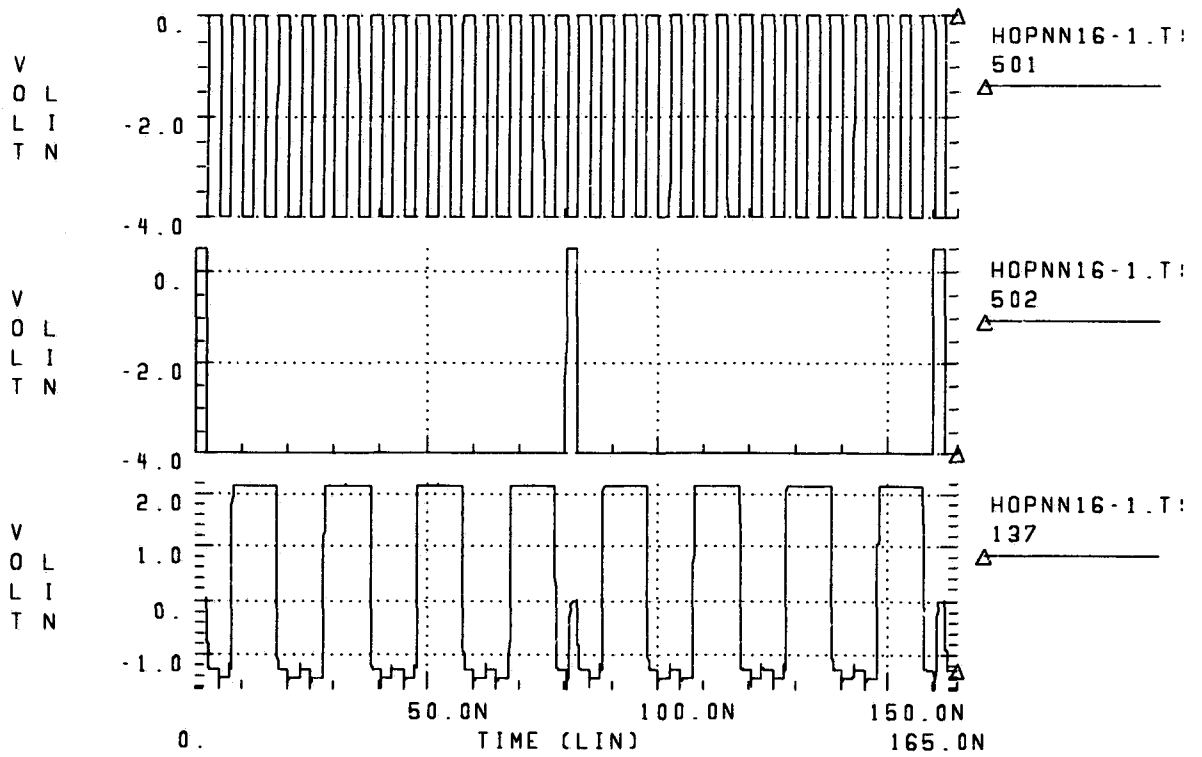


Figure 8.5: Simulation waveforms of the clocks ϕ_1 and ϕ_2 and the output of the activation circuit corresponding to the first stored pattern.

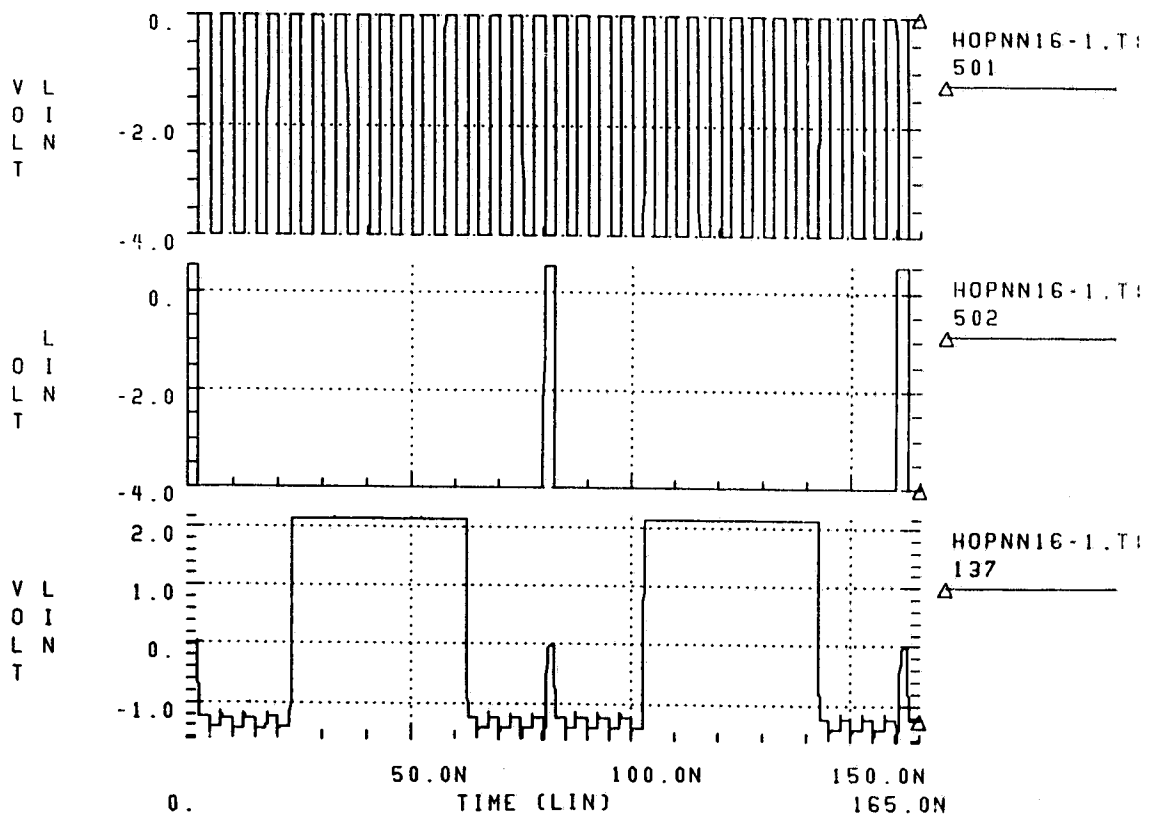


Figure 8.6: Simulation waveforms of the clocks ϕ_1 and ϕ_2 and the output of the activation circuit corresponding to the second stored pattern.

8.2.2 Physical Layout

Fig. 8.7 shows the microphotograph of the ANN IC. The standard-cell approach [78] was used in the layout. The physical layout of the transconductance amplifier, the activation circuit, and the dual-inverter were accomplished at the transistor level using Cadence Edge that is supported by the NT/BNR MESFET device library. The layout of the chip was then accomplished hierarchically and checked using a design-rule checker in Cadence incorporated with the NT/BNR design rules. The ANN IC contains totally 1303 MESFETs and Schottky diodes and occupies an active area of $3.1 \times 1.5 \text{ mm}^2$. The dies were packaged in the 84-pin PGAs.

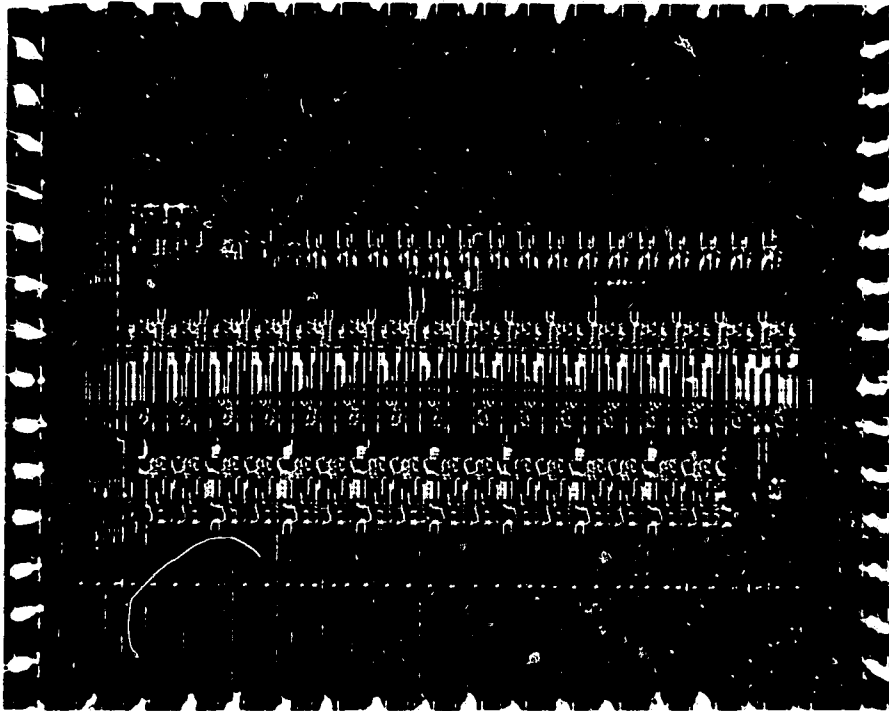


Figure 8.7: Microphotograph of the ANN chip.

8.2.3 Fabrication

The ANN IC was fabricated using the NT/BNR 0.8 μm GaAs depletion-mode MESFET process. The MESFET fabricated by the process exhibits a pinch-off voltage (V_P) of -1.2 V, a saturation current (I_{DSS}) of 120 A/m, a transconductance (g_m) of 140 S/m, and a cut-off frequency (f_T) of 20 GHz. The process provides two layers of metallization with the sheet resistances of $30 \text{ m}\Omega/\text{cm}^2$ for metal 1 and $20 \text{ m}\Omega/\text{cm}^2$ for metal 2. The capacitor used in the activation circuit was formed in between metal 1 and metal 2. The device models and the fabrication steps of the process are given in Appendixes F and G.

8.3 Measurement of the ANN IC for Open-loop Operation

In order to test the functionality of the neuron, i.e., the weight-summation operation and the activation function, of the chip, we first made a measurement on the open-loop operation of the chip, i.e., when the output of the activation circuit was disconnected from the shift-register and the inputs of the initial neuron states were connected externally to the WSC.

The neuron activation function was tested by measuring the output characteristic of the activation circuit which is connected with the WSC. The 16 inputs of the WSC were assumed to be the binary states of the other neurons (\mathbf{V}). Fig. 8.8 shows the output characteristic of the neuron. V_{out} , the output voltage of the neuron, is a function of the number of “on”-neurons in \mathbf{V} when the weights are constant. In this measurement, all of the weights were 0.7 V while w_{ref} was 0.5 V. It can be seen that the output of the neuron is always saturated, either in the “on” state or in the “off” state, when the difference between the numbers of “on”-neurons and “off”-neurons in

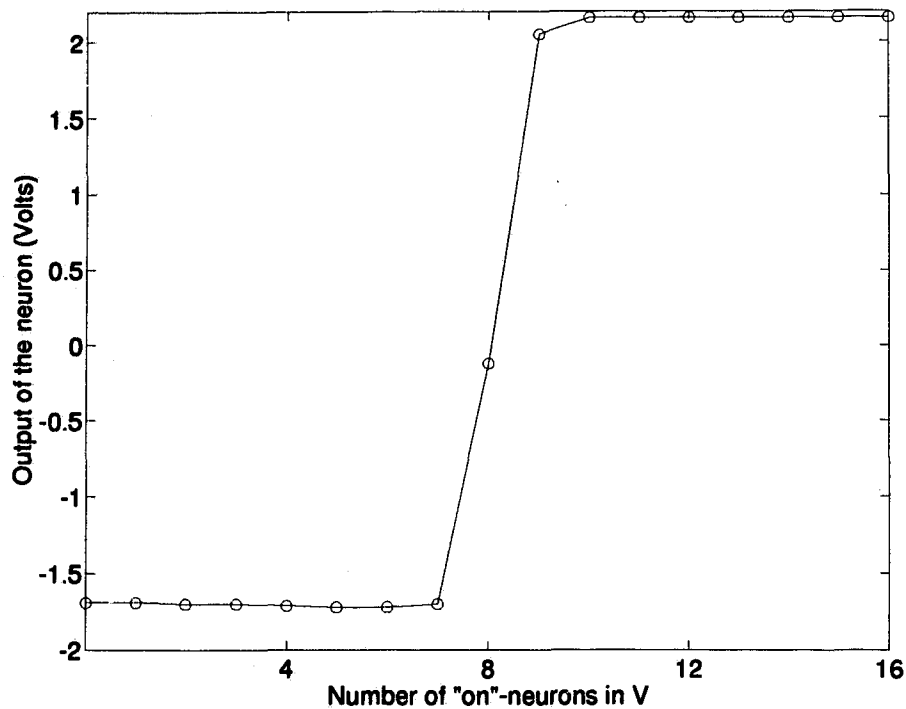


Figure 8.8: Output characteristic of one neuron.

V is greater than two. The fact that the neuron behaves as a hard-limiting function is due to the high gain of the activation circuit.

The linearity of the weight-summation operation for the analog weight was measured using the square-waveform signals in V and W . For simplicity, the 16 transconductance amplifiers were divided into two groups, i.e., eight of the transconductance amplifiers were connected together to the same signal source. We kept the binary inputs in V fixed and only changed the analog voltage of one of the weights in W . The differential output ($V_+ - V_-$) of the WSC with respect to the change in the analog weight is plotted in Fig. 8.9. It is observed that the results are quite linear as long as the voltage range of $(W_i - w_{ref})$ is within ± 0.5 V.

We also ran a series of tests on the chip to evaluate the dynamic function of the Hopfield neural network in the open-loop operation. The test was performed at

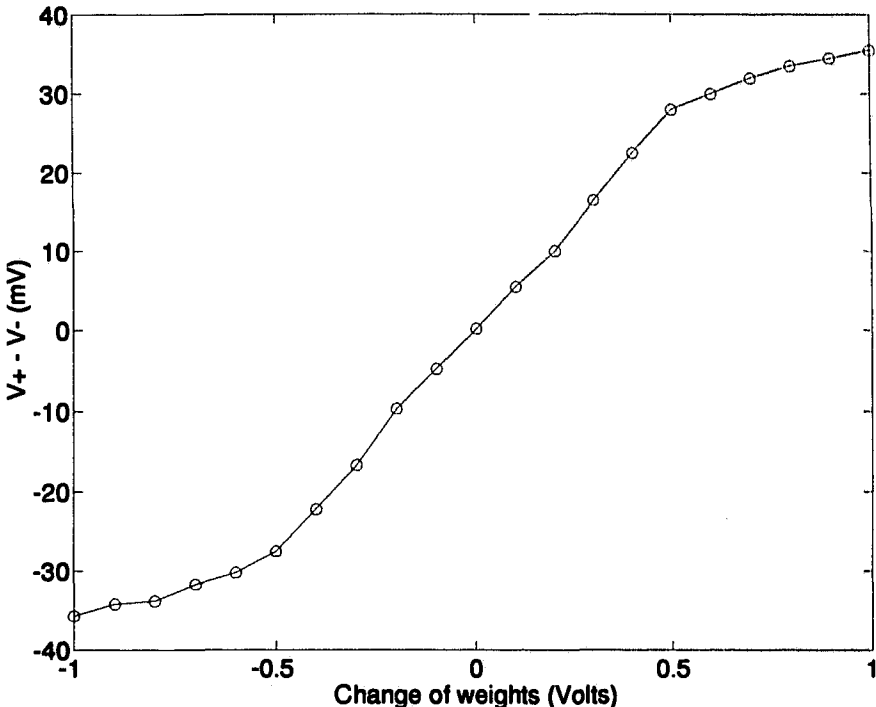


Figure 8.9: Measured linearity of the weight-summation computation for analog weights.

the frequency range of 1 MHz to 100 MHz. Again, for simplicity, some of the pins associated with the weights and the inputs of the neuron state were fed by the same signal sources. The differential output of the WSC and the output of the activation circuit were measured.

Fig. 8.10 shows a group of the input waveforms used in the test and the theoretical output waveforms of the ANN IC for the open-loop operation. The weights $W_1 \sim W_4$ were fed by the same signal source, and so did weights $W_5 \sim W_8$, $W_9 \sim W_{12}$, and $W_{13} \sim W_{16}$. The inputs of the neuron state $V_1, V_2, V_3, V_4, V_{13}, V_{14}, V_{15}$, and V_{16} were fed by the same signal source, and so did $V_5, V_6, V_7, V_8, V_9, V_{10}, V_{11}$, and V_{12} . The differential output of the WSC ($\Delta V = V_+ - V_-$) which is also the input of the activation circuit, is determined by the Hopfield ANN model and is given by:

$$\Delta V = V_+ - V_- = \sum_{i=1}^{16} W_i V_i \quad (8.1)$$

The output of the activation circuit is given by :

$$V_{out} = f\left(\sum_{i=1}^{16} W_i V_i\right) \quad (8.2)$$

where f is the hard-limiting function.

Fig. 8.11 shows the oscillographs of the open-loop measurement. The waveforms within the time slot marked by the vertical cursors correspond to those shown in Fig. 8.10. The four traces in Fig. 8.11(a) are the waveforms of the weights. The first two upper traces in Fig. 8.11(b) are the neuron state inputs. The differential output of the WSC and the output of the activation circuit are displayed in the lower traces of Fig. 8.11(b) and (c).

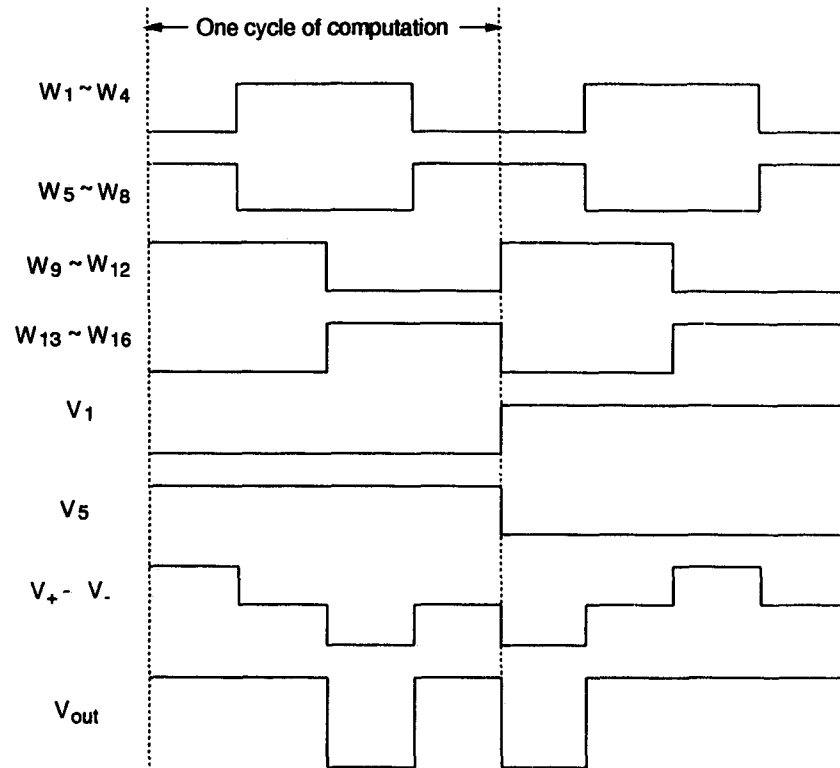
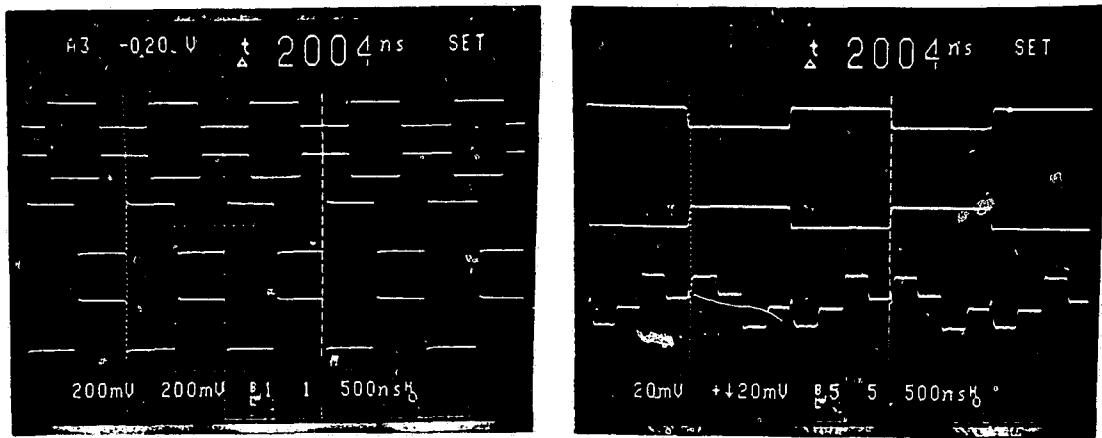
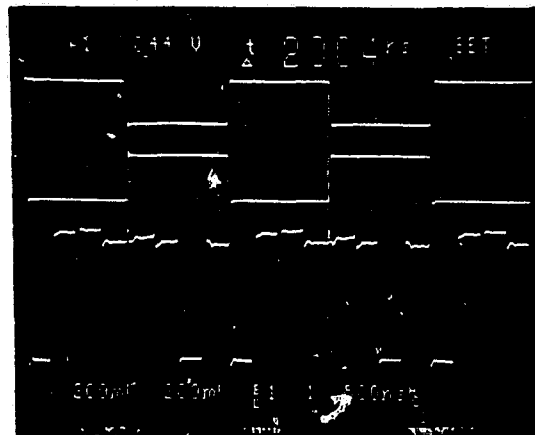


Figure 8.10: Input waveforms used in the test and the theoretical output waveforms of the ANN IC for the open-loop operation.



(a)

(b)



(c)

Figure 8.11: Oscillographs of the open-loop measurement.

8.4 Measurement of the ANN IC as An Associative Memory

The functionality of the ANN IC as an associative memory was tested under the close-loop operation where the output of the activation circuit was fed into the shift-register and the network was controlled by two clocks, ϕ_1 and ϕ_2 . Several groups of the test patterns were used in the test. In the following we assume that the two stored patterns used in the test are the same as those used in the simulation (see Section 8.2).

Fig. 8.12 shows the circuit used to generate the three-level synaptic weights. The three-level weight signal is coded as shown in Table 8.1. The coded binary signals

2	0	-2
11	01	00

Table 8.1: Coding of the three-level weights.

are generated using EPROMs. One 27C64 EPROM can generate 4 weight signals and therefore four EPROMs are required for the 16 weight signals. The 2-bit DAC includes two 4 K resistors and an op-amp NC952. The potentialmeters in the 2-bit DAC are used to adjust the voltage level and amplitude of the weight signal. LS193 is a 4-bit up/down counter and is used to generate the addresses for the 27C64 EPROMs. The oscillographs of the 16 weight signals are shown in Fig. 8.13, where $W_1 \sim W_4$ are displayed in Fig. 8.13(a), $W_5 \sim W_8$ are displayed in Fig. 8.13(b), $W_9 \sim W_{12}$ are displayed in Fig. 8.13(c), and $W_{13} \sim W_{16}$ are displayed in Fig. 8.13(d).

The ANN IC was evaluated by constructing a list of input patterns derived from a stored pattern. One of the input patterns is introduced to the network during

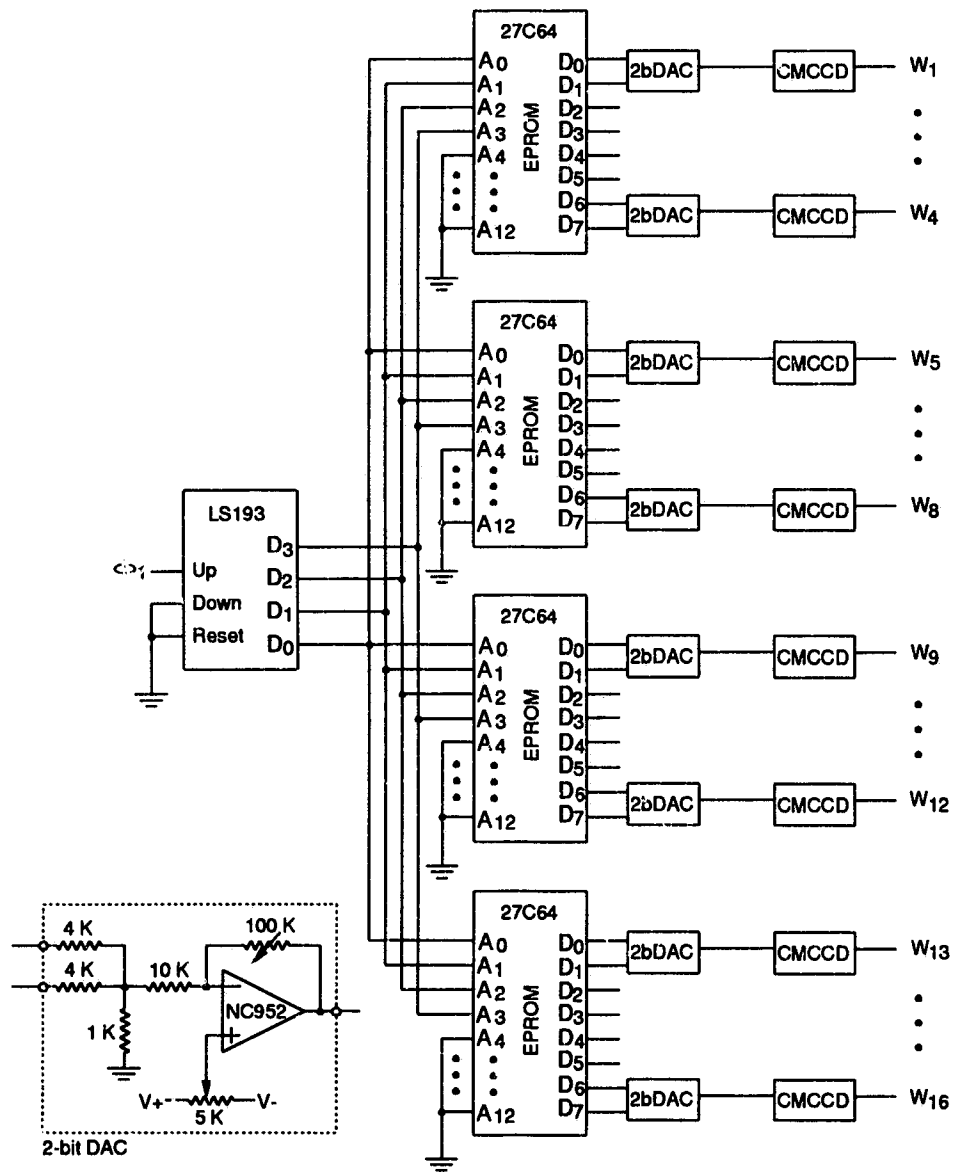


Figure 8.12: Schematic of the circuit used to generate the three-level synaptic weights.

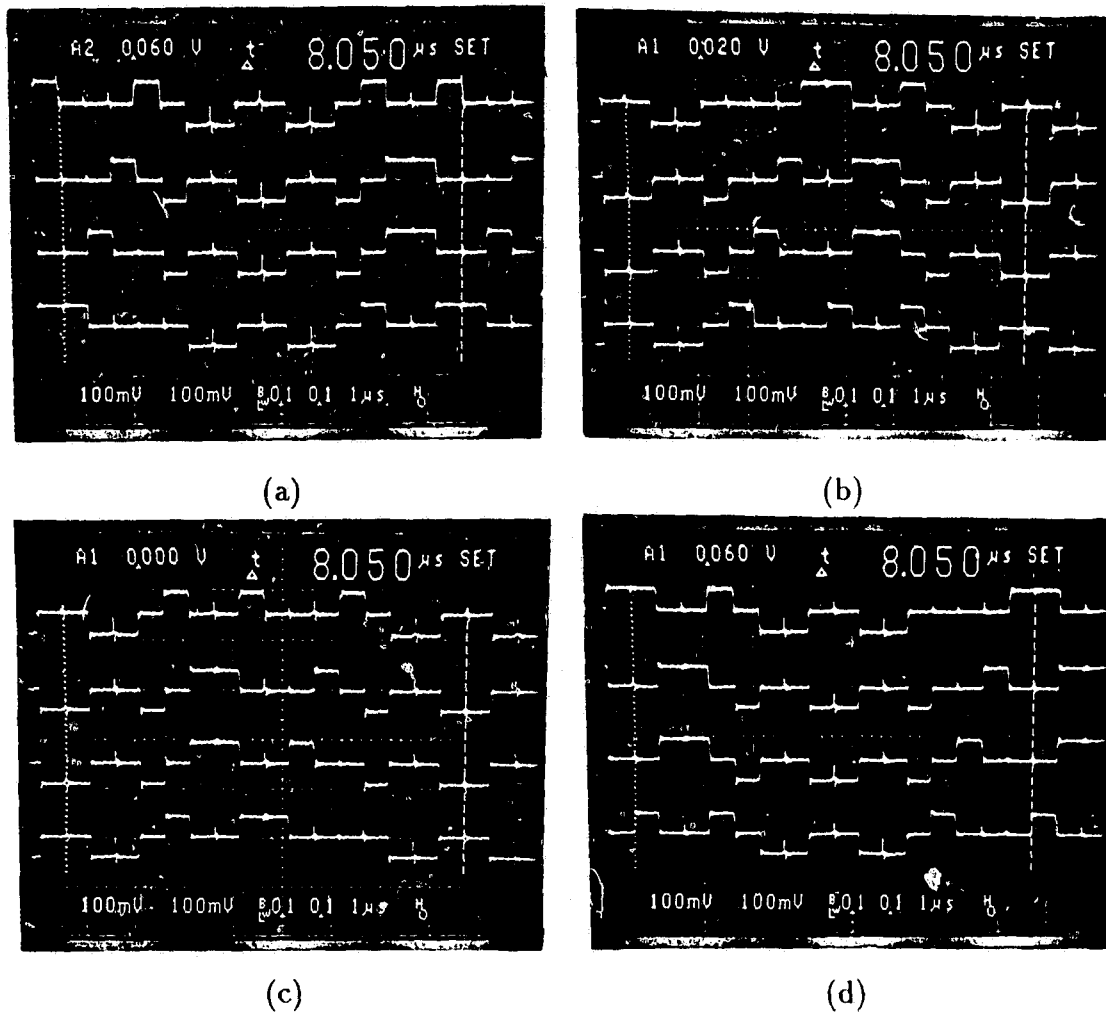


Figure 8.13: Oscillographs of the 16 weight signals.

the first pulse of ϕ_2 . After the network reaches its stable state we can determine if the output pattern is one of the stored patterns. The input patterns are derived by inverting each bit of one of the stored patterns from +1 to -1 or vice versa with a fixed percentage chance (called as the bit-error rate). The results of the test is given by the correct recall rate with respect to the bit-error rate, and is plotted in Fig. 8.14. The results obtained from the Hopfield model [79] in fact coincide very

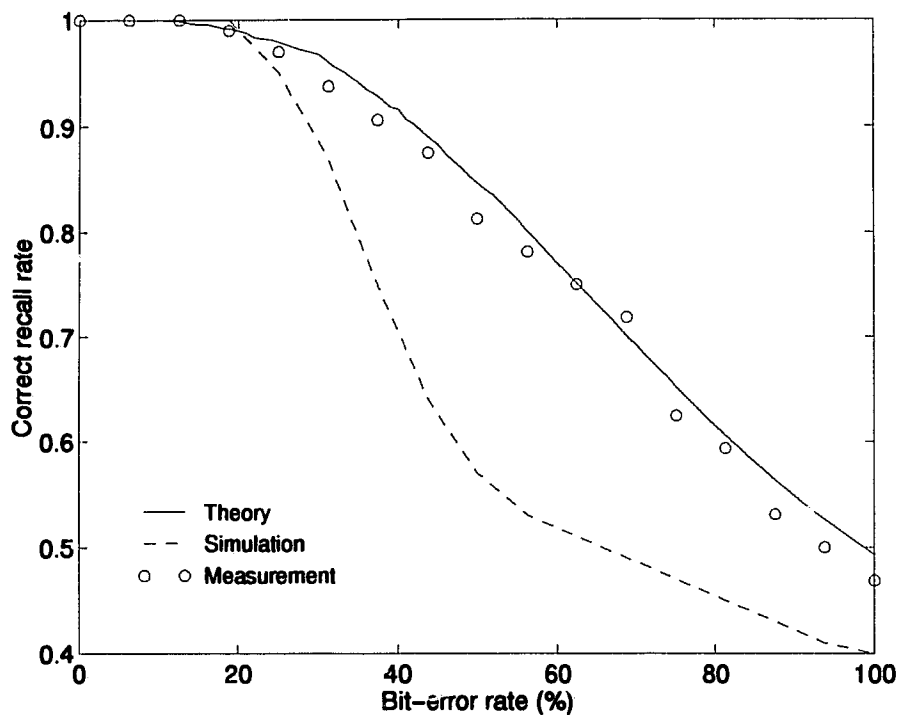


Figure 8.14: Correct recall rate versus bit-error rate.

well with that from the measurement. The deviation between the simulation and the measurement when the bit-error rate is greater than 20% is due to the insufficient simulation times. Therefore, the simulation curve is not accurate. It is noted that even when the percentage bit error rate reaches 100, the correct recall rate is between 0.5 and 0.4. This is due to the fact that the patterns used in the test have a 50% overlap even when all of the bits in one of the patterns are inverted.

As an example, Figs. 8.15 and 8.16 show four different input patterns and the correspondent output patterns of the ANN IC. When the input pattern is closely associated to one of the stored patterns, the chip will recognize it, as indicated by Fig. 8.16(a) and (b). However, when the difference between the input pattern and the stored patterns is large enough (i.e., greater than 25% as observed in Fig. 8.14), the chip may fail to recognize the input pattern and reaches a stable state which is different from all the stored patterns, as observed in Fig. 8.16(c) and (d). These spurious patterns (as called by Lippmann in [39]) represent the major problem suffered from by the Hopfield ANN when it is used as an associative memory. This problem can be reduced or eliminated either by storing a small number of the patterns which share few bits in common, or by a number of orthogonalization procedures [39].

Fig. 8.17 shows the oscillographs of the clocks ϕ_1 and ϕ_2 and the output of the chip (V_{out}) for the different input patterns. The subfigures (a), (b), (c), and (d) correspond to the patterns shown in Fig. 8.16. The time slots within the vertical cursors in the oscillographs represent one cycle of computation. Note that the simulation waveforms shown in Figs. 8.5 and 8.6 have been verified functionally by Fig. 8.17 (a) and (b).

The total power dissipation of the ANN IC was found to be 565 mW. The percentage of the power consumed by the subcircuits are summarized in Table 8.2. As can be seen from Table 8.2, the shift- and buffer-registers consumed more than 2/3 of the the total power of the chip.

WSC	Activation circuit	Shift- and Buffer-registers	Other
4	23	68	2

Table 8.2: Percentage of the power consumed by the subcircuits of the ANN chip.

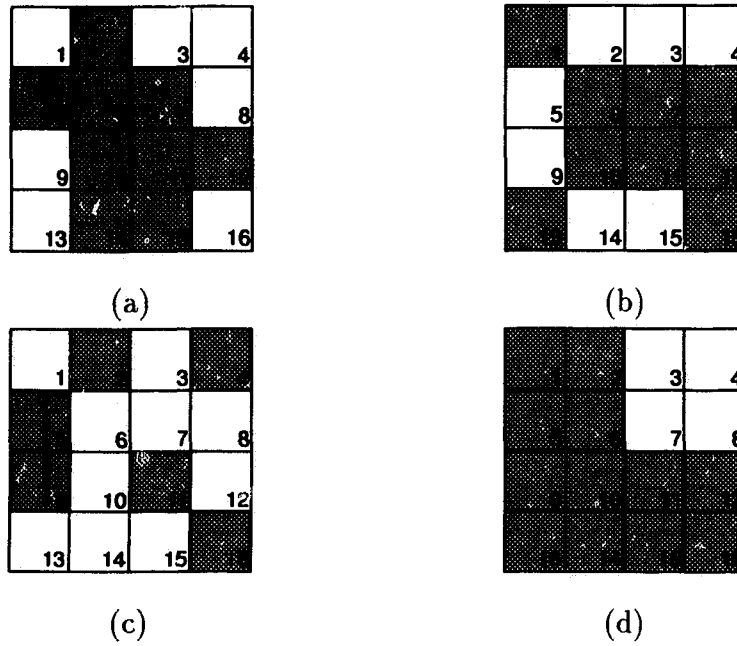


Figure 8.15: Input patterns used in the ANN chip test.

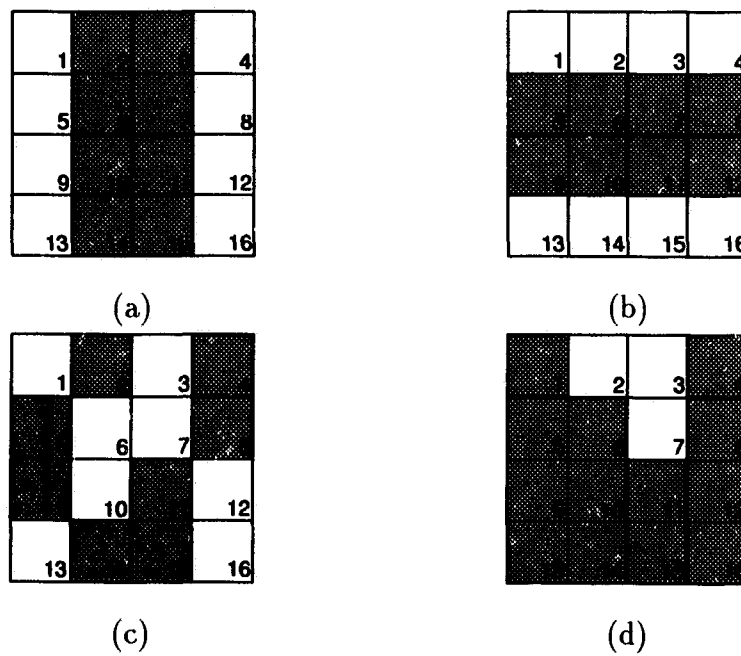


Figure 8.16: Output patterns from the ANN chip.

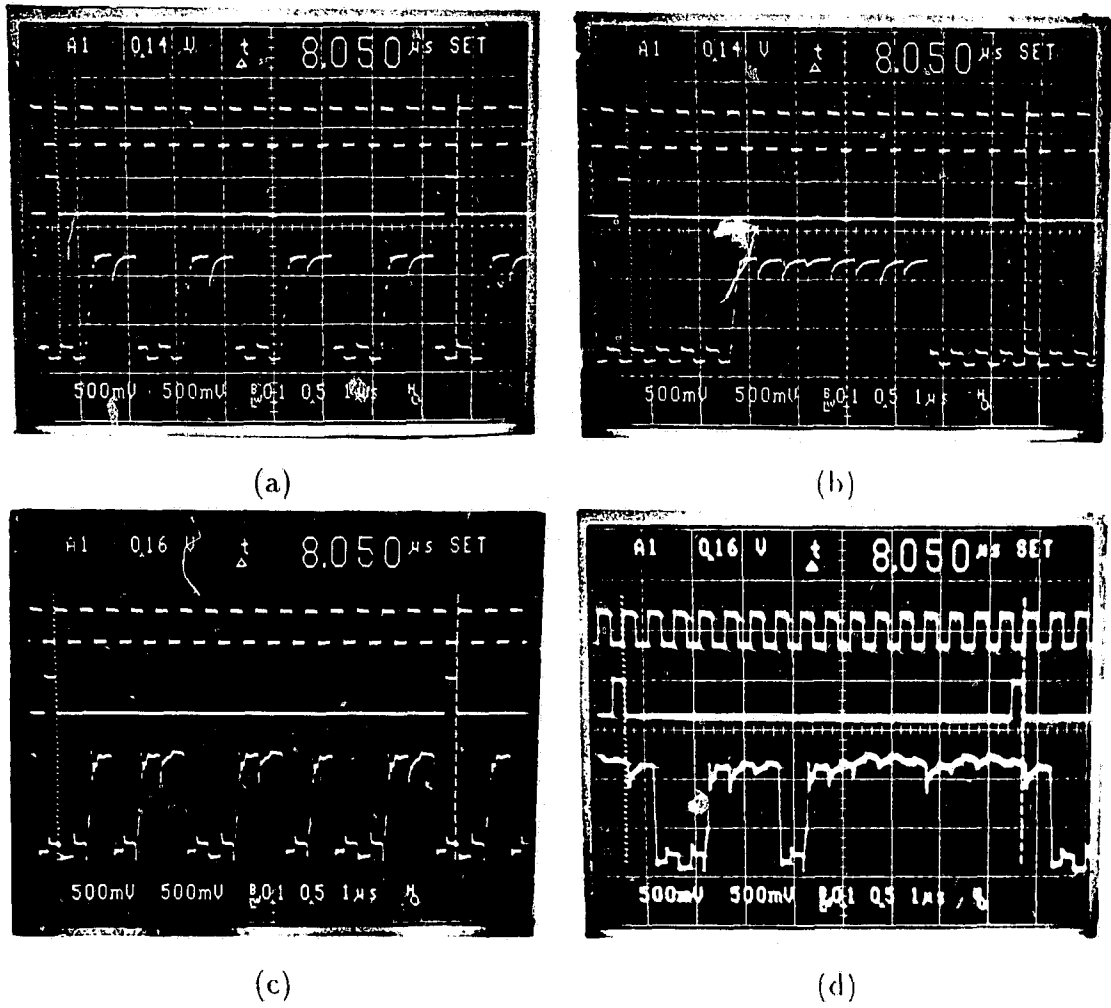


Figure 8.17: Oscillographs of the clocks ϕ_1 and ϕ_2 and the output of the chip for the different input patterns.

8.5 Discussions

The performance of the ANN IC has been successfully demonstrated by a series of test patterns. The problems related to the fabrication and the measurement are discussed in the following.

8.5.1 Fabrication

The original motivation of this work was to have the CMCCD fabricated using the NT/BNR process and then to integrate our MESFET ANN circuit monolithically with 16 16-pixels CMCCDs for the storage of the synaptic weights using the NT/BNR process. However, we found that the NT/BNR process is not suitable for the fabrication of GaAs CCDs. One of the reasons is that the NT/BNR process is strictly optimized for MESFET fabrication. Table 8.3 lists the main process parameters of the NT/BNR GaAs technology. Another technology limitations is the lack of a cer-

Parameter	Value
Active layer thickness	0.1 μm
Implant doping density	$2.1 \times 10^{18} / \text{cm}^3$
Pinch-off voltage	-1.2 V
Min. gate metal space	2 μm
Cermet gate	No

Table 8.3: Main process parameters of the NT/BNR GaAs technology.

met gate process in the NT/BNR GaAs process. Deyhimy *et al.* [70] showed that a GaAs CCD without cermet gate could be fabricated on a deep n-type epitaxial layer ($> 1 \mu\text{m}$) grown on a semi-insulating GaAs substrate, and the doping density for the

n-type epitaxial layer would have to lie within the range from $10^{15} / \text{cm}^3$ to $10^{17} / \text{cm}^3$. They also demonstrated that an interelectrode gap with a length greater than $1 \mu\text{m}$ would result in the formation of potential troughs. The process parameters required for the GaAs CCD fabrication obviously do not match those for the MESFETs as provided by the NT/BNR process which uses a thin and highly doped active layer and has a minimum gate metal space of $2 \mu\text{m}$.

It should be pointed out that not only the process parameters described above but also the masking can be a problem. A CCD has a continuous channel doping under the entire gate metal area in order for the device to function. However, the NT/BNR process is a self-alignment MESFET process. It turns out that there is a n^+ region in each of the spaces between the gate metal since in the NT/BNR process it is the gate metal that acts as the n^+ mask which is used for the formation of the source and the drain of the MESFET. It is therefore not possible to form a CCD using this process.

8.5.2 Linearity of the Weight-summation Operation

The accuracy of the neuron weight-summation operation is of importance for the neural network. It is found from Fig. 8.9 that the measured linearity of the weight-summation operation has slight divergence from the ideal case. This degradation, as we discussed in Chapter 6, seems to be caused by the deviations in the sizes of the matched devices due to the fabrication process. This degradation in linearity can also be observed from the lower trace in Fig. 8.11(b), where the middle voltage levels of the differential outputs of the WSC at different times are slightly different. By optimizing the physical layout of the chip and improving the process, the linearity of the weight-summation operation can be improved.

8.5.3 Glitches in the Output Waveforms of the ANN IC

It is noted from both the simulations and the measurements that the output waveforms of the activation circuit (V_{out}) contain glitches, as can be seen from Figs. 8.11 and 8.17. The glitches occur whenever there is a high-to-low (or low-to-high) transition in the waveform of either the weight or the input of the neuron state. The finite rise time and fall time of the input signals contribute to the glitch generation. The glitches can be generated during the signal multiplication, and amplified by the activation circuit. In general, the glitches are found to be much narrower than the clock signals, and much smaller in amplitude than the output of the activation circuit. Therefore, they can not affect the operation of the chip unless the clock frequency is too high (> 50 MHz).

8.5.4 Limitations on High-frequency Measurement

As indicated in Chapters 6, 7 and 8, the ANN IC and the other subcircuit ICs gave good performance at a frequency up to 50 MHz. Testing at a higher frequency is limited by the parasitic effects of the devices under testing as well as those related to the chip packaging and the input signals.

A GaAs die usually requires LCC (Leadless Chip Carrier) package. However, the LCC packaging available from NT/BNR and CMC (Canadian Microelectronics Cooperation) only has 44 pins which is much less than the number of pins of our chips. The packages used for our GaAs dies are PGA (Pins Grid Array) which is usually used for CMOS die packaging. The PGA does not have proper 50Ω termination, and only works at a speed below 200 Mb/s.

The high frequency test is also limited by the signal generators. The maximum frequency our signal generators produce is 500 MHz (i.e., the rise time and the fall equal 1 ns and the width of the pulse is 0 ns). For a clock pulse with a width of 2 ns

and a rise (or fall) time of 1 ns, the frequency reduces to 160 MHz. The measurement of the ANN chip as an associative memory also depends on the 16 weight signals from the external circuit. The circuit demonstrated its highest operation frequency of 40 MHz which was limited by the highest operation frequency of the NC952 op-amp, the 27C64 EPROMs and the LS193 up/down counter.

It is hopeful that a revised version of the ANN IC would be able to operate at a much higher frequency when these limitations are eliminated.

Chapter 9

Summary and Recommendations

9.1 Summary of the Results

Contributions have been made to develop a design methodology to optimize the GaAs CMCCD structure, and to design GaAs CMCCDs and MESFET circuits for the artificial neural network (ANN) application. The results of this research are summarized in this section.

Two-dimensional models of computing the potential distribution and the signal charge transfer in the CMCCD were developed in Chapter 2. The models were based on the Poisson equation and the continuity equation. The CMCCD was assumed to consist of a uniformly doped active layer and a semi-insulating substrate. The finite difference method was used to solve the equations involved in the numeric models. The two-dimensional charge transfer model was shown to improve the accuracy of computation by an average of 3% over the one-dimensional model developed in [53]. The efficiency of the CMCCD equivalent circuit model developed earlier was improved by using an analytical fringing field expression.

A design methodology was developed to optimize the structures of the two-phase

and the uni-phase CMCCDs, as discussed in Chapter 3. An average square error between the ideal and computed potential distribution was defined as the objective function of the optimization. The potential distribution in the CMCCD channel was computed using the numeric model described in Chapter 2. The Fletcher-Reeves method, a conjugate-direction-approach-based optimization algorithm, was used in the optimization. Charge transfer in the optimized two-phase and uni-phase CMCCDs was simulated to evaluate the charge transfer efficiency with respect to the transfer time, the size of charge packet, and the amplitude of clock voltage. The simulation results indicated that the two-phase CMCCD may potentially operate at the clock frequency of 1.25 GHz with a charge transfer efficiency up to 0.9987 and the uni-phase CMCCD at the clock frequency of 500 MHz with a charge transfer efficiency up to 0.9990. In general, the two-phase CMCCDs have larger charge-handling capability, higher speed than the uni-phase CMCCDs. The effects of the clock waveforms on the charge transfer efficiency were investigated using SPICE simulation incorporated with the equivalent circuit model. The simulation has showed that the rise time plays an important role in determining the charge transfer efficiency, whereas the fall time has almost no influence on the charge transfer.

The fabrication and measurement of the uni-phase CMCCD using the TRIUMF GaAs process was described in Chapter 4. The measurement included the I - V characteristics, the charge transfer efficiency for the different clock frequency, and the effect of the leading edge of the clock voltages on the charge transfer efficiency. The noise-to-signal ratio and the linearity between the charge injection and detection were also measured but presented in Chapter 6. The experimental results demonstrated that the fabricated CMCCD can operate properly within the clock frequency range from 100 KHz up to 100 MHz with a charge transfer efficiency of 0.9810.

A hybrid architecture of the Hopfield ANN implementation using CMCCDs as analog synaptic weight storages was proposed in Chapter 5. This approach merges

parallel multiplication/summation and serial neuron state updating and only needs one adder and one activation circuit, and therefore, reduces significantly the power dissipation, chip area as well as the complexity of the interconnections. The comparison of the parallel, serial, and hybrid ANN architectures suggested that the hybrid architecture ranks first in power dissipation and chip area and second in computation speed among the three architectures. The overview of the operation of the hybrid ANN has shown that if the clock frequency is 100 MHz and 1000 neurons are integrated, the hybrid ANN will update all the neuron states in 10 μ s.

The hybrid architecture consists of a synaptic weight storage array, a weight-summation circuit (WSC), an activation circuit, a N -bit shift-register, and a N -bit buffer-register. The GaAs CMCCD and MESFET designs of these subcircuits were given in Chapters 6 and 7. The weight storage array is composed of N N -bit CMCCD analog memories which include a self-connected N -pixel CMCCD, a 2-to-1 multiplexer to control the write and read operations and a level-shifting circuit. The experimental results demonstrated the function of the CMCCD serial analog memory with very low noise ($S/N = 60$ dB), a large dynamic range (> 38 dB), and low on-chip power dissipation. The WSC was designed using N MESFET transconductance amplifiers, and the activation circuit was designed using an operational amplifier configuration to give the hard-limiting function. Two versions of the shift- and buffer-registers, i.e., CMCCD and MESFET shift- and buffer-registers were designed. The comparison between two versions of the shift- and buffer-registers showed that the CMCCD shift- and buffer-registers consumes much smaller power as well as smaller chip area than the MESFET counterparts. All the MESFET circuits were simulated and fabricated using the NT/BNR 0.8 μ m GaAs depletion-mode MESFET process. The experimental results have verified the circuit design and simulation.

The fabrication and measurement of a GaAs MESFET ANN IC were described in Chapter 8. The ANN IC was fabricated using the NT/BNR 0.8 μ m GaAs depletion-

mode MESFET process. Because the NT/BNR process is not able to support the fabrication of the CCD, the chip only consists of a WSC with 16 transconductance amplifiers, an activation circuit, a 16-bit shift-register, and a 16-bit buffer register. A multi-chip approach was therefore taken and the CMCCDs were fabricated using the TRIUMF GaAs process. The synaptic weights were generated externally using EPROMs and 2-bit DACs. A series of measurement was made to evaluate the performance and the function of the ANN IC. The experimental results have shown that the ANN chip has a good performance in terms of the linearity of the weight-summation computation and the neuron hard-limiting function. By storing two encoded patterns (in the form of analog weights) into the synapses and applying a pattern to its inputs, the chip can reach a stable state which is the nearest match to the input pattern if the difference between the input pattern and one of the stored patterns is small enough ($< 25\%$). Thus we have demonstrated the functionality of the chip as an associative memory. The discussions on the fabrication and the testing of the ANN IC were also given.

9.2 Recommendations for Future Work

9.2.1 ANN Circuit Design

As mentioned earlier, the original motivation of this work was to integrate the MESFET neural processing circuit monolithically with CMCCD analog storage array. Because of the limitations of the GaAs IC fabrication process, we cannot achieve this goal at the moment. Our results seem to indicate that to fully utilize the speed advantage of the GaAs ICs, a CMCCD/MESFET compatible technology is essential.

We have demonstrated the function of the CMCCD serial analog memory (Chapter 6). The performance of the CMCCD serial analog memory at high frequency

can be improved if the level-shifting circuit and the write/read control circuit in the memory loop are integrated monolithically with the CMCCD. To achieve this, a CMCCD/MESFET compatible technology needs to be developed.

It was observed in the measurement that the main power dissipation of the ANN chip is caused by the shift- and buffer-registers and the activation circuit. We have shown that the CMCCD shift- and buffer-registers consume much less power than the MESFET shift- and buffer-registers. The MESFET shift- and buffer-registers can therefore be replaced by the CMCCD counterpart. This substitution will greatly reduce the power dissipation as well as the area of the ANN chip. Again, a CMCCD/MESFET compatible technology is required.

The ANN circuit we developed is actually a mixed-mode IC. With the increase in the applications of GaAs ICs, especially in the optoelectronic applications, more and more mixed-mode very large-scale ICs are required. Therefore it is necessary to develop a methodology for the GaAs mixed-mode IC design. Many approaches for the CMOS mixed-mode IC design have been developed but very few for the GaAs mixed-mode IC was reported.

9.2.2 Synaptic Weights Loading and Learning

In this work, the synaptic weights were loaded into the CMCCDs serially. This is not convenient for the programming and dynamic learning of the synaptic weights. A promising development will replace the CMCCD storage array by CMCCD detector arrays which enable optical parallel loading of the synaptic weights and the initial neuron states, and make the ANN work in an optoelectronic environment.

The number of the patterns stored in the Hopfield ANN (also called as the pattern storage capacity of the ANN) is determined by the learning rules. The learning rule we used (Eq. (5.10)) has limited the pattern storage capacity of the network. In

addition, to tolerate such problems as dynamic range and signal-to-noise ratio of the analog weight memory, the synaptic weights can take only a number of values (or multiple values) instead of fully analog values. This also causes a decrease in the pattern storage capacity. Therefore, a new learning rule needs to be developed to allow a good pattern storage capacity with only restricted number of weight values.

9.2.3 Large Scale Integration of the ANN

An ANN chip with more than 16 neurons can be realised in the same way as we discussed in this thesis. However, because of the limitation of the number of pins provided by the current packaging technology, on-chip multiplexers are needed to insert between the input of the ANN and the shift- and buffer-registers.

It is noted that the hybrid ANN IC has the ability to realize different types of the neural networks in large scale. By cascading two chips , a forward propagation network (or so-called two layer perceptron network) can be achieved. Or, by using two chips in parallel, a network with more interconnections between adjacent layers can be realized.

Bibliography

- [1] W. W. Hooper and W. I. Lehrer, "An epitaxial GaAs field-effect transistor," *Proc. IEEE*, vol. 55, pp. 1237–1245, 1967.
- [2] W. S. Boyle and G. E. Smith, "Charge coupled semiconductor devices," *Bell Syst. Tech. J.*, vol. 49, pp. 587–592, 1970.
- [3] G. F. Amelio, M. F. Tompsett, and G. E. Smith, "Experimental verification of the charge-coupled devices concept," *Bell Syst. Tech. J.*, vol. 49, pp. 593–600, 1970.
- [4] F. L. Schuermeyer, R. A. Belt, C. R. Young, and J. M. Blasingame, "New structure for charge-coupled devices," *Proc. IEEE*, 60, pp. 1444–1445, 1972.
- [5] W. Kellner, H. Bierhenke, and H. Kniepkanp, "A Schottky barrier CCD on GaAs," in *IEDM Tech. Dig.*, pp. 599–602, 1977.
- [6] I. Deyhimy, J. S. Harris, Jr., R. C. Eden, D. D. Edwall, S. J. Anderson, and L. O. Bubelac, "GaAs charge-coupled devices," *Appl. Phys. Lett.*, 32 (6), pp. 383–385, 1978.
- [7] I. Deyhimy, R. C. Eden, A. J. Anderson, and J. S. Harris, Jr., "A 500-MHz GaAs charge-coupled device," *Appl. Phys. Lett.*, 36 (2), pp. 151–153, 1980.

- [8] W. Kellner, U. Ablasmeier, and H. Kniepkamp, "A two-phase CCD on GaAs with 0.3- μm -wide electrode gaps," *IEEE Trans. Electron Devices*, vol. 27, No. 5, pp. 1195–1197, 1980.
- [9] I. Deyhimy, W. A. Hill, and R. J. Anderson, "Continuously clocked 1 GHz CCD," *IEEE Electr. Dev. Lett.*, vol. 2, pp. 70–72, 1981.
- [10] K. B. Nichols and B. E. Burke, "A gallium arsenide overlapping gate charge-coupled device," *IEEE Electr. Dev. Lett.*, vol. 6, pp. 237–240, 1985.
- [11] A. Hayes, J. Davies and E. Eccleston, "Operation of a single phase CCD on GaAs at 560 MHz," *IEE Proc. Pt. I*, vol. 132, No.1, pp. 34–36, 1985.
- [12] L. J. M. Esser and F. L. J. Sangster, in *Handbook on Semiconductors* (C. Hilsum, ed.), vol. 4, pp. 335–421, North-Holland Publishing Co. Amsterdam, 1981.
- [13] R. C. Eden and I. Deyhimy, "Application of GaAs integrated circuit and CCDs for high speed signal processing," in *IEEE Electronics and Aerospace Systems Conf.*, pp. 78–86, 1979.
- [14] R. Sahai, W. A. Hill, B. Mathur, S. P. Pittman, and J. A. Higgins, "GaAs CCDs for analog signal processing," in *IEEE 1986 CICC*, pp. 521–527, 1986.
- [15] R. E. Colbeth, D. V. Rossi, J.-I. Song, and E. R. Fossum, "GHz GaAs CCDs: Promises, problems and progress," in *Proc. SPIE*, vol. 1071, pp. 108–114, 1989.
- [16] R. E. Colbeth, J.-I. Song, D. V. Rossi, and E. R. Fossum, "A recessed-gap capacitive-gate GaAs CCD," *IEEE Electr. Dev. Lett.*, vol. 10, pp. 525–527, 1989.
- [17] J. A. Higgins, R. A. Milano, E. A. Sovero, and R. Sahai, "Resistive-gate GaAs charge-coupled devices," in *GaAs IC Symp. Tech. Dig.*, pp. 49–52, 1982.

- [18] E. Sovero, R. Sahai, W. Hill, and J. Higgins, "Microwave frequency GaAs charge-coupled devices," in *GaAs IC Symp. Tech. Dig.*, pp. 101–104, 1984.
- [19] J.-I. Song and E. R. Fossum, "Characterization of evaporated $C_r - SiO$ cermet films for resistive-gate CCD applications," *IEEE Trans. Electron Devices*, vol. 36, No. 9, pp. 1575–1579, 1989.
- [20] J.-I. Song, D. V. Rossi, S. Xin, W. J. Wang, and E. R. Fossum, "A resistive-gate $Al_{0.3} Ga_{0.7}As/GaAs$ 2DEG CCD with high charge transfer efficiency at 1GHz," *IEEE Trans. Electron Devices*, vol. 38, No. 4, pp. 930–933, 1991.
- [21] J.-I. Song and E. R. Fossum, "Inhibition of charge packet broadening in GaAs charge-coupled devices," *Appl. Phys. Lett.*, vol. 51, No. 9, pp. 1539–1541, 1987.
- [22] D. Rossi, J.-I. Song, E. Fossum, P. Kirchner, G. Pettit, and J. Woodall, "A resistive-gate $In_{0.53} Ga_{0.47} As/InP$ heterostructure CCD," *IEEE Electr. Dev. Lett.*, vol. 12, No.12, pp. 688–690, 1991.
- [23] M. LeNoble, J. V. Cresswell, and L. Young, "The surface potential variation in the interelectrode gap of GaAs cermet-gate charge-coupled devices," *Solid-State Electronics*, vol. 33, No.7, pp. 851–857, 1990.
- [24] M. LeNoble, J. V. Cresswell, and R. R. Johnson, "A two-phase GaAs cermet gate charge-coupled device," *IEEE Trans. Electron Devices*, vol. 37, No.8, pp. 1796–1799, 1990.
- [25] M. V. Howes and D. V. Morhan, *Charge-coupled Devices and Systems*, John Wiley Sons, 1979.
- [26] E. A. Sovero, W. A. Hill, R. Sahai, J. A. Higgins, S. Pittman, E. H. Martin, and R. L. Pierson, Jr., "Transversal filter application of a high speed gallium arsenide CCD," in *GaAs IC Symp. Tech. Dig.*, pp. 92–95, 1983.

- [27] W. A. Hill, E. A. Sovero, J. A. Higgins, E. H. Martin, and S. Pittman, "1GHz sample rate GaAs CCD transversal filter," in *GaAs IC Symp. Tech. Dig.*, pp. 27–30, 1985.
- [28] R. E. Howard, D. B. Schwartz, J. S. Denker, R. W. Epworth, H. P. Graf, W. E. Hubbard, L. D. Jackel, B. L. Straughn, and D. H. Tennant, "An associative memory based on an electronic neural network architecture," *IEEE Trans. Electron Devices*, vol. 34, No. 7, pp. 1553–1555, 1987.
- [29] M. Verleysen, B. Sirletti, A. Vandemenlebroecke, and P. G. A. Jespers, "Neural networks for high-storage content-addressable memory: VLSI circuit and learning algorithm," *IEEE J. Solid-state Circuits*, vol. 24, No. 3, pp. 562–569, 1989.
- [30] N. I. Khachab and M. Ismail, "A nonlinear CMOS analog cell for VLSI signal and information processing," *IEEE J. Solid-state Circuits*, vol. 26, No. 11, pp. 1689–1699, 1991.
- [31] C. R. Schneider and H. C. Card, "Analog CMOS deterministic Boltzmann circuits," *IEEE J. Solid-state Circuits*, vol. 28, No. 8, pp. 907–914, 1993.
- [32] H. P. Graf and L. D. Jackel, "Analog electronic neural network circuits," *IEEE Circuits and Devices Mag.*, vol. 5, No. 4, pp. 44–49, 1989.
- [33] J. M. Zurada, "Analog implementation of neural networks," *IEEE Circuits and Devices Mag.*, vol. 8, No. 5, pp. 36–41, 1992.
- [34] A. M. Chiang, "A CCD programmable signal processor," *IEEE J. Solid-state Circuits*, vol. 25, No. 6, pp. 1510–1517, 1990.
- [35] J. P. Sage, R. S. Withers, and K. E. Thompson, "MNOS/CCD circuits for neural network implementations," in *Proc. IEEE 1989 ICCAS*, pp. 1207–1209, 1986.

- [36] A. Agranat and A. Yariv, "Semi-parallel microelectronics implementation of neural network model using CCD technology," *Electron. Lett.*, vol. 23, No. 11, pp. 580–581, 1987.
- [37] A. J. Agranat, C. F. Neugebauer, R. D. Nelson, and A. Yariv, "The CCD neural processor: a neural network integrated circuit with 65536 programmable analog synapses." *IEEE Trans. Circuits Sys.*, vol. 37, No. 8, pp. 1073–1075, 1990.
- [38] J. J. Hopfield, "Neural network and physical systems with emergent collective computational ability," *Proc. Natl. Acad. Sci.*, vol. 79, pp. 2556–2558, 1982.
- [39] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, vol. 4, pp. 4–19, 1987.
- [40] E. R. Fossum, "Wire transfer of charge packets using a CCD-BBD structure for charge-domain signal processing," *IEEE Trans. Electron Devices*, vol. 38, No. 2, pp. 291–298, 1991.
- [41] E. R. Fossum, S. E. Kemeny, R. A. Bredthauer, and M. LaShell, "Digital programmable gain control circuit for charge-domain signal processing," *IEEE J. Solid-state Circuits*, vol. 26, No. 4, pp. 683–686, 1991.
- [42] S. Satyanarayana, Y. P. Tsividis, and H. P. Graf, "A reconfigurable VLSI neural network," *IEEE J. Solid-state Circuits*, vol. 27, No. 1, pp. 67–81, 1992.
- [43] D. B. Schwartz and R. E. Howard, "A programmable analog neural network chip," in *Proc. IEEE 1988 CICC*, pp. 10.2.1–10.2.4, 1988.
- [44] J. Raffel, J. Mann, R. Bergner, A. Soares, and S. Gilbert, "A generic architecture for wafer-scale neuromorphic system," in *Proc. IEEE First Int. Conf. Neural Networks*, vol. IV, pp. 485–493, 1987.

- [45] R. Tawel, R. Benson, and A. P. Thakoor, "A CMOS UV-programmable non-volatile synaptic array," in *Proc. IEEE IJCNN-91*, pp. 581–585, 1991.
- [46] G. Cauwenberghs, C. F. Neugebauer, and A. Yariv, "An adaptive CMOS matrix-vector multiplier for large scale analog hardware neural network applications," in *Proc. IEEE IJCNN-91*, pp. 507–511, 1991.
- [47] C. F. Neugebauer and A. Yariv, "A parallel analog CCD/CMOS neural network IC," in *Proc. IEEE IJCNN-91*, pp. 447–451, 1991.
- [48] J.-Y. Chang and C.-C. Wu, "Design of Hopfield-type associative memory with maximal basin of attraction," *Electron. Lett.*, vol. 29, No. 24, pp. 2128–2130, 1993.
- [49] P. Treleaven, M. Pacheco and M. Vellasco, "VLSI architectures for neural network," *IEEE Micro*, No. 12, pp. 8–26, 1989.
- [50] A. F. Murray and A. V. W. Smith, "Asynchronous VLSI neural networks using pulse-stream arithmetic," *IEEE J. Solid-state Circuits*, vol. 23, No. 3, pp. 688–697, 1988.
- [51] B. W. Lee and B. J. Sheu, "General-purpose neural chips with electrically programmable synapses and gain-adjustable neurons," *IEEE J. Solid-state Circuits*, vol. 27, No. 9, pp. 1299–1302, 1992.
- [52] M. Wedlake, "On the operations of neural networks with detail of a mixed mode neural network implementation," Research Report, University of Victoria, 1993.
- [53] S. Pennathur, "Analysis and modeling of charge transport in a GaAs CCD," M.Sc. Thesis, University of Victoria, 1991.

- [54] S. Pennathur and H. H. L. Kwok, "Simulation of charge transfer in GaAs cermet-gate CCDs," *IEEE Trans. CAD*, vol. 11, No. 7, pp. 903–910, 1992.
- [55] S. Pennathur and H. H. L. Kwok, "Equivalent circuit for GaAs CCD," *IEE Proc. Pt. G*, vol. 140, No. 6, pp. 377–382, 1993.
- [56] J. M. Carxes, W. F. Kasonocky, and E. G. Ramberg, "Drift-aided fringing field in charge-coupled devices," *IEEE J. Solid-state Circuits*, vol. 6, No. 10, pp. 322–326, 1971.
- [57] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, 1987.
- [58] A. Antoniou, *Introduction to Optimization Theory and Practice*, University of Victoria, pp. 139–154, 1992.
- [59] L. Chen and H. H. L. Kwok, "Optimized design of a 2-phase GaAs CMCCD and its charge transfer simulation," *Solid-State Electronics*, vol. 37, No. 11, pp. 1863–1869, 1994.
- [60] L. Chen and H. H. L. Kwok, "An Optimized structure of a 2-phase GaAs CMCCD," in *Proc. ISDRS*, pp. 519–522, 1993.
- [61] L. Chen and H. H. L. Kwok, "A comparison study of the uni-phase and the two-phase GaAs CMCCDs" submitted to *Solid-State Electronics*.
- [62] L. Chen and H. H. L. Kwok, "Effects of the clock voltage waveforms on the charge transfer inefficiency of high-speed charge-coupled devices," *Solid-State Electronics*, vol. 36, No. 5, pp. 798–800, 1993.
- [63] L. Chen and H. H. L. Kwok, "A hybrid architecture for Hopfield neural network and its GaAs implementation using CCDs as storage elements," submitted to *IEEE Trans. VLSI*.

- [64] R. W. Broderson and S. P. Emmons, "Noise in buried channel charge-coupled devices," *IEEE J. Solid-state Circuits*, vol. 11, No. 2, pp. 147–155, 1976.
- [65] M. Shur, *Physics of Semiconductor Devices*, Prentice Hall, 1990.
- [66] J. D. E. Beynon and D. R. Lamb, *Charge-coupled Devices and Their Applications*, McGraw-Hill, 1980.
- [67] C. M. Miller, *Modern Electronic Communication*, Prentice Hall, 1978.
- [68] C. H. Sequin and M. F. Tompsett, *Charge Transfer Devices*, Academic Press, New York, 1975.
- [69] R. W. Broderson, D. D. Buss, and A. F. Tasch, Jr., "Experimental characterization of transfer efficiency in charge-coupled devices," *IEEE Trans. Electron Devices*, vol. 22, No. 1, pp. 40–46, 1975.
- [70] I. Deyhimy, R. C. Eden, and J. S. Harris, Jr., "GaAs and related heterojunction charge-coupled devices," *IEEE Trans. Electron Devices*, vol. 27, No. 6, pp. 1172–1180, 1980.
- [71] S. Katsu, M. Kazumura, and G. Kano, "Design and fabrication of a GaAs monolithic operational amplifier," *IEEE Trans. Electron Devices*, vol. 35, No. 7, pp. 831–838, 1988.
- [72] L. E. Larson, J. F. Jensen, H. M. Levy, P. T. Greiling, and G. C. Temes, "GaAs differential amplifier," in *GaAs IC Symp. Tech. Dig.*, pp. 19–22, 1985.
- [73] N. Scheinberg, "High-speed GaAs operational amplifier," *IEEE J. Solid-state Circuits*, vol. 22, No. 4, pp. 522–527, 1987.
- [74] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, Holt, Rinehart and Winston, Inc., 1987.

- [75] N. Kanopoulos, *Gallium Arsenide Digital Integrated Circuits*, Prentice Hall, 1989.
- [76] S. I. Long and S. E. Butner, *Gallium Arsenide Digital Integrated Circuit Design*, McGraw-Hill Pub. Co., 1990.
- [77] Y. S. Abu-Mostafa and J. St. Jacques, "Information capacity of the Hopfield model," *IEEE Trans. Inform. Theory*, vol. 31, pp. 461–464, 1985.
- [78] T. E. Dillinger, *VLSI Engineering*, Prentice Hall, pp. 64–69, 1988.
- [79] M. Wedlake, Private communication, 1993.

Appendix A

Derivation of the Electric Field in CMCCD

When the signal charge resides in the CMCCD channel, the channel potential is given by [70]:

$$V(x, y) = \psi(x, y) + \frac{qN_d(x_t - n_s(x, y)/N_d)^2}{2\epsilon_s} \quad (\text{A.1})$$

The first term in Eq. (A.1) is the channel potential in the absence of the signal charge, and the second term is the channel potential only resulted from the signal charge. The electric field in the CMCCD channel is then given by:

$$\begin{aligned} E(x, y) &= -\nabla V(x, y) \\ &= -\nabla \psi(x, y) - \frac{qN_d}{2\epsilon_s} \nabla (x_t - n_s(x, y)/N_d)^2 \\ &= -\nabla \psi(x, y) - \frac{q}{\epsilon_s} (x_t - n_s(x, y)/N_d) \nabla n_s(x, y) \end{aligned} \quad (\text{A.2})$$

Note that $-\nabla \psi(x, y)$ is the fringing field. Eq. (A.2) can be rewritten as:

$$E(x, y) = E_f(x, y) + \frac{q}{\epsilon_s} \left(\frac{n_s(x, y)}{N_d} - x_t \right) \nabla n_s(x, y) \quad (\text{A.3})$$

Appendix B

Equivalent Circuit Model of CMCCD in SPICE3e2

The following is the equivalent circuit model for a unit cell of the CMCCD used in SPICE3e2 simulation. The model is actually defined in SPICE3e2 as a subcircuit, and one pixel of the CMCCD is formed by a number of such unit cells.

```
***** Equivalent-circuit Model of CMCCD for a unit cell *****  
.LIB ccdcell  
.SUBCIRCUIT ccdcell 1 2 201 202 K1  
*Description of the Distributed C  
V1 1 101 0  
B1a 101 0 V=V(201)*(1-1E14*V(201))/(0.579675E-15)  
B1b 0 201 I=I(V1)  
C1 201 0 1  
*Description of the Current Sources  
*B1bb K1 0 V=0.33E4*(V(G1)-V(G4))*0.90  
B1x J1 0 V=4.77e7*exp(-1*sqrt(V(K1)^2+1e-40)/1644)
```

```
+          +3.24e7/(1+(sqrt(V(K1)^2+1e-30)/130.5)^0.32)
B1u L1 0 V=5000/sqrt(1+(5000*V(K1)/V(J1))^2)
B1z M1 0 V=129.5
+          +812*exp(-1*(ln(sqrt(V(K1)^2+1e-30)/3394.8)/ln(1.82))^2)
B1c 1 2 I=4e4*V(L1)*V(K1)*V(201)
B1d 1 2 I=16e8*V(L1)*V(101)*(V(201)-V(202))
B1e 1 2 I=16e8*V(M1)*(V(201)-V(202))
.ENDS ccdcell
.ENDL ccdcell
```

Appendix C

Fletcher-Reeves Method

The Fletcher-Reeves method is a conjugate-direction-approach-based optimization algorithm [57] [58]. Its main feature is that constants α_k for $k = 0, 1, 2, \dots$ are determined by minimizing $f(\mathbf{x} + \alpha \mathbf{d}_k)$ with respect to α using a line search, while the line search direction \mathbf{d}_k is a conjugate direction with respect to $\mathbf{d}_{k-1}, \mathbf{d}_{k-2}, \dots, \mathbf{d}_0$. The Fletcher-Reeves method is more amenable to the minimization of nonquadratic problems since a larger reduction can be achieved in calculating $f(\mathbf{x})$ along \mathbf{d}_k at points outside the neighbourhood of the solution. In addition, this method obviates the calculation of the Hessian matrix $\mathbf{H}(x)$, the second-order partial derivatives of $f(\mathbf{x})$.

Before describing the Fletcher-Reeves method, we define the following. If $f(\mathbf{x}) \in C^1$ where $\mathbf{x} = [x_1 x_2 \dots x_n]^T$, the gradient vector of $f(\mathbf{x})$ is defined as

$$g(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_n} \right]^T = \nabla f(\mathbf{x}) \quad (\text{C.1})$$

If there exists a finite set of distinct nonzero vector $\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_k\}$ such that

$$\mathbf{d}_i^T \mathbf{H} \mathbf{d}_i = 0 \quad (\text{C.2})$$

for all $i \neq j$, where \mathbf{H} is a real symmetric matrix, the vector $\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_k\}$ is said

to be conjugate with respect to \mathbf{H} , and \mathbf{d}_k is said to be a conjugate direction with respect to $\mathbf{d}_{k-1}, \mathbf{d}_{k-2}, \dots, \mathbf{d}_0$.

An implementation of the Fletcher-Reeves algorithm is described as the following.

Step 1 Input \mathbf{x}_0 and ε .

Step 2 Set $k = 0$.

Compute \mathbf{g}_0 and set $\mathbf{d}_0 = -\mathbf{g}_0$.

Step 3 Find α_k , the value of α that minimizes $f(\mathbf{x}_k + \alpha\mathbf{d}_k)$.

Set $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{d}_k$.

Step 4 If $|\mathbf{x}_{k+1} - \mathbf{x}_k| = |\alpha_k\mathbf{d}_k| < \varepsilon$, then do:

output $\mathbf{x}^* = \mathbf{x}_{k+1}$, $f(\mathbf{x}^*) = f(\mathbf{x}_{k+1})$;

stop.

Step 5 If $k = n - 1$, set $\mathbf{x}_0 = \mathbf{x}_{k+1}$, and go to Step 2.

Step 6 Compute \mathbf{g}_{k+1} .

Compute

$$\beta_k = \frac{\mathbf{g}_{k+1}^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{g}_k}$$

Set $\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_k\mathbf{d}_k$.

$k = k + 1$, and repeat from Step 3.

Appendix D

Noise Analysis of CMCCD

Noise in a GaAs CMCCD can be separated into four components: (1) thermal generation noise (or shot noise, dark current noise), (2) bulk trapping noise (or fast interface state noise), (3) input noise and (4) output noise (or *KTC* noise). The first two types of noise are associated with charge transfer and the last two are associated with charge injection and detection.

D.1 Thermal Generation Noise

The thermal generation noise, n_{th} , is due to the thermally-generated leakage current and is given by the square root of the number of thermally-generated electrons in the charge packet as it is transferred down to the CMCCD channel. An expression of n_{th} is given by [25]

$$n_{th} = \sqrt{(J_d A_e N_T)/(q f_c)} \quad (\text{D.1})$$

where J_d is the leakage current density, A_e is the area of the transport electrode, N_T is the number of transfers in the CMCCD and f_c is the clock frequency. For the two-phase and uni-phase CMCCDs, assuming $f_c = 100$ MHz, $N_T = 128 \times 2$, $N_d =$

$5 \times 10^{16} / \text{cm}^3$ and therefore $J_d = 1.5 \times 10^{-7} \text{ A/cm}^2$ [19], the thermal generation noise can be calculated by Eq. (D.1) and are 31 and 32 electrons, respectively.

D.2 Bulk Trapping Noise

Bulk traps are the dominant sources in fast interface state noise for GaAs CMCCD since the signal charge packets are transferred within the bulk of the CCD channel. The bulk trapping noise is given by [64]:

$$n_{trap} = \sqrt{N_T V_{SIG} N_t \exp(-T_t/\tau_e) [1 - \exp(-T_t/\tau_e)]} \quad (\text{D.2})$$

where V_{SIG} is the volume that the charge packet occupies, N_t is the number of bulk states per unit volume and typically $10^{11} / \text{cm}^3$, T_t is the transfer time, and τ_e is the electron emission time constant. It can be seen from Eq. (D.2) that the bulk trapping noise decreases as the size of the charge packet decreases because the volume that a charge packet occupies is a strong function of the size of the charge packet. Therefore, the bulk trapping noise does not impose a serious limitation to the signal-to-noise ratio of the CMCCD. The factor in Eq. (D.2) has a maximum value of 1/2 when $T_t = \tau_e \ln 2$ [64]. Assuming the size of the charge packet is 10^6 electrons, then $V_{SIG} \approx 10^6 / N_d = 0.5 \times 10^{-10} \text{ cm}^3$. Thus the bulk trapping noise for the two-phase and uni-phase CMCCDs is approximately 36 and 51 electrons, respectively.

D.3 Input and Output Noise

The charge injection and detection in a CMCCD frequently employ some means of charging a capacitor to a certain level. When a capacitor is charged through a resistor, the voltage across the capacitor contains some noise due to the fluctuation of the voltage across the resistor. The noise related to the charge injection and detection

is given by [64]:

$$n_{input} = \frac{1}{q} \sqrt{KTC_{inp}} \quad (D.3)$$

$$n_{output} = \frac{1}{q} \sqrt{KTC_{op}} \quad (D.4)$$

where K is Boltzmann constant, T is temperature, C_{inp} is the depletion capacitance of the input gate G_1 in Fig. D.1, and C_{op} is the sum of the capacitance associated with the output ohmic contact and the gate parasitic capacitance of the output buffer, as shown in Fig. D.1. Since the input and output noise is only proportional to the square root of KTC , they are also called KTC noise. At room temperature, Eqs. (D.3) and

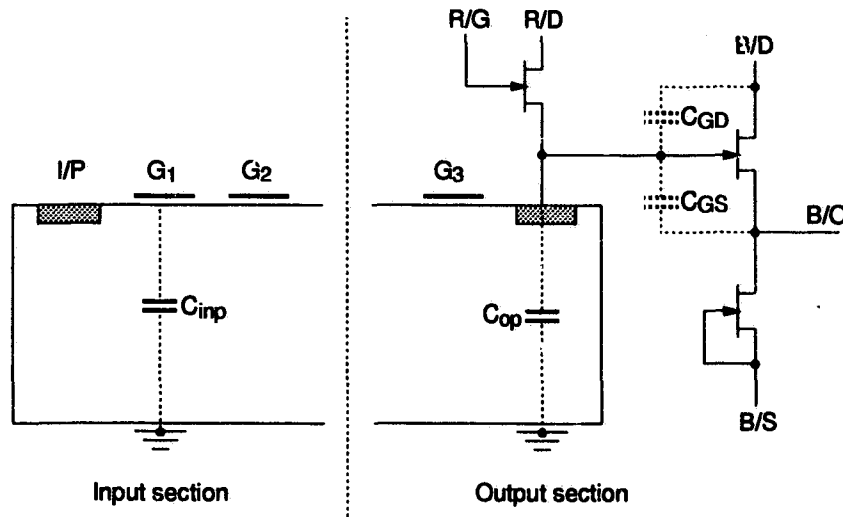


Figure D.1: Input and output sections of the CCD.

(D.4) become:

$$n_{input} = 400 \sqrt{C_{inp} (pF)} \quad (D.5)$$

$$n_{output} = 400 \sqrt{C_{op} (pF)} \quad (D.6)$$

The depletion capacitance C_{inp} can be calculated approximately by:

$$C_{inp} \approx \frac{\epsilon_s}{x_t} A_{G1} \quad (D.7)$$

where x_t is the thickness of the active layer and A_{G_1} is the area of G_1 . And C_{op} is given by:

$$C_{op} = C_{GD} + (1 - G_v)C_{GS} \quad (D.8)$$

where G_v is the voltage gain of the buffer. C_{GD} and C_{GS} are approximately 1.5 fF per micron of channel width for the depletion-mode GaAs process. Assuming the area of G_1 is $500 \mu\text{m}^2$, $x_t = 0.3 \mu\text{m}$, the width of the MESFETs is $50 \mu\text{m}$, then $C_{inp} \approx 0.2 \text{ pF}$ and $C_{op} \approx 0.1 \text{ pF}$. Using the values of C_{inp} and C_{op} , we obtain the input noise $n_{input} = 179$ electrons, and the output noise $n_{output} = 127$ electrons for both the two-phase and the uni-phase CMCCD.

D.4 Signal-to-Noise Ratio

The Signal-to-noise ratio (S/N) of the CMCCD, expressed in terms of signal level within the devices, is determined at the upper end by the maximum charge handling capability and at the lower end by the various noise sources present in the CMCCD channel. For both the two-phase and the uni-phase CMCCDs, the maximum charge packet size is given by [66]:

$$Q_{max} = C_e \psi_{BH} \quad (D.9)$$

where C_e is the depletion capacitance associated with the area of that part of the electrode under which the charge is stored, and ψ_{BH} is the depth of the potential well (or the barrier height). Using the parameters obtained in Chapter 3, we have the maximum charge handling capability of 8.6×10^5 electrons for the two-phase CMCCD and 6.0×10^5 electrons for the uni-phase CMCCD. If we define the S/N of the CMCCD as $20 \log(Q_{max}/Q_{noise})$, then they are 67 dB and 63 dB for the two-phase and the uni-phase CMCCDs, respectively. Table D.1 lists the various noise contributions as calculated above, the total noise, and the S/N for the two-phase and

the uni-phase CMCCDs.

Noise	Two-phase CMCCD	Uni-phase CMCCD
Thermal generation	$31 e^-$	$32 e^-$
Bulk trapping	$36 e^-$	$51 e^-$
Input	$179 e^-$	$179 e^-$
Output	$127 e^-$	$127 e^-$
Total noise	$373 e^-$	$389 e^-$
Max. charge packet size	$8.6 \times 10^5 e^-$	$6.0 \times 10^5 e^-$
S/N ratio	67 dB	63 dB

Table D.1: Noise and S/N of the two-phase and the uni-phase CMCCDs

Appendix E

Stability Analysis of the Activation Circuit

The activation circuit has an operational amplifier configuration. An amplifier is required to be unconditionally stable in almost all of the applications. Therefore, an amplifier should have a one-pole open-loop transfer function. In practical amplifier, however, the ideal one-pole transfer function is not achievable. It is, therefore, necessary to compensate the frequency performance of the amplifier to avoid oscillations. In our design, a Miller compensation capacitor C_c was connected across the gain stage.

The effect of the input stage is to function as a transconductance amplifier with an output current given by $g_{mI}v_{in}$. If the effect of the output source follower is ignored, the activation circuit configured by the two-stage operational amplifier with a Miller compensation capacitor can be modeled by the equivalent circuit [74] as shown in Fig. E.1. The overall transfer function for the equivalent circuit is given by:

$$\frac{v_{out}}{v_{in}} = \frac{g_{mI}g_{mII}r_{I}r_{II}(1 - sC_c/g_{mII})}{1 + sA + s^2B} \quad (\text{E.1})$$

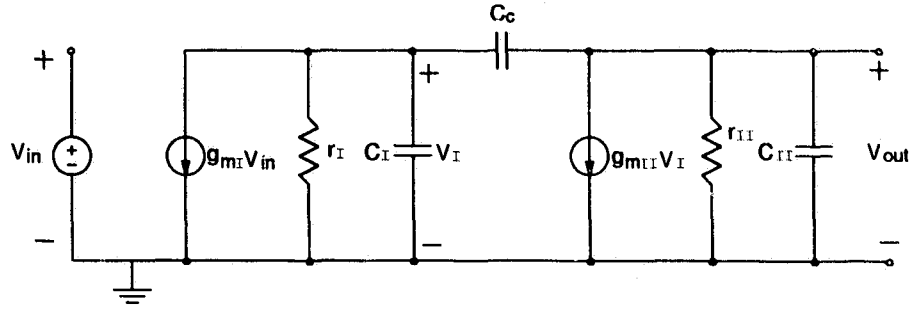


Figure E.1: Small-signal equivalent circuit for a two-stage operational amplifier.

where

$$A = r_I(C_I + C_c) + r_{II}(C_{II} + C_c) + g_{mII}r_I r_{II}C_c,$$

$$B = r_I r_{II}[C_I C_{II} + C_c(C_I + C_{II})]$$

Neglecting the finite value of the gain at low frequencies, the voltage gain in the mid- and high-frequency ranges can be simplified and is given by:

$$\left| \frac{v_{out}(j\omega)}{v_{in}(j\omega)} \right| \approx \frac{g_{mI}}{\omega C_c} \quad (\text{E.2})$$

Eq. (E.2) exhibits a 6 dB/octave rolloff and reaches a unity-gain at $\omega_1 = g_{mI}/C_c$.

The value of C_c is obtained

$$C_c \approx \frac{g_{mI}}{2\pi f_1} \quad (\text{E.3})$$

If g_{mI} and f_1 are assumed to be 4 mS and 1 GHz, respectively, C_c is approximately 0.6 pF.

Appendix F

NT/BNR 0.8 μm GaAs Models

Copying is not allowed by Canadian Microelectronics Corporation and Northern Telecom/Bell Northern Research.

Appendix G

NT/BNR 0.8 μm GaAs Process

Copying is not allowed by Canadian Microelectronics Corporation and Northern Telecom/Bell Northern Research.

VITA

Surname: CHEN

Given Names: LIDONG

Place of Birth: Hefei, China

Date of Birth: Dec. 7, 1961

Educational Institutions Attended:

Anhui University	1979 — 1983
Anhui University	1983 — 1986
University of Victoria	1992 — 1994

Degrees Awarded:

B. Sc.	Anhui University	1983
M. Sc.	Anhui University	1986

Honours and Awards:

The Royal Fellowship Award, The Royal Society, UK.	1991 — 1992
Teaching Assistant Fellowship Awards, Univ. of Victoria,	1992 — 1994

Publications:

Journal Paper

L. Chen and L. Zhang, "Channel ordering and routing for VLSI building blocks layout," *ACTA ELECTRONICS SINICA*, vol. 16, No. 1, pp. 53 — 58, 1988.

L. Chen, S. Pennathur and H. H. L. Kwok, "Effects of the clock voltage waveforms on the charge transfer inefficiency of high-speed charge-coupled devices," *Solid-State Electronics*, vol. 36, No. 5, pp. 798 — 800, 1993.

G. Deliyannides, L. Chen and H. H. L. Kwok, "Design and analysis of a high-speed GaAs comparator circuits," *Int. Journal of Electronics*, vol. 76, No. 2, pp. 241 – 248, 1994.

L. Chen and H. H. L. Kwok, "An optimized design for 2-phase GaAs CMCCD and its charge transfer simulation," *Solid-State Electronics*, vol. 37, No. 11, pp.1863 – 1869, 1994.

L. Chen and H. H. L. Kwok, "A comparison study of the uni-phase and the two-phase GaAs CMCCDs" submitted to *Solid-State Electronics*,

L. Chen and H. H. L. Kwok, "A hybrid architecture for Hopfield neural network and its GaAs implementation using CCDs as storage elements," submitted to *IEEE Trans. VLSI*.

Conference Paper

L. Chen and B. Makin, "Automatic layout generation for digital BiCMOS cells," in *Proc. of IEEE ISCAS*, San Diego, pp. 2569 – 2572, May, 1992.

L. Chen and H. H. L. Kwok, "A GaAs CCD structure for multiplying digital-to-analog conversion," in *Proc. of CCECE*, Vancouver, pp. 811 – 813, Sept., 1993.

L. Chen and H. H. L. Kwok, "An optimized structure of 2-phase GaAs CMCCD," in *Proc. of ISDRS*, Charlottesville, pp. 519 – 522, Dec., 1993.

PARTIAL COPYRIGHT LICENSE

I hereby grant the right to lend my dissertation to users of the University of Victoria Library, and to make single copies only for such users or in response to a request from the Library of any other university, or similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this dissertation for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying of this dissertation for financial gain shall not be allowed without my written permission.

Title of Dissertation:

Design and Implementation of GaAs CCD/MESFET ICs for Artificial Neural Network Application

Author: _____

LIDONG CHEN

January 21, 1995