UNIVERSIDADE FEDERAL DE ITAJUBÁ PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Implementação de Neurônio Artificial em Tecnologia CMOS IBM 130 nm utilizando o Modelo de Izhikevich em Topologia Translinear Dinâmica com Transistores Halo-Implantados.

Odilon de Oliveira Dutra

Itajubá, Junho de 2015

UNIVERSIDADE FEDERAL DE ITAJUBÁ PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Odilon de Oliveira Dutra

Implementação de Neurônio Artificial em Tecnologia CMOS IBM 130 nm utilizando o Modelo de Izhikevich em Topologia Translinear Dinâmica com Transistores Halo-Implantados.

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica como parte dos requisitos para obtenção do Título de Doutor em Ciências em Engenharia Elétrica.

Área de Concentração: Microeletrônica

Orientador: Luís Henrique de Carvalho Ferreira Coorientador: Tales Cleber Pimenta

> Junho de 2015 Itajubá - MG

UNIVERSIDADE FEDERAL DE ITAJUBÁ PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Implementação de Neurônio Artificial em Tecnologia CMOS IBM 130 nm utilizando o Modelo de Izhikevich em Topologia Translinear Dinâmica com Transistores Halo-Implantados.

Odilon de Oliveira Dutra

Tese aprovada por banca examinadora em 05 de Junho de 2015, conferindo ao autor o título de **Doutor em Ciências em Engenharia Elétrica**.

Banca Examinadora:

- Prof. Dr. Luís Henrique de Carvalho Ferreira
- Prof. Dr. Tales Cléber Pimenta
- Prof. Dr. Samuel Euzédice de Lucena
- Prof. Dr. José Feliciano Adami
- Prof. Dr. Leonardo Breseghello Zoccal
- Prof. Dr. Robson Luiz Moreno
- Prof. Dr. Paulo César Crepaldi

Itajubá 2015 "Conhecimento real é saber a extensão da própria ignorância.

"(Confúcio).

Agradeço e dedico este trabalho primeiramente aos meus pais, Nélia e Odilon, por me apoiarem incondicionalmente durante minha vida.

Gostaria ainda de homenagear o meu pai, que mesmo após falecido sei que continua me apoiando e a cada dia me dando forças para continuar sempre em busca de conhecimento.

Também devo agradecer ao restante dos meus familiares e minha querida esposa, pela dedicação e paciência confiadas a mim durante todo este processo.

Em seguida, agradeço ao meu orientador, Prof. Luís Henrique de Carvalho Ferreira, e Coorientador, Tales Cleber Pimenta, pela paciência, confiança e crédito em mim depositado.

Agradeço ao MEC que através do projeto Reuni viabilizou financeiramente este trabalho.

Agradeço ainda à CAPES, à FAPEMIG e ao CNPq.

Agradeço também a MOSIS que através de seu Programa Educacional MOSIS (MEP) viabilizou a difusão do amplificador descrito neste trabalho em pastilha de Silício.

E a quem mais ler este trabalho e sentir que contribuiu para o seu desenvolvimento, meus mais sinceros agradecimentos.

Resumo

Este trabalho descreve a implementação do modelo neural de Izhikevich em circuitos dinâmicos translineares - DTL - através da utilização de dispositivos halo-implantados em tecnologia 130 nm. Um desenho em forma matricial de ordem $m \ge n$ é capaz de, não apenas, substancialmente aumentar a impedância de saída de tais transistores haloimplantados devido à diminuição do efeito LDIBL, mas também melhorar o descasamento e a degradação da transcondutância, tornando-os elementos translineares aplicáveis a circuitos de ultra-baixa-tensão e ultra-baixa-potência. O neurônio proposto foi simulado com sucesso em IBM CMOS 130 nm. O mesmo é capaz de gerar os 20 padrões previstos pelo modelo para neurônios tálamo-corticais como outros trabalhos com implementações diferentes, mas melhorando diversos aspectos como a utilização de baixíssima tensão de alimentação de 250 mV. Mesmo não tendo sido possível a medição dos diversos padrões, a medição de diversos espelhos de corrente implementados no chip para auxílio da polarização do circuito, bem como a caracterização das matrizes, medição do comportamento estático de um filtro passa-baixas DTL que implementa a acomodação do neurônio e também a comparação de um padrão neural fictício com sua simulação demonstraram funcionamentos muito similares aos obtidos nas simulações, indicando que a topologia adotada é uma boa opção para a implementação de circuitos DTL em ultra-baixa-tensão e ultra-baixa-potência.

Palavras-chaves: Neurônio. Izhikevich. Transistores Halo-Implantados. Circuitos DTL. Implante Halo. Ultra-Baixa-Tensão. Ultra-Baixa-Potência.

Abstract

This work describes the current mode implementation of Izhikevich neuron model implemented using halo-implanted devices 130 nm structured within matrices of order $m \ge n$ capable of substantially increasing output impedance of such devices due to the decreasing on LDIBL effect while also improving mismatch and decreasing transconductance degradation, turning such devices applicable into dynamical translinear loops (DTL) circuits operating at ultra-low-voltage and ultra-low-power. The proposed neuron was successfully simulated in 130 nm IBM CMOS process as the DTL circuit topology adopted generates the 20 patterns defined in Izhikevich model as other similar works while improving several aspects as the low supplied voltage used 250 mV. Although neuron patterns could not be measured yet due to test bench issues, current mirrors implemented on chip for circuit biasing purposes, as well as the matrices characterization, measuring of a DTL low-pass filter static behavior and also a comparison between a fictitious neuron pattern and its simulation have shown good matchings against the simulations making us believe the adopted topology is a nice choice for the implementation of DTL circuits in ultra-low-voltage and ultra-low-power.

Key-words: Neuron. Izhikevich Implementation. Halo-Implanted Devices. DTL. Halo Implant. Ultra-Low-Voltage. Ultra-Low-Power.

Lista de ilustrações

Figura 1 –	Neurônio Biológico	20
Figura 2 –	Padrões de Disparos e Rajadas do Modelo de Izhikevich (1)	24
Figura 3 –	Circuito Elétrico Equivalente ao Modelo de Hodgkin-Huxley	25
Figura 4 –	Parâmetros do Modelo de Izhikevich e suas Atuações nos Potenciais de	
	Membrana (v) e Acomodação (u)	28
Figura 5 –	Padrões Excitatórios - Neurônios Corticais.	29
(a)	RS	29
(b)	IB	29
(c)	СН	29
Figura 6 –	Padrões Inibitórios - Neurônios Corticais.	29
(a)	FS	29
(b)) LTS	29
Figura 7 $-$	Princípio DTL em Transistores MOS Operando em Subthreshold e Sa-	
	turado.	31
Figura 8 –	Transistor nMOS Halo Implantado	33
(a)	Corte Transversal.	33
(b)	ISE-TCAD (Sem Gate para Visualização dos Implantes Halo) (2)	33
Figura 9 $-$	Efeitos de Canal Curto e Canal Estreito em Dispositivos Uniforme-	
	mente Dopados na Tensão <i>Threshold</i> (3)	36
Figura 10 –	- Matrizes de Transistores Halo-Implantados.	39
(a)	Matriz nMOS.	39
(b)	Matriz pMOS.	39
Figura 11 –	- Espelhos de Corrente Utilizados para Projeto das Matrizes	40
(a)	Espelho de Corrente	40
(b)	Espelhamento de Corrente Melhorado com Uso de Matrizes	40
Figura 12 –	Coeficientes Angulares das Regressões Lineares dos Espelhamentos	41
(a)	Coeficiente Angular - M_2	41
(b)	Coeficiente Angular - M_3	41
(c)	Coeficiente Angular - M_4	41
Figura 13 –	Coeficientes Lineares das Regressões Lineares dos Espelhamentos	42
(a)	Coeficiente Linear - M_2	42
(b)	Coeficiente Linear - M_3	42
(c)	Coeficiente Linear - M_4	42
Figura 14 -	$-R^2$ das Regressões Lineares dos Espelhamentos	42
(a)	R^2 - M_2	42
(b)) R^2 - M_3	42

(c) $R^2 - M_4$	42
Figura 15 – Coeficientes Angulares das Regressões Lineares dos Espelhamentos -	
Halo Implantados de Tamanho Equivalente à Matriz	43
(a) Coeficiente Angular - M_2	43
(b) Coeficiente Angular - M_3	43
(c) Coeficiente Angular - M_4	43
Figura 16 – Coeficientes Lineares das Regressões Lineares dos Espelhamentos -	
Halo Implantados de Tamanho Equivalente à Matriz	43
(a) Coeficiente Linear - M_2	43
(b) Coeficiente Linear - M_3	43
(c) Coeficiente Linear - M_4	43
Figura 17 – R^2 das Regressões Lineares dos Espelhamentos - Halo Implantados de	
Tamanho Equivalente à Matriz	43
(a) $R^2 - M_2 \dots \dots$	43
(b) $R^2 - M_3$	43
(c) $R^2 - M_4$	43
Figura 18 – Comparação entre Funções de Transferência - p MOS - Matriz x Transferência - p MOS	
sistor Halo de Tamanho Equivalente.	44
Figura 19 – Comparação entre Funções de Transferência n MOS - Matriz x Transis-	
tor Halo de Tamanho Equivalente à Matriz	45
Figura 20 – Comparação entre Transcondu tâncias - p MOS - Matriz $\mathbf x$ Transistor	
Halo de Tamanho Equivalente.	46
Figura 21 – Comparação entre Transcondu tâncias n ${\rm MOS}$ - Matriz x Transistor Halo	
de Tamanho Equivalente à Matriz.	46
Figura 22 – Comparação entre Funções de Saída - p MOS - Matriz $\mathbf x$ Transistor	
Halo-Implantado de Tamanho Equivalente	47
Figura 23 – Comparação entre Funções de Saída n MOS - Matriz x Transistor Halo-	
Implantado de Tamanho Equivalente à Matriz.	48
Figura 24 – Comparação entre Impedâncias - pMOS - Matriz x Transistor Halo-	
Implantado de Tamanho Equivalente	48
Figura 25 – Comparação entre Impedâncias - n MOS - Matriz $\mathbf x$ Transistor Halo-	
Implantado de Tamanho Equivalente	49
Figura 26 – Neurônio Desenvolvido. \ldots \ldots \ldots \ldots \ldots \ldots \ldots	50
Figura 27 – Acomodação.	51
Figura 28 – Comportamento Acomodado do Circuito Dinâmico Translinear. \ldots .	54
(a) Matriz - I_u Acomodado	54
(b) Transistor de Tamanho Equivalente à Matriz - I_u Acomodado	54
Figura 29 – Membrana.	55
Figura 30 – Comparador de Corrente com Histerese.	57

Figura 31 $-$	Função de Transferência do Comparador de Corrente com Histerese	58
Figura 32 –	Variação Estatística 1000 Rodadas - Limite Mínimo Histeres e	59
Figura 33 –	Variação Estatística 1000 Rodadas - Limite Máximo Histerese	59
Figura 34 –	Circuito de Reset do Integrador da Membrana.	60
Figura 35 –	Padrão RS - Simulação SPICE vs. Modelo	65
Figura 36 –	Padrão CH - Simulação SPICE vs. Modelo.	65
Figura 37 –	Padrão IB - Simulação SPICE vs. Modelo	66
Figura 38 –	Padrão FS - Simulação SPICE vs. Modelo	66
Figura 39 –	Padrão LTS - Simulação SPICE vs. Modelo	67
Figura 40 –	Padrão PS - Simulação SPICE vs. Modelo	67
Figura 41 –	Análise de Monte Carlo - Tempo entre Potenciais de Ação	69
(a)	RS	69
(b)	FS	69
(c)	СН	69
(d)	IB	69
(e)	LTS	69
Figura 42 $-$	Calibração da Capacitância da SMU	70
Figura 43 –	Neurônio Implementado.	83
(a)	Neurônio - Layout.	83
(b)	Neurônio - Foto.	83
Figura 44 –	Modelo de Neurônio Artificial de Izhikevich Implementado com Matri-	
	zes de Transistores Halo-Implantados.	84
Figura 45 $-$	Circuito de Acomodação Isolado (Filtro Passa-Baixas DTL) Implemen-	
	tado com Matrizes de Transistores Halo-Implantados	85
Figura 46 –	Esquemático de Ligação para teste de Bancada - Circuito Neurônio	87
Figura 47 $-$	Buffer Utilizado para Geração de Polarizações dos Espelhos por Tensão	
	- Circuito Neurônio.	88
Figura 48 –	Ensaio para Medições do Neurônio	88
Figura 49 –	Esquemático de Ligação para teste de Bancada - Circuito Neurônio	90
Figura 50 –	Circuito de Reset da Acomodação Preciso.	91

Lista de tabelas

Tabela 1 –	- Conversão de Parâmetros - Do Modelo Izhikevich para Circuito DTL	63
Tabela 2 –	- Resultados e Comparação	71
Tabela 3 –	- Conversão de Parâmetros - Do Modelo Izhikevich para Circuito DTL.	92

Lista de abreviaturas e siglas

DIBL	Barreira Induzida do Dreno (Drain Induced Barrier Lowering)
СН	Chattering
DITS	Deslocamento da Tensão de Limiar (Drain Induced Threshold Shift)
DTL	Circuitos Dinâmicos Translineares (Dynamic Translinear Loops)
FS	Disparo Rápido (Fast Spiking)
IB	Rajada Intrínseca (Intrinsically Bursting)
LDIBL	Diminuição da Barreira Induzida do Dreno em Dispositivos de Canal Longo (Long-Channel Drain Induced Barrier Lowering)
LTS	Disparo Sub-Limiar (Low Threshold Spiking)
RS	Disparo Regular (Regular Spiking)
HRSMU	Unidade de Fonte e Medição de Alta Resolução <i>(High Resolution Source Meter Unit)</i>
STC	Código de Espaçamento de Tempo (Space Time Code)

Lista de símbolos

D_n	Coeficiente de difusão
I_{1u}	Corrente do circuito de acomodação oriunda da conversão do parâmetro a e b do modelo de Izhikevich
I_{2u}	Corrente do circuito de acomodação oriunda da conversão do parâmetro a e b do modelo de Izhikevich
I_{1v}	Corrente do circuito da membrana oriunda da conversão do parâmetro a e b do modelo de Izhikevich
I_{2v}	Corrente do circuito de acomodação oriunda da conversão do parâmetro a e b do modelo de Izhikevich
I_c	Corrente correspondente ao Parâmetro \boldsymbol{c} do modelo de Izhikevich
I_{DC}	Corrente de Offset DC do modelo de Izhikevich
I_D	Corrente de Dreno do Transistor
I_d	Corrente correspondente ao Parâmetro d do modelo de Izhikevich
I_S	Corrente Característica Reversa
\dot{I}_u	Corrente que Representa o Potencial de Acomodação em Modo Corrente
I_u	Corrente que Representa a Derivada do Potencial de Acomodação em Modo Corrente
\dot{I}_v	Corrente que Representa a Derivada do Potencial de Membrana em Modo Corrente
I_v	Corrente que Representa o Potencial de Membrana em Modo Corrente
k	Constante de Boltzmann = $1,38 \times 10^{-23} [J/K]$
L	Comprimento de Canal
n	Parâmetro dependente da tecnologia utilizada. Equivale a 1,16 para IBM 130nm

 n_e Concentração de elétrons

n_{io}	Concentração de portadores intrínseca do Silício
q	Carga Fundamental do Elétron = 1,60218 $\times 10^{-19}$ coulomb
T	Temperatura Absoluta
\dot{u}	Derivada do Potencial de Acomodação
u	Potencial de Acomodação
U_T	Tensão Térmica kT/q
\dot{v}	Derivada do Potencial de Membrana
v	Potencial de Membrana
V_{DS}	Tensão entre Dreno e Fonte
V_{Th}	Tensão de Limiar
W	Largura de Canal
μ_n	Mobilidade dos portadores
φ_{FN}	Potencial de Fermi do elétron

Sumário

1	Introdução	16
1.1	Motivação	16
1.2	Objetivo	17
1.3	Conteúdo	17
2	Revisão da Literatura	19
2.1	O Neurônio	19
2.1.1	Funcionamento do Neurônio	20
2.1.2	Neurônios Tálamo-Corticais	21
2.1.3	Sinapses	22
2.1.3.1	Sinapses Químicas	22
2.1.3.2	Sinapses Elétricas	22
2.1.4	Padrões Neurais	23
2.2	Modelos	23
2.2.1	Modelo de Hodgkin-Huxley	24
2.2.2	Modelo de Izhikevich	26
2.2.2.1	Categorização das Sequências de Potencias de Ação	28
2.3	Circuitos DTL	29
2.4	Implante Halo	31
2.4.1	Construção	32
2.4.2	Modelos de Corrente de Dreno	32
2.4.3	Efeitos dos Implantes Halo	35
3	Proposta	38
3.1	Matrizes de Transistores Halo-Implantados	38
3.1.1	Implementação das Matrizes	40
3.1.2	Curvas Características	44
4	Aplicação	50
4.1	Acomodação	50
4.1.1	Circuito de Acomodação - Análise como Filtro Translinear Passa-Baixa	52
4.2	Membrana	54
4.3	Circuito de Reset	56
4.3.1	Comparador com Histerese	56
4.3.2	Reset do Integrador da Membrana	59
4.3.3	Reset do Integrador da Acomodação	61
4.3.4	Cálculo dos Parâmetros para Demais Padrões	63

4.4	Simulações e Medições	64
4.4.1	Padrões de Disparo de Potenciais de Ação	64
4.4.2	Análise Estatística	68
4.4.3	Medidas dos Comportamentos de Disparo e Rajadas de Potenciais de	
	Ação dos Neurônios	68
4.5	Compilação de Figuras de Mérito	71
5	Conclusão	73
5.1	Trabalhos Futuros	74

Referências					•						•	•		•			•			•		•	•	•	•					•			7	76
-------------	--	--	--	--	---	--	--	--	--	--	---	---	--	---	--	--	---	--	--	---	--	---	---	---	---	--	--	--	--	---	--	--	---	----

		•	79
 			80
 			80
 			82
 			86
 			86
 		•	89
 • •		•	91
 			93
 			93
 			95
 			96
 			98
 		. 1	100
 · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·		· · · · · · · · · · · · · · · · · · ·

1 Introdução

1.1 Motivação

Nos dias de hoje há um crescente interesse no entendimento dos cérebros dos mamíferos como consequência dos grandes avanços da neurobiologia. Governos de todo mundo, em especial o governo americano e a União Europeia, têm investido uma grande quantidade de recursos em projetos que visam desenvolver *clusters* computacionais capazes de simular modelos matemáticos criados para representar o comportamento dinâmico de neurônios.

No entanto, modelos matemáticos neurais demandam um poder computacional muito grande, inviabilizando simulações de redes neurais baseadas em modelos realmente representativos. Tais simulações podem durar dezenas de dias para representar apenas 1 segundo do comportamento neural biológico mesmo nos mais avançados *clusters* do mundo.

Dessa forma, muitos trabalhos têm sido publicados na direção de circuitos analógicos totalmente implementados em circuito integrado (4, 5, 6) capazes de emular neurônios através de seus modelos matemáticos de maneira analógica.

Contudo, a abordagem analógica traz consigo muitos outros problemas. Muitos modelos matemáticos como o integra e dispara oferecem uma descrição biológica do comportamento dinâmico do neurônio muito pobre e, em contrapartida, podem ser facilmente implementados em circuito. Outros modelos, como por exemplo o modelo de Hodgin-Huxley (7), modelo de Fitzhugh-Nagumo (8) e modelo de Mihalas-Niebur (9), dentre outros, oferecem avançada descrição biológica dos neurônios com o custo de circuitos complicadíssimos de se implementar em silício.

O modelo de Izhikevich (1) representa um equilíbrio entre estas duas vertentes. O mesmo apresenta uma descrição completa do comportamento dinâmico dos neurônios corticais e tálamo-corticais, capaz de emular os 20 padrões descritos na literatura, de uma forma possível de se implementar em circuitos integrados em tecnologia MOS polarizados abaixo da tensão de limiar implementados em topologia de circuitos translineares.

Entretanto, é comum encontrar na literatura citações à não aplicabilidade de dispositivos halo-implantados em circuitos analógicos (3) devido, principalmente, à degradação da transcondutância (10) e impedância de saída dos mesmos (11, 12), bem como a maior variação estatística que tais dispositivos apresentam (13), devido principalmente ao efeito de diminuição da barreira induzida de dreno em dispositivos de canal longo (*Long Channel Drain Induced Barrier Lowering* - LDIBL). Desta forma, esta tese vem desenvolver um trabalho que, através de uma técnica de implementação de circuitos integrados com transistores halo-implantados supere os problemas acima relatados comprovando a utilidade de tais transistores como elementos translineares utilizáveis em aplicações biomédicas.

1.2 Objetivo

O objetivo desta tese é comprovar a utilidade de transistores halo-implantados organizados matricialmente (11, 12) como elementos translineares. Espera-se atingir este objetivo através de uma análise no circuito correspondente à variável de acomodação do modelo, desenvolvida como um simples filtro passa-baixas em topologia translinear e implementada em silício, que permitirá uma análise mais simples do desempenho dos transistores halo-implantados utilizados como componentes translineares, isolando-se outros fatores que possam comprometer o desempenho do circuito. Esta análise e fatores serão discutidos no decorrer desta tese no Capítulo 4.

Será implementado, também no Capítulo 4, um circuito equivalente a um neurônio cortical descrito no modelo de Izhikevich (1) que corroborará com a análise da eficiência das matrizes de transistores halo-implantados como elementos translineares. Através de medições e simulações em BSIM4v3, serão comparados os diferentes padrões de disparos e rajadas de potenciais de ação previstos pelo modelo e gerados pelo circuito desenvolvido.

1.3 Conteúdo

Esta tese divide-se em cinco capítulos mais apêndices. Cada capítulo se divide em seções para melhor organização do conteúdo de forma a se obter um perfeito entendimento do trabalho realizado. O Capítulo 1 apresenta esta introdução. O restante do trabalho será organizado da seguinte forma:

• Capítulo 2 - Revisão da Literatura

Neste capítulo serão abordados os conteúdos teóricos necessários ao desenvolvimento do trabalho. Assim, serão abordados:

- O funcionamento de um neurônio biológico e como a informação é processada e flui através de um conjunto de neurônios;
- Modelo matemático representativo do neurônio biológico desenvolvido por Hodgin-Huxley (7) por se tratar do modelo mais complexo presente na literatura e de onde vários outros modelos se derivaram e também o modelo de Izhikevich (1) utilizado como base para o desenvolvimento do circuito;

- Princípio dinâmico translinear (DTL) utilizado para construção dos circuitos;
- Conceitos dos transistores halo-implantados, tais quais aspectos construtivos e características operacionais.
- Capítulo 3 Proposta

Este capítulo apresentará a tese em si, definindo em detalhes os problemas a serem resolvidos, a definição das matrizes de transistores halo-implantados e suas características capazes de resolverem ou ao menos minimizarem os efeitos dos implantes halo e assim tornar possível a utilização destes dispositivos em aplicações translineares de ultra-baixa-tensão e ultra-baixa-potência.

Capítulo 4 - Aplicação

Este capítulo apresentará um filtro de primeira ordem passa-baixa que implementa a variável de acomodação do modelo de Izhikevich (1) com comparações entre o comportamento modelado e medições do mesmo a fim de se demonstrar a viabilidade dos transistores halo-implantados como componentes translineares. Será abordado também o circuito desenvolvido capaz de emular o modelo em questão, com comparações entre resultados medidos e simulados com os resultados obtidos através da simulação do modelo em MATLAB®.

• Capítulo 5 - Conclusão

Por fim serão apresentadas as conclusões do trabalho desenvolvido e trabalhos futuros identificados durante a realização do mesmo.

• Apêndice A - Artigos Publicados

Neste apêndice são apresentados os artigos oriundos desta tese.

• Apêndice B - Layout do Circuito Integrado Desenvolvido

Neste Apêndice são apresentados os *layouts* desenvolvidos na tecnologia IBM130nm. Também está presente uma foto do circuito integrado fabricado.

• Apêndice C - Arranjos para Testes e Medições

Este apêndice apresenta como foram realizados os testes e medições, tanto do circuito da acomodação analisado como filtro translinear passa-baixa quanto para o circuito completo que emula o neurônio modelado por Izhikevich.

• Apêndice D - *Reset* Preciso da Acomodação

Este apêndice apresenta um circuito alternativo e preciso para o *reset* da acomodação. Com ele é possível determinar analiticamente o valor da corrente de *reset* do circuito de acomodação.

2 Revisão da Literatura

Neste capítulo serão abordados os conteúdos teóricos necessários ao desenvolvimento do trabalho. Assim, na Seção 2.1 será abordado o funcionamento de um neurônio biológico e como a informação é processada e flui através de um conjunto de neurônios. Depois, na Seção 2.2 serão apresentados o modelo matemático representativo do neurônio biológico desenvolvido por Hodgkin-Huxley (7) por se tratar do modelo mais complexo presente na literatura e de onde vários outros modelos se derivaram e também o modelo de Izhikevich (1) utilizado como base para o desenvolvimento desta tese. Na Seção 2.3 é abordado o princípio dinâmico translinear (DTL) utilizado para construção dos circuitos. Finalmente será apresentado na Seção 2.4 conceitos dos transistores halo-implantados, tais quais aspectos construtivos e características operacionais.

2.1 O Neurônio

O neurônio é a célula do sistema nervoso responsável pela condução do impulso nervoso. Estima-se que existam 86 bilhões de neurônios no sistema nervoso humano (14). Ele é constituído, basicamente, pelo corpo celular, núcleo celular, dendritos (prolongamentos numerosos e curtos do corpo celular que recebem mensagens oriundas de outros neurônios), axônio (prolongamento que transmite o impulso nervoso vindo do corpo celular) e bainha de mielina. O neurônio pode ser considerado a unidade básica da estrutura do cérebro e do sistema nervoso. Na Fig. 1 vê-se a estrutura celular do neurônio. A membrana exterior toma a forma de vários ramos extensos chamados dendritos que recebem sinais elétricos de outros neurônios, e também de uma estrutura chamada axônio que envia sinais elétricos a outros neurônios. O espaço entre o dendrito de um neurônio e os axônios de outro é o que se chama de fenda sináptica. Nestas fendas o impulso nervoso de um neurônio é transferido a outro, em processo conhecido como sinapse, por uma variedade de substâncias químicas chamadas neurotransmissores. Estas sinapses estão representadas na Fig. 1 por círculos que destacam o espaço intercelular entre dentritos de dois ou mais neurônios.

Existem regiões no cérebro onde os neurônios se interconectam em maior número e se especializam para realizar tarefas mais definidas. Uma destas regiões, que contém o neurônio modelado por (1) e que será utilizado por esta tese é o córtex cerebral. O córtex cerebral é um tecido fino composto essencialmente por uma rede de neurônios densamente interligados tal que nenhum neurônio está a mais do que algumas sinapses de distância de qualquer outro neurônio. Esta região do cérebro é de grande importância para a engenharia biomédica por se tratar da região cerebral que recebe e processa toda a informação recebida através das terminações nervosas espalhadas pelo corpo (14). Em outras palavras, esta região é responsável pela *interface* entre nervos e cérebro ou em termos de engenharia, entre sensores e processador.

Os neurônios recebem continuamente impulsos elétricos nas sinapses de seus dendritos vindos de milhares de outras células. Os impulsos geram ondas de corrente elétrica (excitatória ou inibitória, cada uma num sentido diferente) através do corpo da célula até uma zona chamada de zona de disparo, no começo do axônio. É aí que as correntes atravessam a membrana celular para o espaço extracelular e que a diferença de potencial que se forma na membrana determina se o neurônio dispara ou não.



Figura 1 – Neurônio Biológico.

2.1.1 Funcionamento do Neurônio

O neurônio é uma célula altamente especializada na transmissão de informações, na forma de impulsos nervosos. Os impulsos nervosos são fenômenos eletroquímicos que utilizam certas propriedades e substâncias da membrana plasmática, que permitem que seja criado e transmitido um impulso elétrico.

Um neurônio em repouso é uma célula que possui uma diferença de potencial entre o seu citoplasma e o líquido extracelular. Esta diferença de potencial é criada graças ao acúmulo seletivo de íons de potássio (K^+) e sódio (Na^+) , que ocorre pela ação de bombas que criam uma diferença de concentração. Esta diferença de concentração é controlada por canais de íons de potássio e sódio, gerando uma tensão negativa (de aproximadamente -58 mV no interior de neurônios humanos), que pode variar entre espécies animais (14). Este valor, como será visto no Capítulo 2, Seção 2.2.2, no modelo que será usado neste trabalho é considerado de -60 mV pois foi realizado tendo como base o córtex cerebral de ratos.

Este estado de equilíbrio (ou estado de polarização do neurônio) dura até o momento em que um potencial de ação abre os canais iônicos de potássio e sódio, alterando a concentração destes íons. Esta modificação gera um potencial positivo dentro do neurônio, chegando aos +40 mV ou mais (dependendo do organismo estudado, porém como será visto no Capítulo 2, Seção 2.2.2, no modelo que será usado neste trabalho o máximo valor do potencial de membrana é de +30 mV.). Este desequilíbrio gera um efeito cascata, que é o potencial de ação. Usualmente o potencial de ação se inicia no começo no axônio (zona de disparo) e se propaga até as vesículas sinápticas, gerando a descarga de neurotransmissores.

Após ter ocorrido o potencial de ação, imediatamente os canais iônicos de sódio e potássio começam a restabelecer o equilíbrio anterior, com uma tensão negativa no interior do neurônio e positiva fora dele. O neurônio precisa, então, de um brevíssimo tempo para reconstituir seu estado pré-descarga, e durante este tempo ele não consegue efetuar outro potencial de ação. Este período de latência chama-se período refratário. Logo em seguida, o neurônio adquire sua capacidade para efetuar outro potencial de ação, estabelecendo um ciclo.

2.1.2 Neurônios Tálamo-Corticais

Também chamados de neurônios Associativos ou Conectores ou ainda Inter Neurônios, estão basicamente presentes no córtex cerebral. É o grupo de neurônios mais numeroso. Como o nome indica, estes neurônios transmitem o sinal desde os neurônios sensitivos ao sistema nervoso central. Liga também neurônios motores entre si.

O modelo de Izhikevich, utilizado nesta tese para construção do circuito integrado capaz de gerar os padrões neurais, utilizou do estudo em particular destes neurônios, em córtex cerebrais de ratos, para a geração de suas equações matemáticas como será visto no Capítulo 2, Seção 2.2.2. Esta tese tomou como referência este modelo justamente por representar o comportamento de tais neurônios, dentre algumas outras características a serem expostas nos próximos capítulos. A geração de circuitos capazes de emular os comportamentos neurais dos neurônios corticais abre caminho para a criação de *interfaces* entre os tecidos nervosos e próteses, uma grande área de estudo em dias atuais na área de engenharia biomédica.

2.1.3 Sinapses

Sinapses nervosas são os pontos onde as extremidades de neurônios vizinhos se encontram e o estímulo passa de um neurônio para o seguinte por meio de mediadores químicos, os neurotransmissores. A sinapse é considerada uma estrutura formada por membrana pré-sináptica, fenda sináptica e membrana pós-sináptica (14).

As sinapses ocorrem na *interface* das terminações nervosas chamadas axônios, usualmente com os dendritos de outro neurônio, mas pode haver contato com o corpo celular e mesmo com outros axônios (menos comum já que a célula neural conta com a bainha de mielina que isola outras regiões do corpo celular diferentes de dentritos e apêndices dos axônios). O contato físico em sinapses químicas não existe realmente, pois há um espaço entre elas, denominado de fenda sináptica, onde ocorre a ação dos neurotransmissores. Dos axônios, são liberadas substâncias (neurotransmissores), que atravessam a fenda e estimulam os receptores pós-sinápticos.

A literatura aponta a existência de dois tipos de sinapses neuronais: as sinapses químicas e as sinapses elétricas. Ambos os tipos de sinapses transmitem o potencial de ação para outros neurônios, diferindo apenas no mecanismo de comunicação (químico ou elétrico).

2.1.3.1 Sinapses Químicas

As sinapses químicas consistem na maioria das sinapses presentes no sistema nervoso. Ela consiste numa fenda presente entre o axônio do neurônio que está transmitindo a informação (neurônio pré-sináptico) e o neurônio que receberá uma descarga de neurotransmissores, o receptor (neurônio pós-sináptico) (14).

Quando o impulso nervoso atinge as extremidades do axônio, libertam-se para a fenda sináptica os neurotransmissores, que se ligam a receptores da membrana da célula seguinte, desencadeando o impulso nervoso, que, assim, continua a sua propagação.

A chegada do impulso nervoso até o botão sináptico, que é a parte do neurônio pré-sináptico que irá liberar os neurotransmissores, provocará uma reação de liberação de vesículas sinápticas, carregadas com neurotransmissores. Estas substâncias passarão pela fenda sináptica atingindo sítios receptores dos dendritos dos neurônios pós-sinápticos, o que provavelmente irá gerar um potencial de ação provocando um impulso nervoso, que passará pelo corpo celular e prosseguirá até o axônio.

2.1.3.2 Sinapses Elétricas

Alguns neurônios comunicam-se através de sinapses menos comuns, que são as sinapses elétricas, que são junções muito estreitas entre dois neurônios. Estas junções comunicantes são constituídas por proteínas chamadas de conéxons, que permitem uma continuidade entre as células e dispensa, em grande medida, o uso de neurotransmissores. Este tipo de sinapse reduz muito o tempo de transmissão do impulso elétrico entre os neurônios, sendo ideal para comportamentos que exigem rapidez de resposta. Organismos como lagostins, que necessitam fugir com velocidade de predadores, possuem sinapses elétricas em vários circuitos (14).

Outros sistemas que se beneficiam com a sincronização de neurônios também utilizam este tipo de sinapse, como por exemplo neurônios do tronco encefálico, que controlam o ritmo da respiração e em populações de neurônios secretores de hormônios. Esta sincronização facilita a descarga hormonal na corrente sanguínea. Estas junções também chamadas de abertas estão em abundância no músculo cardíaco (discos intercalares) e músculo liso (corpos densos) (14).

2.1.4 Padrões Neurais

Neurônios corticais e tálamo-corticais apresentam 20 padrões de disparos e rajadas de potenciais de ação como mostrados na Fig. 2. Cada padrão é definido por uma sequência única de potenciais de ação. Dependendo da velocidade com que os potenciais de ação são gerados, os padrões são denominados de disparos ou rajadas de potenciais de ação.

O foco principal deste trabalho concentra-se num circuito capaz de emular todos os 20 padrões visualizados na Fig. 2 já que estes neurônios tem uma aplicabilidade muito grande na utilização em engenharia biomédica. Este circuito será desenvolvido com base no modelo de Izhikevich (1) capaz de emular todos estes padrões.

2.2 Modelos Neurais

Ao longo do tempo vários modelos matemáticos surgiram na tentativa de representar o comportamento biológico de células neurais dos mamíferos. Ao final da segunda guerra mundial, estes modelos tomaram por base o advento da neurociência, cuja ciência teve um grande avanço ao final da grande guerra. Desde então alguns modelos tiveram maior ou menor destaque na literatura. A seguir, dois modelos serão apresentados:

- o modelo de Hodgkin-Huxley, devido sua importância histórica e por permitir uma maior compressão do funcionamento da célula neural,
- e o modelo de Izhikevich, poderoso do ponto de vista em riqueza de detalhes porém simples e computacionalmente eficiente, muito utilizado em implementações em circuito integrado nos dias de hoje.



Figura 2 – Padrões de Disparos e Rajadas do Modelo de Izhikevich (1).

2.2.1 Modelo de Hodgkin-Huxley

O primeiro modelo capaz de representar todos os 20 padrões de células neurais documentas pela neurociência foi desenvolvido por Alan Hodgkin e Andrew F. Huxley em 1952 (7). Basicamente o modelo descreve as interações elétricas entre íons de potássio e sódio e os impulsos nervosos presentes no meio circundante às células nervosas. Um



esquema elétrico correspondente ao modelo pode ser visto na Fig. 3.

Figura 3 – Circuito Elétrico Equivalente ao Modelo de Hodgkin-Huxley.

O modelo diz que as correntes podem fluir através da membrana do neurônio tanto através da carga da capacitância da membrana (representada por C_m na Fig. 3) quanto pelo movimento dos íons de sódio e potássio através das condutâncias g em paralelo à capacitância C_m da membrana. A corrente iônia é dividida em componentes formadas por íons de potássio e sódio (I_{Na} e I_K , na Fig. 3). Uma pequena corrente de fuga formada por cloretos e outros íons também é formada, como representada na Fig. 3 por I_L . A corrente iônica de cada íon é determinada pelo produto de sua respectiva condutância ge a diferença entre os potenciais de membrana V_m e o potencial de equilíbrio para o íon representado na Fig. 3 por E. Assim:

$$I_K = g_k n^4 (V_m - E_K) , (2.1)$$

$$I_{Na} = g_{Na}m^3h(V_m - E_{Na}) , (2.2)$$

$$I_L = g_L (V_m - E_L) . (2.3)$$

Assim, a equação diferencial para o circuito da Fig. 3 pode ser derivada como

$$C_m \frac{dV_m}{dt} = I_{ext} - I_K - I_{Na} - I_L .$$
 (2.4)

A equação (2.4) adiciona as probabilidades de ativação e desativação $n, m \in h$ dadas pelas equações diferenciais abaixo:

$$\frac{dn}{dt} = \alpha_n (1-n) - \beta_n n = \frac{n_\infty - n}{\tau_n}, \text{ ativação lenta de potássio,}$$
(2.5)

$$\frac{dm}{dt} = \alpha_m (1-m) - \beta_m m = \frac{m_\infty - m}{\tau_m}, \text{ ativação rápida de sódio},$$
(2.6)

$$\frac{dh}{dt} = \alpha_h (1-h) - \beta_h h = \frac{h_\infty - h}{\tau_h}, \text{ ativação lenta de sódio.}$$
(2.7)

A taxa de abertura de canal $\alpha(V_m)$ e fechamento $\beta(V_m)$ são definidas como

$$\alpha_n(V_m) = \frac{10 - V_m}{100 \left[\exp\left(1 - V_m/10\right) - 1\right]} ,$$

$$\beta_n(V_m) = \frac{1}{8} \exp(-V_m/80)$$
(2.8)

е

$$\alpha_m(V_m) = \frac{25 - V_m}{10 \left[\exp\left(2.5 - V_m/10\right) - 1 \right]} ,$$

$$\beta_m(V_m) = 4 \exp(-V_m/18)$$
(2.9)

е

$$\alpha_h(V_m) = \frac{7}{100} \exp\left(-V_m/20\right) ,$$

$$\beta_h(V_m) = \frac{1}{1 + \exp\left(3 - V_m/10\right)} .$$
(2.10)

O modelo apresenta 4 variáveis $(V_m, n, m \in h)$. A ativação do sódio m segue as dinâmicas da tensão de membrana V_m . De igual forma, a inativação de sódio 1 - h e a ativação do potássio n se aproximam.

No entanto, apesar do modelo ser capaz de descrever todos os 20 comportamentos neurais descritos na literatura (vide Fig. 2), o equacionamento acima nos mostra o quão complexo ele é. Esta complexidade torna o modelo inviável na realização de circuitos simples que o desenvolvam sem grandes custos de área e consumo de energia, características indispensáveis em desenvolvimento em tecnologia CMOS integrável em pastilha de silício. Assim, usando as características do modelo, várias versões simplificadas foram criadas como FitzHugh-Nagumo (8), Morris-Lecar (15), Mihalas–Niebur (9), bem como várias implementações em circuitos destes modelos, (16, 17, 18), respectivamente.

Destes modelos oriundos do modelo de Hodgkin-Huxley, no início dos anos 2000 surgiu o modelo de Izhikevich, o qual vem se destacando pela simplicidade matemática e poder na geração dos 20 padrões neurais descritos na Fig. 2. Este é o modelo utilizado nesta tese e por isso na Seção 2.2.2 ele será descrito com mais detalhes.

2.2.2 Modelo de Izhikevich

O modelo adotado para a implementação do neurônio artificial é o modelo de Izhikevich (1). A opção por este modelo se dá pela representação muito próxima da realidade biológica com uma modelagem simples do ponto de vista matemático e consequentemente da implementação em circuito eletrônico.

O modelo é descrito por duas equações diferenciais de primeira ordem que descrevem o potencial de membrana (v) e acomodação (u), (2.11) e (2.12) respectivamente, com um estado de pós *reset* para as mesmas (2.13).

$$\dot{v} = 0,04v^2 + 5v + 140 - u + I_{DC} , \qquad (2.11)$$

$$\dot{u} = a(bv - u) . \tag{2.12}$$

Se $v \geq 30$ mV, então

$$\begin{cases} v = c , \\ u = u + d . \end{cases}$$
(2.13)

O potencial (v) representa o potencial da membrana do neurônio e (u) representa a variável de recuperação desta membrana, que na realidade descreve a ativação de correntes iônicas de íons de potássio (K^+) e desativação de íons de sódio (Na^+) , provendo assim uma realimentação negativa à membrana (v). Após o disparo atingir seu ápice em aproximadamente 30 mV, o potencial da membrana e a variável de recuperação são reiniciadas de acordo com (2.13). Correntes de sinapse ou correntes DC injetadas são dadas via variável I_{DC} e significam as correntes que estimulam os neurônios. Apesar de parecer estranho que uma corrente seja somada a diferenças de potenciais na equação (2.11), deve-se lembrar que esta modela matematicamente o neurônio e, na verdade, a variável I_{DC} já recupera o valor das correntes sinápticas integradas pela membrana neural, o que resultaria em um valor de tensão.

A equação diferencial $0, 04v^2+5v+140$ de (2.11) é obtida por regressão da dinâmica do disparo de um neurônio cortical de forma que o potencial da membrana v esteja em escala de mV e o tempo esteja em ms. O potencial de repouso do modelo está entre -70 e -60 mV dependendo do valor de b.

Os parâmetros utilizados em (2.11) e (2.12) e suas funcionalidades na modelagem do comportamento dinâmico dos neurônios são descritas a seguir. O efeito que cada parâmetro efetua nos potenciais de ação são mostrados graficamente na Fig. 4. Alterando-se estes parâmetros geram-se as 20 diferentes sequências de potenciais de ação visualizadas na Fig. 2.

- Parâmetro a este parâmetro descreve a escala de tempo da variável de recuperação u. Quanto menor seu valor mais devagar é a recuperação. Um valor típico é 0,02;
- Parâmetro b descreve a sensibilidade da variável de recuperação a flutuações subthreshold do potencial de membrana v. Maiores valores acoplam $v \in u$ mas fortemente

resultando em possíveis oscilações subthreshold e dinâmicas de disparo de baixo threshold. Um valor típico é 0,2;

- Parâmetro c descreve o valor do potencial de membrana v reset após o disparo. Um típico valor é -65 mV;
- Parâmetro d descreve o reset após o disparo da variável de recuperação u. Um valor típico é 2.



Figura 4 – Parâmetros do Modelo de Izhikevich e suas Atuações nos Potenciais de Membrana (v) e Acomodação (u).

2.2.2.1 Categorização das Sequências de Potencias de Ação

Os 20 padrões da Fig. 2 podem ser divididos em dois tipos: excitatórios e inibitórios. Dentro desses dois tipos pode-se ainda subclassificar 5 comportamentos neurais mais importantes, 3 excitatórios e 2 inibitórios, como descrito a seguir.

• Excitatórios

Disparo Regular - Regular Spiking (RS) (Fig. 5a);
Rajada Intrínseca - Intrinsically Bursting (IB) (Fig. 5b);
Chattering (CH) (Fig. 5c).

• Inibitórios

Disparo Rápido - *Fast Spiking* (FS) (Fig. 6a); Disparo Sub-Limiar - *Low Threshold Spiking* (LTS) (Fig. 6b). Apesar do modelo descrito em (1) e o circuito desenvolvido serem plenamente capazes de gerar todos os 20 padrões mostrados na Fig. 2, esta tese se concentrará na emulação destas 5 sequências de potenciais de ação mais importantes visualizadas nas Figs. 5 e 6.



Figura 5 – Padrões Excitatórios - Neurônios Corticais.



Figura 6 – Padrões Inibitórios - Neurônios Corticais.

A adoção deste modelo numa implementação de circuito integrado MOS com dispositivos polarizados abaixo da tensão de limiar permite a utilização de uma topologia de circuito DTL (19) que, juntamente com o uso de dispositivos halo-implantados, permitirá o desenvolvimento de um neurônio artificial de baixíssimo consumo de potência, como será visto no Capítulo 4.

2.3 Circuitos DTL

Existe um grande interesse no desenvolvimento de circuitos DTL nos dias de hoje. Este interesse é devido, basicamente, à adaptabilidade que esses circuitos possuem para aplicações de ultra-baixa-tensão de alimentação, baixa potência, grande faixa de operação linear, altas frequências de resposta e, principalmente, por funcionarem sem a necessidade de polarização (20).

Este tipo de circuito teve grandes aplicações com uso de transistores bipolares devido a sua função de transferência característica de origem exponencial. Nas duas últimas décadas, vêm sendo usados também em circuitos integrados em tecnologia MOS, com transistores polarizados abaixo da tensão de limiar já que, nesta região de operação, o dispositivo MOS opera de forma semelhante a um transistor bipolar, como em (2.14).

$$I_{DS} = I_S \left(\frac{W}{L}\right) \exp\left(\frac{V_{GS} - V_{TH}}{nU_T}\right) \left[1 - \exp\left(\frac{-V_{DS}}{U_T}\right)\right] , \qquad (2.14)$$

na qual:

- V_{DS} é a tensão entre dreno e fonte;
- V_{GS} é a tensão entre *gate* e fonte;
- V_{TH} é a tensão de limiar independente da polarização de corpo;
- U_T é a tensão térmica kT/q;
- I_S é a corrente característica reversa;
- *n* é um parâmetro dependente da tecnologia utilizada. Equivale a 1,16 para o processo IBM 130 nm;
- W é a largura de canal;
- L é o comprimento de canal.

Caso o transistor opere saturado, tem-se $V_{DS} \gg U_T$. Assim, pode-se escrever:

$$I_D = I_S \left(\frac{W}{L}\right) \exp\left(\frac{V_{GS}}{nU_T}\right) \ . \tag{2.15}$$

Pode-se ainda incluir os efeitos geométricos na corrente de saturação reversa I_S . Assim,

$$I_D = I_S \exp\left(\frac{V_{GS}}{nU_T}\right) \ . \tag{2.16}$$

Com a utilização de (2.16) é possível desenvolver circuitos translineares DTL com transistores MOS, como visto a seguir.

A Fig. 7 é usada para explicar o princípio DTL em transistores MOS. O princípio dinâmico é conseguido pela colocação de um capacitor C como referência de tensão no



Figura 7 – Princípio DTL em Transistores MOS Operando em Subthreshold e Saturado.

gate do transistor MOS. Uma outra fonte DC pode ou não ser adicionada em conjunto ao capacitor já que sua derivada será zero e não influenciará no comportamento do circuito. Aplicando (2.16) vem:

$$I_{Out} = I_S \exp\left(\frac{V_C + V_{DC}}{nU_T}\right) . \tag{2.17}$$

Derivando (2.17),

$$\dot{I}_{Out} = I_{Out} \frac{1}{nU_T} \dot{V}_C$$
 (2.18)

Como a derivada da tensão em um capacitor é proporcional à sua corrente, tem-se:

$$\dot{V}_C = \frac{I_{Cap}}{C} \ . \tag{2.19}$$

Finalmente, substituindo (2.19) em (2.18) tem-se a equação que descreve o princípio DTL aplicado a transistores MOS:

$$\dot{I}_{Out} = I_{Out} \frac{1}{nU_T} \frac{I_{Cap}}{C}$$
 (2.20)

A equação (2.20) viabiliza computar a derivada de um sinal através deste mesmo sinal, desde que se tenha um dispositivo que opere de forma exponencial. Ela formará a base para todo o desenvolvimento dos circuitos no Capítulo 3 que se utilizam do princípio translinear para desenvolver um artefato físico que consiga emular o modelo de Izhikevich.

2.4 Implante Halo

A diminuição do comprimento de canal de transistores MOS é uma busca constante da tecnologia de circuitos integrados. Tal diminuição permite a integração de uma quantidade maior de dispositivos por pastilha de silício. Dessa forma, circuitos digitais têm seu desempenho ampliado em termos de aumento de frequência de operação, diminuição da potência dissipada e aumento da complexidade da arquitetura.

No entanto, tal miniaturização causa efeitos indesejáveis na operação do transistor MOS, a serem discutidos na Seção 2.4.3. Estes efeitos são muito significativos em circuitos analógicos. Para se diminuir tais efeitos, começou-se a implantar ou depositar uma nova região do transistor denominada de implante halo ou, mais ainda, implante *pocket* ou implante bolso. A construção de tais implantes é brevemente descrita na Seção 2.4.1.

2.4.1 Construção

Um implante halo é a adição de uma fina região altamente dopada de mesmo tipo do substrato contra as paredes de junção para limitar a extensão de regiões de depleção, visto na Fig. 8. Nesta figura, dois implantes halo são mostrados, um junto à fonte e outro junto ao dreno.

A construção basicamente se dá após a formação do contato de *gate* e geralmente ao mesmo tempo da extensão do dreno fracamente dopada. Os implantes são depositados ou implantados ionicamente, dependendo do processo.

A Fig. 8b apresenta um esquema 3D de um transistor nMOS halo-implantado desenvolvido em ISE-TCAD, ainda sem a adição do polissilício de *gate* para melhor visualização dos implantes halo (2).

2.4.2 Modelos de Corrente de Dreno

Ao longo da última década têm surgido alguns trabalhos que modelam a corrente de dreno de dispositivos halo-implantados (21, 22, 23).

Porém, todos os modelos analíticos até então carecem de uma simplicidade tal qual é dada pela equação (2.16) que modela a corrente de dreno para transistores MOS operando em inversão fraca e saturados.

Para ilustrar a complexidade destes equacionamentos, a seguir é colocado um dos modelos mais recentes presentes na literatura, utilizado na criação dos parâmetros que descrevem o BSIM4v3 (23).

Primeiramente, o modelo parte das seguintes premissas:

- É aplicada a aproximação de Boltzmann;
- Não há corrente fluindo através do dielétrico do *gate* e não há recombinação na região de canal, o qual leva a uma corrente constante de canal;



(b) ISE-TCAD (Sem *Gate* para Visualização dos Implantes Halo) (2). Figura 8 – Transistor nMOS Halo Implantado.

- O potencial de Fermi φ_{FP} é constante no canal;
- Não há carga de interface;
- Não há corrente fluindo na direção vertical do canal;
- Não há confinamento quântico.

É importante ressaltar que as aproximações acima citadas, embora acrescentem grande erro na modelagem, principalmente no que diz respeito a correntes de fuga, são utilizadas em (23) no intuito de se facilitar a análise e possibilitar a criação de uma modelagem analítica mais simples da corrente de dreno em dispositivos halo-implantados. Assim, tem-se que a densidade de corrente no canal é dada por

$$J_n(x,y) = -qn_e(x,y)\mu_n(x,y)\frac{d\varphi_{FN}(x)}{dx} , \qquad (2.21)$$

na qual:

- q é a carga elementar do elétron;
- n_e é a concentração de elétrons;
- μ_n é a mobilidade dos portadores;
- φ_{FN} é o potencial de Fermi do elétron;
- x denota a distância lateral com relação ao dielétrico do gate;
- y denota a distância vertical com relação ao dielétrico do gate;
- a coordenada (x, y) = (0, 0) corresponde ao ponto abaixo do limite do *gate* do lado da região de fonte dentro do Silício (8a).

Como a corrente é definida como a soma de todos os elétrons presente no canal dada a densidade de corrente, tem-se:

$$I(x) = \int_0^{t_{Si}} J_n(x, y) dy = -q \frac{d\varphi_{FN}(x)}{dx} \int_0^{t_{Si}} n(x, y) \mu_n(x, y) dy .$$
 (2.22)

A equação (2.22) define a corrente de canal por unidade de largura do canal e t_{Si} é a espessura do óxido abaixo do *gate*.

Aplicando as considerações descritas, pode-se escrever a corrente de difusão por unidade de largura de canal como

$$I_D = -\frac{qn_{io}^2}{G_{CH}} \left[1 - \exp\left(-\frac{V_{DS}}{U_T}\right) \right] , \qquad (2.23)$$

e G_{CH} é dado por

$$G_{CH} = \int_0^L \left\{ \int_0^{t_{Si}} \left[\frac{n_{ie}(x,y)}{n_{io}} \right]^2 \left[\frac{D_n(x,y)}{p(x,y)} \right] dy \right\}^{-1} dx , \qquad (2.24)$$

em que:

- n_{io} é a concentração de portadores intrínseca do silício;
- D_n é o coeficiente de difusão;
- U_T é a tensão térmica;

Através das equações (2.23) e (2.24) vê-se que a determinação da corrente de dreno para dispositivos halo-implantados é complexa e, ainda mais, é dependente de variáveis nem sempre fornecidas pelo fabricante, como a concentração e densidade de portadores, bem como fatores geométricos como o comprimento dos implantes halo.

Com modelos complexos, torna-se inviável a análise analítica de circuitos com transistores MOS que contenham tais implantes. Este fato justifica o método adotado no Capítulo 3 para o projeto dos elementos translineares utilizados na construção do circuito proposto através de simulações sucessivas para ajuste fino do comportamento de tais elementos. Outra justificativa é que o método analítico de avaliação da corrente de dreno descrita acima (23), apesar da complexidade e ser a base do modelo BSIM4 de versões anteriores a BSIM4v3, não conseguia prever a corrente de dreno de forma correta para algumas situações (24). O modelo de simulação BSIM4v3 já incorpora as modificações previstas em (24) e é o modelo utilizado nas simulações desta tese.

Como também será visto, no Capítulo 3 será demonstrado que o arranjo destes transistores MOS halo-implantados em matrizes nos permitirá adotar um modelo para a corrente de dreno na forma de (2.16) com um coeficiente de determinação maior que 99,8%.

2.4.3 Efeitos dos Implantes Halo

Comprimentos de canal ultrapequenos (dezenas a poucas centenas nanométricas) aumentam os efeitos de canal curto dos transistores uniformemente dopados. Ou seja, o controle do mesmo através de sua polarização de *gate* cede espaço para a polarização do dreno, causando uma diminuição da tensão de limiar, efeito conhecido como diminuição da barreira induzida do dreno (*Drain Induced Barrier Lowering* - DIBL) (3).

O DIBL ocorre nos dispositivos uniformemente dopados devido ao compartilhamento de cargas entre o *gate* e as junções entre o canal e as regiões de fonte e dreno.

Ao se aplicar uma tensão no dreno do dispositivo MOS convencional de canal longo, a barreira de potencial da junção de fonte praticamente não sofre alterações. Porém, para transistores de canal curto, essa barreira de potencial pode sofrer alterações (redução da barreira de potencial de fonte) dependendo do potencial aplicado ao dreno.

Elevando-se a tensão entre dreno e fonte, a região de depleção de dreno aumenta, fazendo com que o *gate* perca parte do controle das cargas de depleção, gerando uma redução da tensão de limiar do dispositivo. Este efeito é visualizável na Fig. 9. Quanto mais curto ou mais estreito é o canal, maior a variação na tensão de limiar do dispositivo.

Para se mitigar este efeito, tecnologias MOS nanométricas têm oferecido opções de implementação de transistores MOS com implantes halo buscando restabelecer o controle do dispositivo por sua polarização de *gate* e assim diminuindo os efeitos de canal curto.



Figura 9 – Efeitos de Canal Curto e Canal Estreito em Dispositivos Uniformemente Dopados na Tensão *Threshold* (3).

No entanto, esses implantes geram uma série de outros efeitos. Apenas a redução dos efeitos de canal curto mostrados na Fig. 9, a que os implantes halo se propõem, seria de grande utilidade em circuitos MOS analógicos. Porém, há uma contrapartida. Em geral estes dispositivos sofrem de um efeito de canal longo conhecido como diminuição da barreira induzida do dreno em dispositivos de canal longo (*Long Channel Drain Induced Barrier Lowering* - LDIBL) (24, 25).

Este efeito torna a magnitude da impedância de saída destes transistores haloimplantados uma ordem de magnitude menor (pelo menos) se comparada aos dispositivos uniformemente dopados para os comprimentos de canal normalmente utilizados em circuitos analógicos ($L \ge 5L_{Minimo}$), o que influencia negativamente o ganho intrínseco do dispositivo (11, 12).

Outro efeito não desejado para circuitos analógicos é a degradação da transcondutância destes dispositivos (13, 10), tornando-os dispositivos pobres para aplicação como elementos translineares práticos, como será discutido no Capítulo 3. Esta característica inviabiliza sua utilização em circuitos DTL operando com correntes diminutas, como é o caso desta tese.

Há ainda um outro efeito gerado pelos implantes halo. A corrente de fuga apresentada por estes dispositivos é muito elevada (26), chegando a dezenas ou centenas de picoamperes. Isto inviabiliza a utilização de transistores halo-implantados em circuitos analógicos que operem com correntes de polarização extremamente baixas e, no caso de circuitos DTL, manipulem sinais extremamente pequenos em amplitude.

A topologia matricial de associação destes transistores, a ser apresentada no Capítulo 3, foi primeiramente desenvolvida em (27), porém para regiões de operação linear e inversão forte, e apenas para transistores regulares ou uniformemente dopados. Nesta tese, no entanto, esta técnica matricial será utilizada no Capítulo 3 para transistores
halo-implantados operando em inversão fraca. As análises feitas apresentam resultados promissores na diminuição do efeito LDIBL (e suas consequências) dos transistores halo-implantados. Com as matrizes foi possível utilizar transistores halo-implantados como componentes translineares práticos, em um circuito DTL operando com correntes na faixa de dezenas de picoamperes, como será visto nos resultados obtidos no Capítulo 4.

3 Proposta

Este trabalho propõe como solução para os problemas descritos no Capítulo 2, Seção 2.4, a construção de elementos translineares através da disposição matricial de transistores halo-implantados. Em (11, 12) já foi mostrado que tal construção é capaz de mitigar o problema da baixa impedância intrínseca destes componentes, que deteriora o ganho de amplificadores. A seguir será demonstrado que outras características negativas destes implantes halo na construção de circuitos analógicos também são mitigadas.

3.1 Matrizes de Transistores Halo-Implantados

A utilização de associações de transistores em matrizes foi primeiramente abordada em (27). Através de análise algébrica é criada uma modelagem para a implementação de matrizes de transistores capazes de melhorar o ganho dos dispositivos através do aumento da impedância de saída dos mesmos. No entanto, toda a análise é feita para transistores uniformemente dopados (sem implantes halo) e ainda operando na região de triodo e saturação em inversão forte.

Recentemente, em (11, 12), a mesma abordagem de matrizes foi utilizada para transistores halo-implantados operando em inversão fraca. Neste caso, esta abordagem mostrou a mesma eficiência no aumento de impedância de saída, diminuindo consideravelmente o efeito LDIBL dos mesmos.

A mesma abordagem é utilizada nesta tese. A implementação dos transistores halo-implantados em forma de matrizes m por n, como mostrado na Fig. 10, supera os problemas mencionados no Capítulo 2, Seção 2.4, que os dispositivos halo-implantados apresentam. Dessa forma, essa abordagem matricial resulta em:

- Aumento da impedância de saída (impedâncias obtidas na ordem de gigaohms, aumentando o desempenho dos espelhos de corrente dos laços translineares);
- Diminuição da alta fuga de corrente de dreno (necessário para circuitos operando com baixíssimas correntes, na ordem de dezenas de picoamperes, como é o caso desta tese);
- Diminuição da degradação da transcondutância (aumenta a robustez a ruídos externos ao circuito).

Assim é possível fabricar circuitos MOS operando em modo corrente alimentado por apenas 250 mV, manipulando correntes diminutas que proporcionam a possibilidade



Figura 10 – Matrizes de Transistores Halo-Implantados.

de manipulação ou mesmo emulação dos lentos sinais biológicos sem a necessidade de utilização de grandes capacitâncias de integração.

A seguir é demonstrado como as matrizes foram projetadas para utilização como componentes translineares e resultados de medições são confrontados com simulações a fim de se comprovar o efeito que a utilização das matrizes tem na correção e mitigação dos efeitos gerados em transistores halo-implantados, como visto no Capítulo 2, Seção 2.4.

3.1.1 Implementação das Matrizes

A Fig. 10 apresenta as matrizes $n \in p$, respectivamente, construídas para a utilização no circuito do neurônio proposto.

A dimensão m representa o número de transistores em série. A dimensão p representa o número de transistores em paralelo. Desta forma, forma-se uma matriz de transistores de m linhas por p colunas.

Para o projeto de cada uma destas matrizes foi utilizado um espelho de corrente como apresentado na Fig. 11a. Neste simples circuito de espelhamento, uma corrente de polarização I_{dc} polariza um espelho n em M_1 . A corrente espelhada em M_2 polariza um espelho p em M_3 de forma a se obter a mesma corrente de polarização I_{dc} em M_4 .



Figura 11 – Espelhos de Corrente Utilizados para Projeto das Matrizes.

Este circuito foi utilizado para o correto dimensionamento das matrizes $n \in p$. Através de simulações utilizando BSIM4v3, as correntes de dreno de $M_2 \in M_4$ foram obtidas e comparadas a I_{dc} e as dimensões das matrizes foram ajustadas de forma a se obter o melhor espelhamento possível (obviamente levando em consideração a área utilizada). Este procedimento foi utilizado por não haver modelagem algébrica da corrente de dreno independente de fatores de processo muitas vezes indisponíveis para projetistas de circuitos integrados para dispositivos halo-implantados (21, 22, 23). Desta forma, sem uma modelagem simples da corrente de dreno como (2.16), torna-se impraticável o desenvolvimento de um modelo de dimensionamento da matriz similar ao realizado em (27) para transistores MOS sem implantes halo.

O dimensionamento das matrizes é então feito de forma a se obter o melhor espelhamento de corrente possível no meio da faixa esperada de excursão da corrente de dreno de cada transistor (entre 0 e 150 pA). Como pode ser observado na Fig. 11b, o erro de espelhamento é maior nos extremos (próximo a 0 pA e a 150 pA) e menor na região central (em torno de 75 pA).

Para se diminuir o descasamento estatístico de cada matriz, a dimensão p deve ser aumentada. Em contrapartida, obtém-se a diminuição da impedância de saída das mesmas, diminuindo o desempenho como fontes de corrente. Para reverter a perda de impedância, incrementa-se a dimensão m, obtendo em contrapartida, a necessidade de uma tensão de alimentação maior para suprir não mais a polarização do *gate* em saturação sub-limiar, mas sim para suprir as próprias correntes envolvidas no circuito.

Através de várias simulações, o dimensionamento apresentado na Fig. 10 foi o melhor obtido em termos de desempenho como espelhos de corrente com descasamento estatístico aceitável para tensão de alimentação $V_{DD} = 250$ mV.

Para se avaliar este desempenho, foi realizada uma simulação de Monte Carlo com 300 rodadas avaliando-se a variação DC da corrente de polarização I_{dc} da Fig. 11a. Para cada uma destas simulações, foram realizadas regressões lineares e calculados os coeficientes angulares e lineares das resultantes, bem como o fator R_2 de cada uma destas regressões, para os transistores M_2 , M_3 e M_4 .



Figura 12 – Coeficientes Angulares das Regressões Lineares dos Espelhamentos.

Os coeficientes angulares em torno do valor 0,97 (Fig. 12) em conjunto com um fator R^2 muito próximo de 1 (Fig. 14) demonstram que, para a faixa de corrente de trabalho (0 a 150 pA), as matrizes funcionam como bons espelhos, já que o valor ideal para o coeficiente angular seria 1, para uma correspondência de um para um entre I_{dc} e um fator R^2 próximo a 1 indica que as regressões lineares das correntes espelhadas por M_2 , M_3 e M_4 podem ser realmente expressas por retas.

Já coeficientes lineares das regressões lineares entre 2,5 pA e 4,5 pA (Fig. 13) demonstram que as correntes de fuga para as matrizes são realmente baixas, já que este coeficiente representa o cruzamento da equação da reta obtida com o eixo y, ou seja, o valor das correntes espelhadas quando a corrente de polarização I_{dc} é zero.



Figura 13 – Coeficientes Lineares das Regressões Lineares dos Espelhamentos.



Figura $14 - R^2$ das Regressões Lineares dos Espelhamentos.

Por questão de comparação, foram feitas as mesmas simulações para os espelhos de corrente mostrados na Fig. 11a porém substituindo as matrizes n e p por transistores n e p halo-implantados de tamanho efetivo de canal similar ao tamanho efetivo de cada matriz n e p. Estas simulações foram realizadas no intuito de mostrar que transistores halo-implantados não são bons espelhos de corrente quando se opera em faixas de correntes muito pequenas (da ordem de dezenas a centenas de picoamperes) como é o caso do circuito implementado nesta tese.

Como pode ser observado na Fig. 15 o coeficiente angular médio fica bem abaixo de 1 como no caso das matrizes mostradas anteriormente, indicando que tais transistores não fornecem correto espelhamento de corrente (para cada corrente de polarização I_{dc} aplicado em M1, uma corrente diferente é espelhada em M_2 , M_3 e M_4 , mesmo tendo fator R^2 muito próximos de 1 (Fig. 17), o que indica que as regressões lineares captam apropriadamente a natureza linear do espelhamento).

Já analisando os coeficientes lineares das regressões lineares da Fig. 16, vê-se que as correntes de fuga estão muito acima de um valor tolerável para o projeto em questão visto que correntes de fuga na ordem de 500 pA é algumas vezes maior que a maior corrente esperada para o projeto de 150 pA.

Na Seção 3.1.2 a seguir, as curvas características das matrizes e de transistores halo-



Figura 15 – Coeficientes Angulares das Regressões Lineares dos Espelhamentos - Halo Implantados de Tamanho Equivalente à Matriz.



Figura 16 – Coeficientes Lineares das Regressões Lineares dos Espelhamentos - Halo Implantados de Tamanho Equivalente à Matriz.



Figura 17 – R^2 das Regressões Lineares dos Espelhamentos - Halo Implantados de Tamanho Equivalente à Matriz.

implantados de tamanho equivalente às matrizes foram levantadas a fim de corroborar com o acima descrito no intuito de provar que as matrizes de transistores halo-implantados são componentes translineares melhores que transistores halo-implantados de tamanho equivalente às matrizes.

3.1.2 Curvas Características

Através de medições realizadas em espelhos $p \in n$ difundidos em silício para auxílio nas medições do circuito principal, as matrizes de transistores halo-implantados $n \in p$ foram caracterizadas em termos de função de transferência e transcondutância, como também função de saída e impedância de saída.

Foram extraídas as funções de transferência das matrizes $p \in n$, como observado nas Fig. 18 e Fig. 19, respectivamente. Estes dados foram confrontados com dados de simulações para as mesmas matrizes e também para um transistor único, de dimensões de canal equivalentes à soma das dimensões de cada canal dos elementos que compõem cada matriz. Os resultados são vistos a seguir.

• Função de Transferência e Transcondutância

Transistores halo-implantados sofrem do efeito de LDIBL (21, 22, 23). Basicamente, este efeito deteriora a transcondutância dos transistores (13) a ponto de, como visto na Figs. 18 e 19, para as correntes envolvidas neste projeto, praticamente perder-se o efeito transistor. Ou seja, a transcondutância fica tão degradada que variações muito pequenas de V_{GS} têm um efeito enorme na corrente de dreno I_D . Este efeito inviabiliza muitas aplicações analógicas pois torna os circuitos muito susceptíveis a ruído.



Figura 18 – Comparação entre Funções de Transferência - pMOS - Matriz x Transistor Halo de Tamanho Equivalente.



Figura 19 – Comparação entre Funções de Transferência nMOS - Matriz x Transistor Halo de Tamanho Equivalente à Matriz.

Nas Figs. 18 e 19 visualiza-se também um outro problema de transistores haloimplantados. A alta corrente de fuga (na ordem de algumas dezenas de pA para a matriz p e na ordem de centenas de pA para a matriz n) presente nestes dispositivos também inviabiliza sua utilização em circuitos operando com baixíssimas correntes. Para as matrizes, estas correntes de fuga são praticamente desprezíveis.

Há ainda uma outra vantagem das matrizes em relação à utilização de transistores halo-implantados únicos. Como pode ser visto nas Figs. 18 e 19, o descasamento estatístico é menor para os transistores associados em matrizes.

Como será visto no Capítulo 4, estas características obtidas pela utilização das matrizes viabilizam a aplicação de transistores halo-implantados não só a projetos de circuitos analógicos (11, 12), mas também operando como componentes translineares operando em circuitos DTL.

Nas Figs. 20 e 21 são plotadas as transcondutâncias em função de V_{GS} para facilitar a visualização da maior taxa de variação de I_D em função de V_{GS} para os transistores halo-implantados, quando utilizados sem a associação em matrizes. Observa-se, também por essas curvas, a recuperação da transcondutância obtida pela associação matricial de transistores halo-implantados.



Figura 20 – Comparação entre Transcondutâncias - pMOS - Matriz x Transistor Halo de Tamanho Equivalente.



Figura 21 – Comparação entre Transcondutâncias n
MOS - Matriz x Transistor Halo de Tamanho Equivalente à Matriz.

• Função de Saída e Impedância de Saída

As curvas de saída das matrizes $p \in n$ podem ser vistas nas Fig. 22 e Fig. 23. As mesmas foram extraídas para polarizações de *gate* equivalentes em corrente, já que as medições foram realizadas nos espelhos desenvolvidos no circuito integrado, e não em matrizes isoladas. Assim, para uma corrente de polarização de 150 pA, V_{SG} é de aproximadamente 195 mV para matriz $p \in V_{GS}$ é de aproximadamente 100 mV para a matriz n.

Novamente os dados medidos foram confrontados para as matrizes e para transistores halo-implantados, $p \in n$, isolados e com tamanho de canal equivalente à soma das dimensões das áreas de todos os canais dos transistores que compõem cada matriz.

Para as matrizes p, como mostrado na Fig. 24 percebe-se um aumento de impedância em torno de uma ordem de grandeza. Já para as matrizes n, como visualizado na Fig. 25, percebe-se um aumento de impedância de saída da ordem de duas ordens de grandeza. Ambos os resultados são expressivos, considerando que implementações DTL práticas com dispositivos CMOS necessitam que os elementos translineares tenham grande desempenho como fontes e espelhos de corrente.



Figura 22 – Comparação entre Funções de Saída - pMOS - Matriz x Transistor Halo-Implantado de Tamanho Equivalente.

As curvas respectivas às matrizes mostram um ótimo casamento, tanto para as matrizes p como para as matrizes n, como mostrado nas Fig. 24 e Fig. 25, respectivamente. Elas também demonstram que a impedância de saída é aumentada



Figura 23 – Comparação entre Funções de Saída nMOS - Matriz x Transistor Halo-Implantado de Tamanho Equivalente à Matriz.



Figura 24 – Comparação entre Impedâncias - pMOS - Matriz x Transistor Halo-Implantado de Tamanho Equivalente.



Figura 25 – Comparação entre Impedâncias - nMOS - Matriz x Transistor Halo-Implantado de Tamanho Equivalente.

em algumas ordens de grandeza, se comparada à impedância dos transistores de tamanho equivalente. Este comportamento já era esperado já que a topologia matricial mitiga os efeitos LDIBL criados pelos implantes halo, como foi visto pelo comportamento das funções de transferência.

4 Aplicação

O neurônio implementado é visualizado na Fig. 26. O mesmo é composto por dois circuitos principais que representam a membrana e a acomodação do neurônio biológico real e de uma circuitaria de apoio para se obter os *resets* do modelo. Estes circuitos de apoio são representados por um comparador de corrente com histerese e um reset do integrador da membrana.



Figura 26 - Neurônio Desenvolvido.

Os dois circuitos principais (membrana e acomodação) estão presentes na literatura (19). Porém, neste mesmo artigo o autor faz ressalvas que os transistores utilizados não são halo-implantados (são utilizados transistores de óxido de *gate* espesso, que na tecnologia adotada corresponde a dispositivos uniformemente dopados) para se evitar justamente os efeitos descritos no Capítulo 2, Seção 2.4.

Neste trabalho, contudo, são utilizados justamente os dispositivos de óxido de *gate* delgados, que no processo de fabricação IBM utilizado, correspondem aos dispositivos halo-implantados. Através da implementação do circuito correspondente ao modelo de Izhikevich e também da análise separada do circuito de acomodação, espera-se comprovar a melhoria de desempenho dos transistores halo-implantados a ponto de permitir utilizá-los como componentes translineares em circuitos DTL de ultra-baixa-tensão e ultra-baixa-potência.

Este capítulo também apresenta os desenvolvimentos adotados para o circuito de reset não especificado em (19).

4.1 Acomodação

A variável de acomodação (u) descrita pelo modelo de Izhikevich (2.12) pode ser implementada por um circuito denominado célula de Gilbert. Trata-se de um filtro passa-

baixa translinear que pode ser visto na Fig. 27 e cujo princípio básico está descrito no Capítulo 2, Seção 2.3.



Figura 27 – Acomodação.

Basicamente, o equacionamento do circuito é feito na forma:

$$I_{1u}I_v = (I_{Cu} + I_{2u} - I_{1u})I_u$$

$$I_{1u}I_v = (I_{2u} - I_{1u})I_u + I_{Cu}I_u$$

$$I_{Cu}I_u = I_{1u}I_v - (I_{2u} - I_{1u})I_u .$$
(4.1)

Do princípio dos circuitos DTL descrito no Capítulo 2, Seção 2.3, tem-se que

$$I_{Cu} = C_u n U_T \frac{\dot{I}_u}{I_u} . aga{4.2}$$

Substituindo 4.2 em 4.1,

$$C_{u}nU_{T}\frac{\dot{I}_{u}}{I_{u}}I_{u} = I_{1u}I_{v} - (I_{2u} - I_{1u})I_{u}$$

$$C_{u}nU_{T}\dot{I}_{u} = I_{1u}I_{v} - (I_{2u} - I_{1u})I_{u} .$$
(4.3)

A equação 4.3 pode ser facilmente comparada à equação 2.12 que descreve o potencial de acomodação no modelo.

De posse da equação que rege o funcionamento da membrana, calculam-se as correntes I_{1u} , I_{2u} e a capacitância de integração da acomodação C_u .

Primeiramente, adotou-se $I_{1u} = 10$ pA para se manter as constantes de tempo do circuito equivalentes à constante de tempo prevista no modelo, sem que seja necessária a utilização da capacitância de integração C_u em valores muito elevados que ocupariam muita área de silício. O valor de 10 pA também se mostrou funcional considerando as correntes de fuga envolvidas nas matrizes, descritas no Capítulo 3.

Assim:

е

$$nU_T = 0.03$$
.

 $I_{1u} = 10 \text{ pA}$

Para o projeto, usou-se também o padrão de disparo de potenciais de ação denominado Disparos Regulares - RS, em que a = 0.02 e b = 0.2. Então, igualando os termos de (2.12) a (4.3) tem-se:

$$C_u = \frac{I_{1u}}{abnU_T} = 83,3 \text{ pF}$$
 (4.4)

е

$$I_{2u} = 6I_{1u} = 60 \text{ pA} . (4.5)$$

E importante ressaltar que a tecnologia adotada para implementação do circuito integrado possui camadas que possibilitam a integração de altos valores de capacitância com a utilização de uma área pequena e por isso, neste trabalho, não foi utilizado escalonamento de tempo no modelo como feito em (19), artifício frequentemente utilizado na literatura em implementações analógicas de modelos neurais.

A fim de corroborar com a análise das matrizes de transistores halo-implantados funcionando como componentes translineares, estudou-se este circuito da acomodação, a priori sem considerar seu *reset*, ou seja, analisou-se o circuito apenas como um filtro passa-baixas. No entanto, para se isolar os efeitos que apenas os dispositivos translineares incorporam ao circuito, foi feita apenas uma análise estática do filtro, como será visto a seguir.

4.1.1 Circuito de Acomodação - Análise como Filtro Translinear Passa-Baixa

O circuito de acomodação apresentado na Fig. 27 nada mais é que um filtro passa baixas translinear desenvolvido através de um circuito DTL. Este circuito é o laço dinâmico translinear mais simples existente. Assim, este circuito se torna bastante útil para se avaliar separadamente o comportamento das matrizes de transistores halo-implantados como elementos translineares.

Tomando a equação (4.3) e comparando-a com a equação que compreende a acomodação no modelo de Izhikevich (2.12), percebe-se que as duas se apresentam na mesma forma. Assim, pode-se dizer que os parâmetros $a \in b$ do modelo correspondem em laço translinear, respectivamente:

$$a = \frac{I_{2u} - I_{1u}}{C_u n U_T}$$
(4.6)

$$b = \frac{I_{1u}}{I_{2u} - I_{1u}} \ . \tag{4.7}$$

Ambas as equações (4.6) e (4.7) apresentam o termo $I_{2u} - I_{1u}$. Desta forma, caso I_{1u} e I_{2u} sejam variados na mesma proporção, a diferença continua inalterada, mantendo a proporção entre os parâmetros a e b do modelo. Isto é posto pois se quer analisar apenas o impacto do desempenho das matrizes no laço translinear e assim pode-se dizer que para uma mesma proporção entre as correntes I_{1u} e I_{2u} o circuito DTL da Fig. 27 deverá responder com o mesmo valor acomodado, variando apenas a dinâmica até atingir este valor. Como a capacitância C_u é fixa, faremos medições no circuito da Fig. 27 da seguinte forma:

- Polariza-se o circuito com I_{1u} e I_{2u} iguais a 10 pA e 60 pA, respectivamente (valores escolhidos arbitrariamente que correspondem a valores típicos para operação de padrões de disparos regulares do neurônio);
- Varia-se a corrente I_v ;
- Espera-se que I_u acomode;
- Mede-se o valor acomodado de I_u .

Os valores de I_u medidos foram confrontados com os dados obtidos através de (4.3) simulados em MATLAB®. A Fig. 28a mostra que o laço translinear apresentado na Fig. 27 funciona muito bem com as matrizes de transistores halo-implantados como componentes translineares. O valor acomodado de I_u responde praticamente da mesma forma que o previsto pelo modelo. Apenas para valores acima de 130 pA a diferença entre os valores medidos e simulados aumenta significantemente. Isto acontece pois correntes maiores geram maiores V_{DS} e, consequentemente, acabam por atingir valores próximos a V_{DD} . Dessa forma, a corrente I_u começa a aumentar mais lentamente. No caso do projeto isto não é um problema pois a faixa de trabalho para correntes é de aproximadamente 130 pA. Dentro da faixa de projeto o erro máximo entre I_u medido e simulado não atinge 6%.



Figura 28 – Comportamento Acomodado do Circuito Dinâmico Translinear.

Dentro da faixa média de correntes (onde o circuito irá operar a maior parte do tempo), no entanto, este erro é muito menor e praticamente não atinge 1%.

Na Fig. 28b foram plotados a resposta de I_u pela variação de I_v para o modelo e medidas com o laço translinear implementado com matrizes (exatamente como na Fig. 28a) e também a resposta simulada do mesmo laço translinear da Fig. 27 desenvolvido com transistores halo-implantados de tamanho equivalentes a matriz. Basicamente devido às altas correntes de vazamentos que estes dispositivos apresentam, a reta de resposta de I_u se apresenta bastante deslocada da reta obtida pelo modelo. Já a diferença entre as inclinações é basicamente dada pelas distorções destes dispositivos devido sua baixa impedância intrínseca e altíssima transcondutância que amplifica as diferenças de tensão V_{GS} entre os espelhos devido aos descasamentos.

4.2 Membrana

A membrana (v) descrita pelo modelo de Izhikevich (2.11) também pode ser implementada pela célula de Gilbert, como o circuito da acomodação anterior, com uma modificação para representar o termo não linear de (2.11). Esta modificação é obtida pela utilização de uma realimentação positiva de I_v através de M_{15} e M_{16} na Fig. 29 representada no equacionamento por m, que na realidade é o fator de espelhamento entre M_{15} e M_{16} .

Basicamente, o equacionamento do circuito é dado por

$$(I_{in} + I_{DC} - I_u)I_{1v} = (I_{2v} + I_{Cv} - I_{1v} - I_v)I_v .$$

$$(4.8)$$



Figura 29 – Membrana.

Dividindo (4.8) por I_{1v} , tem-se

$$(I_{In} + I_{DC} - I_u) = (I_{2v} + I_{Cv} - I_{1v} - I_v) \frac{I_v}{I_{1v}} .$$
(4.9)

Do princípio DTL desenvolvido no Capítulo 2, Seção 2.3, I_{Cv} é dado por

$$I_{Cv} = C\dot{V} = C\dot{V}_{GS_{11}} = C_v n U_T \frac{\dot{I}_v}{I_v} .$$
(4.10)

Assim obtém-se a equação que descreve o circuito equivalente do modelo da membrana:

$$\frac{C_v n U_T}{I_{1v}} \dot{I_v} = \frac{m}{I_{1v}} I_v^2 - \frac{I_{2v} - I_{1v}}{I_{1v}} I_v + I_{in} + I_{dc} - I_u .$$
(4.11)

No entanto, o modelo Izhikevich pressupõe equações matemáticas em escala de tempo de ms e potenciais em mV. Porém o modelo descreve números absolutos. Para a adaptação do modelo matemático em um sistema eletrônico real, será adotado:

$$v = I_v - 100 , \qquad (4.12)$$

$$u = I_u - 100b . (4.13)$$

As adaptações realizadas em (4.12) e (4.13) são feitas para que todas as correntes envolvidas no circuito tenham um único sentido e assim a implementação seja simplificada (19).

Substituindo (4.12) e (4.13) em (4.11) tem-se:

$$\dot{I}_v = 0,04(I_v - 100)^2 + 5(I_v - 100) + 140 + I_{in} - I_u$$
$$\dot{I}_v = 0,04I_v^2 - 3I_v + 40 + 100b + I_{in} - I_u .$$
(4.14)

Percebe-se que (4.14) se assemelha a (2.11). Assim, igualando-se os termos de ambas as equações e tomando $nU_T = 0.03$ para o processo adotado,

$$\frac{C_v n U_T}{I_{1v}} = 1 \text{ ms} . ag{4.15}$$

Fazendo $I_{1v} = 25$ pA por ser um valor razoável para as faixas de espelhamento de corrente adotada nas matrizes desenvolvidas no Capítulo 3, tem-se $C_v = 3,3$ pF.

Assim, pode-se calcular os parâmetros de projeto restantes:

$$\frac{m}{I_{1v}} = 0,04 \Rightarrow m = 1 , \qquad (4.16)$$

$$\frac{I_{2v}}{I_{1v}} - 1 = 3 \Rightarrow I_{2v} = 4I_{1V} = 100 \text{ pA} ,$$
 (4.17)

$$Idc = 40 + 100b = 40 + 20 = 60 \text{ pA}$$
. (4.18)

4.3 Circuito de Reset

O reset presente no modelo de Izhikevich é emulado em circuito através de três blocos principais. O primeiro deles é um comparador com histerese capaz de identificar os valores máximo (em torno de 130 pA) e mínimo (parâmetro c) que a variável I_v da membrana poderá assumir. Os outros dois blocos funcionam como conversores de um sinal de corrente em tensão para efetivamente reiniciar os integradores da membrana e acomodação. Tais blocos são descritos a seguir.

4.3.1 Comparador com Histerese

Para a implementação do neurônio faz-se necessária a implementação de um comparador com histerese. Tal circuito é mostrado na Fig. 30.

Do modelo de Izhikevich (1) transformado para modo corrente, I_{Ref} é a máxima corrente I_v que representa o potencial de membrana v. Na Fig. 30, I_{Ref} é a corrente de dreno gerada na matriz M_3 pela aplicação de V_{IREF} como $V_{GS_{M3}}$.

Idealmente, $V_{GS_{nFET}}$ para 130 pA deveria ser igual a $V_{GS_{pFET}}$ para -130 pA, ou seja, igual a 125 mV = $V_{DD}/2$, para que o estágio inversor pudesse ser projetado para transição em $V_{DD}/2$ e assim obter maior margem de ruído.



Figura 30 – Comparador de Corrente com Histerese.

No entanto, a fim de se utilizar as mesmas matrizes NFET 4x36 e pFET 4x32 projetadas para o neurônio no Capítulo 3 e como as mesmas não são perfeitamente casadas, $V_{threshold}$ do inversor foi projetado para $V_{DD} - V_{GS_{pFET}} = V_{GS_{nFET}}$.

Assim sendo, o *threshold* do inversor deve ser de 137 mV, o que ainda gera uma margem de ruído de 113 mV, considerável para uma tensão de alimentação de 250 mV.

A utilização de dois inversores é necessária para a implementação da topologia de comparador com histerese. Ambos geram sinais antissimétricos para acionamentos das chaves de transmissão FET_{33} posicionadas logo acima das matrizes M_2 e M_3 na Fig. 30 e assim acionando ora M_2 , ora M_3 , de acordo com o nível gerado pela comparação entre I_C ou I_{Ref} e I_{Signal} . Esse processo realimenta o sistema e gera os sinais de *resets* do neurônio, V_{Reset} da Fig. 30.

Desta forma, o comparador com histerese apresenta uma função de transferência como visto na Fig. 31. O gráfico mostra as curvas do primeiro e segundo inversores, onde a curva do segundo inversor corresponde à tensão de *reset* V_{Reset} da Fig. 30.

A implementação do circuito comparador com a utilização de matrizes de transistores halo-implantados apresenta uma distribuição estatística para descasamento bastante uniforme para ambos os níveis de comparação, em torno de 30 pA e 130 pA, para os limites de comparação inferior e superior, respectivamente. A distribuição amostral dos limites mínimo e máximo da histerese, respectivamente, são vistos nas Fig. 32 e Fig. 33. Esta análise estatística é importantíssima pois os pontos de comparação do comparador com histerese tem grande influência no comportamento do neurônio desenvolvido.



Figura 31 – Função de Transferência do Comparador de Corrente com Histerese.

O limite inferior corresponde ao parâmetro c do modelo descrito em (2.12) ou I_c em laço translinear, ou seja, variações no limite inferior do comparador tem consequências diretas no circuito de acomodação e consequentemente na sequência de potenciais de ação emuladas pelo circuito.

Já o limite superior do comparador tem efeitos menores no comportamento do circuito, mas ainda relevantes. Se a corrente I_v atinge valores muito elevados (muito acima dos 130 pA estimados), a corrente de acomodação I_u também aumenta. Como este aumento em I_v é na forma de um disparo decorrente da instabilidade da membrana (aumenta muito rapidamente), I_u também aumenta rapidamente (apesar deste aumento ser mais lento que o aumento em I_v devido à capacitância de integração C_u muito maior que a capacitância de integração C_v). Dessa forma, com o reset acontecendo para o nível inferior do comparador (I_c), a variável de acomodação I_u estaria em valores muito diferentes do correto. No caso de I_u estar acima do previsto, o circuito de acomodação deixa os disparos de potenciais de ação mais lentos. Caso I_u esteja abaixo dos valores previstos, o circuito de acomodação deixa os potenciais de ação mais rápidos.



Figura 32 – Variação Estatística 1000 Rodadas - Limite Mínimo Histerese.



Figura 33 – Variação Estatística 1000 Rodadas - Limite Máximo Histerese.

4.3.2 Reset do Integrador da Membrana

O reset da membrana deve ser feito no seu respectivo capacitor de integração C_v . Porém, como todo o circuito é projetado para operar em modo corrente, o valor de tensão a ser aplicado em C_v deve ser oriundo de uma corrente. Neste caso, a corrente I_c irá corresponder ao parâmetro c em (2.13).

Como circuitos DTL MOS operando na condição sub-limiar funcionam de forma a obter uma corrente de dreno com uma relação exponencial a uma tensão V_{GS} (como mostrado no Capítulo 2, Seção 2.3) aplicada ao *gate* do transistor, faz-se necessário, agora, fazer o processo contrário. Através de uma corrente I_c (corrente de *reset* do neurônio equivalente ao parâmetro c do modelo), deve-se obter uma tensão relacionada para ser aplicada como *reset* de C_v . Isto é feito com o circuito mostrado na Fig. 34.



Figura 34 – Circuito de Reset do Integrador da Membrana.

Da equação que descreve o modelo da corrente de dreno em um dispositivo MOS operando em região de sub-limiar e saturado (2.16), pode-se obter V_{GS} na forma

$$V_{GS} = nU_T ln(I_D) . aga{4.19}$$

Então, equacionando o circuito mostrado na Fig. 34 para V_V tem-se:

$$V_{V} = V_{GS_{M13}} - V_{GS_{M14}} = V_{GS_{M12}}$$

$$V_{V} = nU_{T} \left\{ ln \left[\frac{I_{c}}{I_{S}} \right] - ln \left[\frac{I_{IV}}{I_{S}} \right] \right\}$$

$$V_{V} = nU_{T} ln \left[\frac{I_{c}}{I_{1V}} \right]. \qquad (4.20)$$

O resultado de (4.20) é o mesmo obtido caso seja equacionado o circuito da Fig. 29 do ponto de vista de $V_{GS_{M11}}$, como em (4.21), que é a tensão que polariza M_{10} para gerar a corrente de dreno I_v .

$$V_{GS_{M11}} = nU_T ln(I_v)$$

$$V_{GS_{M11}} = V_{C_V} + V_{GS_{M10}}$$

$$V_{C_V} = V_{GS_{M11}} - V_{GS_{M10}}$$

$$V_{C_V} = nU_T \left\{ ln \left[\frac{I_c}{I_S} \right] - ln \left[\frac{I_{IV}}{I_S} \right] \right\}$$

$$V_{C_V} = nU_T ln \left[\frac{I_c}{I_{IV}} \right]. \qquad (4.21)$$

A igualdade de (4.20) e (4.21) comprova que a tensão V_V gerada pelo circuito de reset mostrado na Fig. 34 é a tensão que deve ser aplicada ao integrador do circuito da membrana mostrado na Fig. 29 para gerar a corrente I_v do reset.

4.3.3 Reset do Integrador da Acomodação

Nesta tese foi utilizado como circuito para *reset* do integrador da acomodação apenas uma chave de transmissão acionada pelo comparador com histerese mostrado na Fig.30. Esta chave aciona uma fonte de corrente cujo valor corresponde ao parâmetro d, ou I_d já no domínio da corrente (Fig. 26).

Esta técnica diferencia-se do circuito de *reset* do integrador da membrana devido ao fato de, segundo (2.13) o *reset* da acomodação apenas adicionar uma carga ao integrador da acomodação, ao passo que o *reset* da membrana deve gerar um valor totalmente novo ao integrador da membrana, necessitando de um circuito como na Fig. 34.

Assim sendo, para adicionar mais carga ao integrador da acomodação, insere-se I_d ao mesmo no momento do *reset* sem remover a carga inicial presente.

Apesar de simplificar o projeto, inclusive diminuindo o número de matrizes necessárias ao desenvolvimento do circuito, esta técnica impossibilita o cálculo algébrico da conversão do parâmetro d para I_d . Equacionando o circuito da acomodação presente na Fig. 27 para I_u , tem-se

$$I_u = I_S \exp(\alpha V_{GS_{M4}}) . \tag{4.22}$$

Mas:

$$V_{GS_{M4}} = V_{GS_{M6}} + V_{Cu} . ag{4.23}$$

Então, substituindo (4.23) em (4.22), tem-se

$$I_u = I_S \exp\left[\alpha (V_{GS_{M6}} + V_{Cu})\right] .$$
(4.24)

 I_u é a corrente de acomodação no momento imediatamente anterior ao reset e V_{Cu} é a tensão no capacitor C_u também imediatamente anterior ao reset.

Para se computar o valor de I_d correspondente ao parâmetro d do modelo presente em (2.13) denominou-se I_u imediatamente após o fim do *reset* de I_u' . Assim:

$$I_{u}' = I_{S} \exp\left[\alpha \left(V_{GS_{M6}} + V_{Cu} + \int_{t_{1}}^{t_{2}} I_{d} dt\right)\right] , \qquad (4.25)$$

em que:

- $t_2 t_1$ é o intervalo de tempo de duração do *reset*;
- $\int_{t_1}^{t_2} I_d dt$ é o acréscimo de tensão em V_{Cu} devido a I_d .

Aplicando o logaritmo neperiano a (4.25) vem:

$$lnI'_{u} = \alpha \left\{ V_{GS_{M6}} + \left[V_{Cu} + \frac{1}{C_{u}} \int_{t_{1}}^{t_{2}} I_{d}dt \right] \right\} + lnI_{S} .$$
(4.26)

Resolvendo para I_d :

$$\alpha \left. \frac{I_d^2}{2} \right|_{t_1}^{t_2} = \ln I'_u - \ln I_S - \alpha V_{GS_{M6}} - \alpha V_{Cu}$$

$$\alpha \left. \frac{I_d^2}{2} \right|_{t_1}^{t_2} = \ln \left(\frac{I'_u}{I_S} \right) - \alpha \left(V_{GS_{M6}} - V_{Cu} \right)$$

$$I_d = \sqrt{\frac{2}{\alpha} \left[\ln \frac{I'_u}{I_S} - \alpha \left(V_{GS_{M6}} + V_{Cu} \right) \right]}_{t_1}^{t_2}.$$
(4.27)

Analisando (4.27) percebe-se que existem duas grandezas que não podem ser definidas previamente num dado momento de tempo. São elas:

- $t_2 t_1$ (intervalo de tempo do *reset*) este tempo pode variar de acordo com o padrão sendo executado pelo circuito;
- V_{Cu} a carga anterior ao *reset* no capacitor de integração da acomodação varia de acordo com o padrão sendo executado pelo circuito.

Além disso, mesmo que fosse possível prever com exatidão o intervalo de tempo de reset, seria necessário saber o valor exato da tensão V_{Cu} no tempo imediatamente anterior ao reset pois a função de transferência de M6 não é linear e sim exponencial. Desta forma, um acréscimo de ΔV_{Cu} que pudesse ser calculado por (4.27) teria efeitos diferentes em I'_u , dependendo do próprio valor de I_u . Desta forma, os valores de I_d correspondentes a cada parâmetro apresentado na Tabela 1 são empíricos. Várias simulações foram realizadas a fim de se obter o valor de I_d que apresentasse a maior correspondência entre as simulações e os padrões obtidos pela modelagem de (2.11), (2.12) e (2.13). Ainda assim, de acordo com os resultados obtidos na Seção 4.4, os padrões simulados assemelham-se bastante aos padrões obtidos através do modelo.

No Apêndice D é apresentado um circuito similar ao circuito de *reset* da membrana (Fig. 34) que tornaria possível calcular os valores exatos de I_d de cada padrão neural. Este não foi implementado devido à necessidade de acréscimo de 6 matrizes (4 matrizes nMOS e 2 pMOS), o que custaria maior área ao circuito. Visto que a intenção futura da topologia é a implementação de redes de neurônios com milhares de células, este custo não foi assumido já que a implementação realizada apresentou resultados satisfatórios como será visto na Seção 4.4.

4.3.4 Cálculo dos Parâmetros para Demais Padrões

De posse do modelo de Izhikevich convertido para circuito translinear DTL descrito, calcularam-se os parâmetros de polarização para os demais padrões neurais subcategorizados em excitatórios e inibitórios discutidos no Capítulo 2, Seção 2.2.2.

A Tabela 1 compila os parâmetros $a, b, c \in d$ descritos no modelo e suas correntes correspondentes em laço DTL para os cinco principais padrões de disparo e rajada de potencias de ação. Todas as conversões dos parâmetros do modelo em correntes foram realizadas seguindo o projeto descrito nas Seções 4.1 e 4.2, com exceção da corrente I_d . Como explicado na Seção 4.1 deste capítulo, esta corrente presente na Tabela 1 foi calibrada empiricamente. No Apêndice D é apresentada uma nova tabela de parâmetros convertidos para corrente (Tabela 3) com todos os parâmetros calculados analiticamente. Neste caso, o circuito de *reset* da acomodação a ser utilizado seria o apresentado no Apêndice D.

Tabela 1 – Conversão de Parâmetros - Do Modelo Izhikevich para Circuito DTL.

	Conversão de Parâmetros									
Padrões Neurais	а	b	I_{1u} /	I_{2u} [pA]	\mathbf{c} / I_c	[pA]	\mathbf{d} / I_d	[pA]	$I_{dc} = 40$ ·	+ 100b [pA]
	Modelo	Modelo	DTL		Modelo	DTL	Modelo	DTL	Modelo	DTL
Disparo Regular (RS)	0,02	0,2	10	60	-65	35	8	827	60	60
Rajada Intrínseca (IB)	0,02	0,2	10	60	-55	45	4	882	60	60
Chattering (CH)	0,02	0,2	10	60	-50	50	2	270	60	60
Disparo Rápido (FS)	0,1	0,2	50	300	-65	35	2	1	60	60
Disparo Sub-Limiar (LTS)	0,02	0,25	$12,\!5$	62,5	-65	35	2	100	65	65

4.4 Simulações e Medições do Neurônio Desenvolvido

De posse de todos os parâmetros previstos pelo modelo de Izhikevich (a, b, c e d) nas equações (2.11), (2.12) e (2.13) convertidos para correntes (Tabela 1) polarizou-se o circuito com os valores obtidos. Simulações em BSIM4v3 e medições foram feitas. Os resultados obtidos confrontam dados de simulação do circuito desenvolvido e obtidos pelo modelo de Izhikevich para padrões de disparo e rajadas de potenciais de ação. Resultados de simulação foram utilizados pois o equipamento utilizado para as polarizações e medições adiciona capacitâncias muito elevadas que interferem no comportamento translinear do circuito desenvolvido. No entanto, será apresentada também uma resposta medida do circuito desenvolvido comparada à resposta simulada, caso fossem colocadas nas simulações as mesmas capacitâncias do equipamento utilizado (vide Apêndice C). Os resultados são apresentados a seguir.

4.4.1 Padrões de Disparo de Potenciais de Ação

Como descrito no Capítulo 2 , Seção 2.2, os padrões gerados pelos neurônios corticais podem ser classificados em duas classes principais:

- Padrões Excitatórios: RS, IB e CH.
- Padrões Inibitórios: FS e LTS.

Estes são os padrões de disparo e rajada utilizados para análise do trabalho realizado por se tratarem dos principais potenciais de ação neurais. Entretanto o circuito também é capaz de gerar os demais comportamentos, como descritos no Capítulo 2, Seção 2.2.

Os resultados foram plotados em conjunto com as respostas obtidas através do modelo para comparação nas Figuras 35, 36, 37, 38 e 39.

• Disparos Regulares - RS;

De acordo com o modelo de Izhikevich (1), o padrão RS se caracteriza pelo disparo de alguns potenciais de ação com período curto e, em seguida, disparos de potenciais de ação com períodos mais longos. Essa mudança de tempo entre os potenciais de ação é chamada de Adaptação de Frequência de Disparo.



Figura 35 – Padrão RS - Simulação SPICE vs. Modelo.

• Chattering - CH;

Neurônios disparam rajada de potenciais de ação muito próximos, de alta frequência, que podem atingir cerca de 40 Hz.



Figura 36 – Padrão CH - Simulação SPICE vs. Modelo.

• Rajada Intrínseca - IB.

Neurônios disparam rajada de potenciais de ação seguidos por repetitivos potenciais de ação únicos.



Figura 37 – Padrão IB - Simulação SPICE vs. Modelo.

• Disparos Rápidos - FS;

Executando o padrão FS, neurônios disparam trens de potenciais de ação com alta frequência, praticamente sem Adaptação de Frequência de Disparo.



Figura 38 – Padrão FS - Simulação SPICE vs. Modelo.

• Disparos de Baixo Limiar - LTS;

Neurônios disparam trens de potenciais de ação em alta frequência mas com clara Adaptação de Frequência de Disparo.

Como pode ser visto nas Figs. 35, 36, 37, 38 e 39, os padrões obtidos assemelham-se muito aos padrões obtidos por Izhikevich em seu modelo matemático, salvo algumas



Figura 39 - Padrão LTS - Simulação SPICE vs. Modelo.

diferenças.

Para uma comparação mais precisa entre potenciais de ação, a Fig. 40 mostra um único potencial de ação sendo disparado. Este padrão é chamado de disparo de fase (*Phasic Spike*). Nele é possível visualizar em detalhe o potencial de ação gerado pelo circuito. Tanto a resposta simulada quanto a modelada respondem com similaridade, o que mostra a validade da utilização das matrizes de transistores halo-implantados como componentes translineares.



Figura 40 - Padrão PS - Simulação SPICE vs. Modelo.

4.4.2 Análise Estatística

Para a análise estatística do tempo entre disparos de potenciais de ação foram realizadas 100 simulações transientes de cada padrão. Um programa foi desenvolvido com auxílio do MATLAB® para identificar os picos dos potenciais de ação e determinar o intervalo de tempo entre eles. O código do programa pode ser visto no Apêndice E. Em sequência, histogramas (Fig. 41) para cada padrão de disparo e rajada de potenciais de ação foram obtidos.

Os resultados estatísticos obtidos mostram boa distribuição amostral do tempo entre os potenciais de ação disparados, apesar de algumas amostras se desviarem de 3σ para todos os padrões, como visto na Fig. 41. No entanto, de acordo com Izhikevich os neurônios biológicos apresentam variação estatística de comportamento. Em seu trabalho (1), uma rede neural de 1000 neurônios é simulada. Para a montagem da rede, cada um dos neurônios é instanciado com parâmetros $a, b, c \in d$ que seguem uma distribuição randômica. Esta distribuição randômica é capaz de cobrir todos os parâmetros para todos os tipos de neurônios da rede mas também cobrem as variações que um mesmo tipo de neurônio tem em relação a seu par. Em seu trabalho, a distribuição randômica é centrada nos parâmetros que ajustam o neurônio para o padrão de disparo regular. Desta forma, mesmo obtendo amostras com intervalo de tempo entre disparos incoerentes para um dado padrão, numa rede de milhares de neurônios esse não seria um problema relevante. Mais ainda, numa rede de milhares de neurônios, todos os parâmetros $a, b, c \in d$ convertidos para modo corrente, ou seja, I_{1u} , I_{2u} , I_{1v} , I_{2v} , $I_c \in I_d$ poderiam ser fixados em valores centrados no ajuste de disparo regular (vide Tabela 1) e a variação estatística do próprio circuito seria suficiente para ajustar os neurônios para os diversos padrões.

4.4.3 Medidas dos Comportamentos de Disparo e Rajadas de Potenciais de Ação dos Neurônios

Para realização das medidas, utilizaram-se 3 HRSMUs (Unidade de Fonte e Medição de Alta Resolução - *High Resolution Source Meter Unit*) de um analisador de semicondutores Agilent B1500. Uma delas para suprir a corrente I_d e as outras duas para polarizar I_{2u} e I_{2v} , respectivamente. O esquema completo da ligação de teste do circuito é mostrado no Apêndice C.

No entanto, como estas HRSMUs apresentam capacitâncias de carga que podem chegar até 1 nF, elas têm impacto relevante no funcionamento do circuito por adicionarem capacitâncias de ordem superior às capacitâncias de integração da acomodação e membrana, $C_u \in C_v$, respectivamente.

Assim, foi feita uma medida com todos os parâmetros ajustados para geração do padrão de disparo regular presente na Tabela 1 obtendo, como esperado, um padrão



Figura 41 – Análise de Monte Carlo - Tempo entre Potenciais de Ação.

diferente do padrão obtido por simulação e modelagem na Fig. 35. Adicionou-se então capacitâncias em paralelo às fontes I_{2u} e I_{2v} representadas na Fig. 26 às simulações, variando-se os valores até encontrar padrão semelhante ao obtido por medição. O resultado é mostrado na Fig. 42.

O valor obtido para as capacitâncias das HRSMUs do B1500 para geração de comportamentos do circuito similares entre simulação e medição ficou em 200 pF. Este



Figura 42 – Calibração da Capacitância da SMU.

valor é condizente com o valor sugerido pelo fabricante (capacitância de carga máxima de 1 nF).

No entanto, por ser um valor elevado se comparado a C_u e C_v (83,3 pF e 3,3 pF, respectivamente), a constante de tempo do circuito foi completamente modificada e, como pode ser visto na Fig. 42, os disparos acontecem em ordem de segundos e não em milissegundos, como no modelo e simulações presentes na Fig. 35

A distorção do comportamento do circuito também ocorre devido às capacitâncias de carga da HRSMU adicionadas serem um valor constante, ou seja, elas adicionam um offset de mesmo valor tanto a C_u quanto a C_v . Desta forma, perde-se a proporcionalidade entre as capacitâncias de integração, o que distorce os padrões de disparo e rajadas gerados, já que as relações para se obter os parâmetros de polarização do circuito (4.15) e (4.4) não serem mais respeitadas.

4.5 Compilação de Figuras de Mérito

As principais características obtidas pelo circuito são compiladas na Tabela 2. Os resultados são comparados com (19) a fim de se avaliar os resultados obtidos neste trabalho com um trabalho similar recente, no que diz respeito à implementação do circuito em topologia translinear e não nos elementos translineares em si (nesta tese foram utilizadas matrizes de transistores halo-implantados e em (19), utilizados transistores uniformes de óxido de gate espesso).

		Esta Tese	(19)
	Tecnologia	IBM 130 nm	45 nm
	Unidade []	Simulado	Simulado
Tensão de Alimentação	$\pm V$	250 mV	-
Corrente Total	nA	1,26	_
Área (Caps / Total)	mm^2	0,03/0,042	0,00298/0,0021
Potência Média	pW	316,2	-
Energia / Potencial de ação	рJ	$1,\!6$	< 1
Escalonamento de Amplitude	-	Sim	Sim
Escalonamento de Tempo	-	1x	10x

Tabela 2 – Resultados e Comparação.

Assim sendo, apesar de algumas figuras de mérito importantes não poderem ter sido comparadas devido à ausência das mesmas em (19), figuras importantes para este tipo de projeto puderam ser relativamente comparadas. No entanto, é necessário ainda fazermos algumas considerações ao se analisar a Tabela 2:

- O circuito desenvolvido (Fig. 26) não utiliza escalonamento no tempo, requisitando maiores capacitâncias de integração. Isto explica o fato do circuito ocupar uma área consideravelmente maior que (19).
- Outro fato que colabora com a maior utilização de área neste trabalho é a utilização das correntes numa ordem de grandeza maior para se diminuir os efeitos das correntes de fuga nos transistores em estruturas parasitas na implementação, principalmente devido ao fato de transistores halo-implantados da tecnologia adotada possuírem uma espessura de óxido no *gate* menor que os transistores normais denominados FET33 na tecnologia utilizada.

Ainda é necessário ressaltar que algumas figuras de mérito poderiam ser melhoradas, como por exemplo a diminuição da área utilizada para a confecção do circuito desta tese. O projeto adotado foi bastante conservador no que diz respeito ao tamanho das matrizes elaboradas. O mesmo circuito poderia ser construído com matrizes menores, em ambas as dimensões, e ainda atenderia as necessidades do circuito translinear que implemente o modelo de Izhikevich, porém com contrapartida de ser mais susceptível a variações de processo e descasamento mais severas (estudo crítico não realizado em (19)).

Apesar da necessidade de relativização dos trabalhos comparados para se efetivamente compará-los, pode-se perceber que as matrizes de transistores halo-implantados efetivamente cumprem seu papel como elementos translineares, comprovando a possibilidade de utilização destes dispositivos em circuitos analógicos desenvolvidos neste tipo de topologia.
5 Conclusão e Trabalhos Futuros

Neste trabalho demonstrou-se a utilização de transistores halo-implantados na construção de circuitos DTL. Apesar de haver trabalhos publicados com a abordagem de matrizes de transistores halo-implantados na confecção de circuitos analógicos como em (11, 12), a utilização em circuitos DTL ainda é inédita na literatura.

A motivação para o uso destes dispositivos se deve, principalmente, à possibilidade de redução da tensão de alimentação em circuitos de ultra-baixa-tensão e ultrabaixa-potência, indo ao encontro às aspirações do grupo de microeletrônica da Unifei, que busca desenvolver conhecimento e trabalhos na área biomédica, incluindo dispositivos implantáveis.

No entanto, como discutido no Capítulo 2, Seção 2.4, os implantes halo trazem uma série de efeitos negativos a circuito analógicos. O uso destes transistores como elementos translineares só é possível devido à adoção da abordagem matricial que mitiga os efeitos LDIBL destes dispositivos, como mostrado no Capítulo 3. Nele, mostrou-se que a curva característica do transistor V_{GS} vs. I_D se torna mais suave já que a transcondutância das matrizes é melhorada. Isto torna possível a utilização das matrizes em laços translineares operando baixíssimas correntes sem tornar o circuito extremamente susceptível a variações inoportunas de V_{GS} . Esta é uma vantagem muito grande já que, quanto menores as correntes operadas nos laços translineares, o circuito é mais capaz de processar ou mesmo emular sinais lentos (centenas de milihertz a hertz) sem o acréscimo de capacitâncias de integração de valores insustentáveis à implementação em circuito integrado. Esta é uma grande vantagem na área biomédica devido à natureza lenta dos sinais biológicos.

Também foi demonstrado no Capítulo 3 que as matrizes diminuem as altas correntes de fuga que os transistores halo-implantados apresentam, tornando possível a aplicação destes dispositivos em circuitos que operem sinais extremamente pequenos, como dezenas de picoamperes.

Ainda no Capítulo 3, demonstrou-se que a altíssima impedância de saída gerada pelas matrizes permite a utilização de tais dispositivos na construção de espelhos de corrente de alto desempenho, característica também necessária à implementação de laços translineares práticos.

Para demonstrar a funcionalidade da proposta matricial, desenvolveu-se no Capítulo 4 um laço translinear simples (filtro passa-baixas) capaz de permitir a análise isolada da influência das matrizes como elementos translineares. Medições no circuito mostraram um erro de no máximo 6% dentro da faixa de operação projetada para o circuito (de 30 a 150 pA) para os valores acomodados das correntes (análise estática). No entanto, para a faixa de operação onde o circuito operará a maior parte do tempo, este erro é menor que 1%.

Ainda no Capítulo 4 também foi desenvolvido o circuito completo de um neurônio, emulando o modelo matemático de Izhikevich (1). As simulações obtidas foram encorajadoras. Apesar de não ter sido possível a medição dos padrões neurais de disparo e rajada de potenciais de ação devido à influência dos equipamentos de medição utilizados nas capacitâncias de integração dos laços translineares, foi realizada uma comparação entre a simulação adicionada dos efeitos capacitivos do equipamento e a medição. Como mostrado na Fig. 42, as respostas obtidas são bem similares. Este fato colabora para avaliarmos o funcionamento do circuito implementado como bastante positivo.

No geral, os resultados positivos obtidos nesta tese contribuem não só ao grupo mas também à comunidade científica, com diversos artigos publicados, sendo um em revista e três em congressos internacionais (vide Apêndice A). Os resultados inéditos apresentados abrem uma fronteira ainda pouco explorada que é a utilização de transistores halo-implantados na confecção de circuitos eletrônicos analógicos e, especificamente como apresentado neste trabalho, em circuitos DTL.

5.1 Trabalhos Futuros

Durante a realização deste trabalho, alguns pontos foram observados para trabalhos futuros. São eles:

• Realização de um teste de bancada que não influencie no circuito

Para se medir com exatidão o circuito confeccionado, será necessária uma nova difusão que implemente fontes de correntes e mais espelhos do tipo nFET auxiliares que possam polarizar os nós I_{2u} e I_{2v} , da Fig. 26, sem interferir no laço translinear. Assim, sem adicionar capacitâncias em paralelo a C_u e C_v , o comportamento do neurônio se dará na forma como projetado e as medidas de todos os padrões de disparo e rajadas de potenciais de ação poderão ser realizadas.

• Interconexão de vários neurônios em um único circuito

Este trabalho pode vir a se tornar uma nova tese de doutorado. A interconexão de vários neurônios num único circuito poderia contribuir bastante na comunidade acadêmica. Inicialmente, poderia se tentar interconectar 100 neurônios e comparar com os resultados simulados por Izhikevich em seu artigo (1). Esta comparação poderia, inclusive, ajudar a comprovar se o descasamento intrínseco dos transistores halo-implantados arranjados como matrizes realmente se equipara à variação alea-tória modelada por Izhikevich em sua rede neural, mencionada no Capítulo 4, Seção 4.4.2.

• Utilização dos neurônios como moduladores assíncronos

Já existe na literatura trabalhos (28), que utilizam circuitos que emulam comportamentos neurais como forma de se modular informação de forma assíncrona. A realização de um trabalho similar poderia gerar uma dissertação de mestrado.

• Associação de neurônios a sensores diversos

Poderiam ser realizados trabalhos, em nível de iniciação científica, interligando o circuito de acomodação do neurônio a sensores (temperatura, pressão, etc) a fim de se variar a realimentação negativa da membrana e assim variar o comportamento do neurônio em função da grandeza medida pelo sensor. Uma pesquisa na literatura médica poderia ser feita em conjunto para se verificar se o sistema nervoso periférico reaje da mesma forma que o circuito implementado.

Referências

1 IZHIKEVICH, E. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, v. 14, n. 6, p. 1569 – 1572, nov. 2003. Citado 13 vezes nas páginas 7, 16, 17, 18, 19, 23, 24, 26, 29, 56, 64, 68 e 74.

2 REGNIER, A.; PORTAL, J.; BOUCHAKOUR, R. NMOS electrical model for halo implant study. *Journal of Non-Crystalline Solids*, v. 353, n. 5–7, p. 620 – 624, 2007. Citado 3 vezes nas páginas 7, 32 e 33.

3 EL-KAREH, B. *Silicon Devices and Process Integration*. [S.l.]: Springer, 2009. Citado 4 vezes nas páginas 7, 16, 35 e 36.

4 WIJEKOON, J.; DUDEK, P. Spiking and bursting firing patterns of a compact VLSI cortical neuron circuit. In: *International Joint Conference on Neural Networks*. *IJCNN* 2007. [S.l.: s.n.], 2007. p. 1332–1337. Citado na página 16.

5 DEMIRKOL, A.; OZOGUZ, S. A low power VLSI implementation of the Izhikevich neuron model. In: *IEEE 9th International New Circuits and Systems Conference (NEW-CAS)*,. [S.l.: s.n.], 2011. p. 169–172. Citado na página 16.

6 WIJEKOON, J.; DUDEK, P. Simple analogue VLSI circuit of a cortical neuron. In: *Electronics, Circuits and Systems, 2006. ICECS '06. 13th IEEE International Conference on.* [S.l.: s.n.], 2006. p. 1344–1347. Citado na página 16.

7 HODGKIN, A. F. H. A. L. A quantitative description of membrane current and its application to conduction and excitation in nerves. J. Physics, v. 177, p. 500–544, 1952. Citado 4 vezes nas páginas 16, 17, 19 e 24.

8 FITZHUGH, R. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, v. 1, n. 6, p. 445–446, 1961. Citado 2 vezes nas páginas 16 e 26.

9 GERSTNER, W.; KISTLER, W. Spiking Neuron Models: Single Neurons, Populations, Plasticity. [S.l.]: Cambridge Univ. Press, 2002. Citado 2 vezes nas páginas 16 e 26.

10 BULUCEA, C. et al. Physics, technology, and modeling of complementary asymmetric MOSFETs. *IEEE Transactions on Electron Devices*, v. 57, n. 10, p. 2363–2380, Oct 2010. Citado 2 vezes nas páginas 16 e 36.

11 FERREIRA, L. H. C.; SONKUSALE, S. R. A 60-dB gain OTA operating at 0.25-V power supply in 130-nm digital CMOS process. *IEEE Transactions on Circuits and Systems I: Regular Papers*, v. 61, n. 6, p. 1609–1617, June 2014. Citado 6 vezes nas páginas 16, 17, 36, 38, 45 e 73.

12 COLLETTA, G. D.; FERREIRA, L. H. C.; PIMENTA, T. C. A 0.25-v 22-ns symmetrical bulk-driven OTA for low-frequency G_m -C applications in 130-nm digital CMOS process. *Analog Integrated Circuits and Signal Processing*, Springer US, v. 81, n. 2, p. 377–383, 2014. Citado 6 vezes nas páginas 16, 17, 36, 38, 45 e 73.

13 MCANDREW, C.; DRENNAN, P. Analysis of halo implanted MOSFETs. In: *Technical Proceedings of the 2007 NSTI Nanotechnology Conference and Trade Show*. [S.l.]: Nano Science and Technology Institute, 2007. v. 3, p. 594 – 598. Citado 3 vezes nas páginas 16, 36 e 44.

14 PURVES GEORGE J. AUGUSTINE, D. F. W. C. H. A.-S. L. J. O. M. S. M. W. D. (Ed.). *NEUROSCIENCE*. 3rd edition. ed. [S.l.]: Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts U.S.A., 2004. Citado 4 vezes nas páginas 19, 20, 22 e 23.

15 MORRIS, C.; LECAR, H. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal*, v. 35, n. 1, p. 193–213, 1981. Citado na página 26.

16 SANCHEZ-SINENCIO, E.; LINARES-BARRANCO, B. Circuit implementation of neural FitzHugh-Nagumo equations. In: *Circuits and Systems, 1989.*, *Proceedings of the 32nd Midwest Symposium on.* [S.l.: s.n.], 1989. p. 244–247 vol.1. Citado na página 26.

17 BEHDAD, R. et al. Electrical Morris-Lecar neuron. In: Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE. [S.l.: s.n.], 2013. p. 5001–5004. Citado na página 26.

18 FOLOWOSELE, F.; HAMILTON, T.; ETIENNE-CUMMINGS, R. Silicon modeling of the Mihalas Niebur neuron. *Neural Networks, IEEE Transactions on*, v. 22, n. 12, p. 1915–1927, Dec 2011. Citado na página 26.

19 RANGAN, V. et al. A subthreshold aVLSI implementation of the Izhikevich simple neuron model. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). [S.l.: s.n.], 2010. p. 4164–4167. Citado 6 vezes nas páginas 29, 50, 52, 55, 71 e 72.

20 SERRANO-GOTARREDONA, T.; LINARES-BARRANCO, B.; ANDREOU, A. A general translinear principle for subthreshold MOS transistors. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, v. 46, n. 5, p. 607–616, 1999. Citado na página 30.

21 HO, C. et al. An analytical subthreshold current model for pocket-implanted NMOS-FETs. *IEEE Transactions on Electron Devices*, v. 50, n. 6, p. 1475–1479, 2003. Citado 3 vezes nas páginas 32, 40 e 44.

22 PANG, Y.-S.; BREWS, J. Models for subthreshold and above-threshold currents in 0.1- μm pocket n-MOSFETs for low-voltage applications. *IEEE Transactions on Electron Devices*, v. 49, n. 5, p. 832–839, 2002. Citado 3 vezes nas páginas 32, 40 e 44.

23 HUETING, R. J. E.; HERINGA, A. Analysis of the subthreshold current of pocket or halo-implanted nMOSFETs. *IEEE Transactions on Electron Devices*, v. 53, n. 7, p. 1641–1646, 2006. Citado 5 vezes nas páginas 32, 33, 35, 40 e 44.

24 ROY, A.; MUDANAI, S.; STETTLER, M. Mechanism of long-channel drain-induced barrier lowering in halo MOSFETs. *IEEE Transactions on Electron Devices*, v. 58, n. 4, p. 979–984, April 2011. Citado 2 vezes nas páginas 35 e 36.

25 CAO, K. Advanced Compact Modeling of MOSFETs. Tese (Doutorado) — University of California at Berkeley, 2002. Citado na página 36.

26 ROY, K.; MUKHOPADHYAY, S.; MAHMOODI-MEIMAND, H. Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits. *Proceedings of the IEEE*, v. 91, n. 2, p. 305–327, Feb 2003. Citado na página 36.

27 GALUP-MONTORO, C.; SCHNEIDER, M.; LOSS, I. Series-parallel association of FET's for high gain and high frequency applications. *Solid-State Circuits, IEEE Journal* of, v. 29, n. 9, p. 1094–1101, 1994. Citado 3 vezes nas páginas 36, 38 e 40.

28 YENDURI, P.; GILBERT, A.; ZHANG, J. Integrate-and-fire neuron modeled as a low-rate sparse time-encoding device. In: *Third International Conference on Intelligent Control and Information Processing (ICICIP)*. [S.l.: s.n.], 2012. p. 507–512. Citado na página 75.

Apêndices

APÊNDICE A – Publicações

Neste Apêndice são apresentadas as publicações oriundas deste trabalho.

A.1 Publicações

Como resultados acadêmicos oriundos desta tese um artigo foi publicado em revista Qualis A2 e dois artigos foram publicados em congressos internacionais IEEE. Um outro artigo também foi publicado, em conjunto com outro doutorando, que explora as características da matriz de transistores halo-implantados. A seguir os detalhes de cada publicação:

- Publicações em Revistas
 - Implementation of an ultra-low-power dynamic translinear loop at 0.25-V with halo-implanted 130-nm MOSFETs.
 - * Jornal: Analog Integrated Circuits and Signal Processing Springer
 - * Ano: 2015
 - $\ast\,$ Qualis A2 $\,$
- Publicações em Congressos Internacionais Relacionados a esta Tese
 - A Sub-Threshold Halo Implanted MOS Implementation of Izhikevich Neuron Model
 - * Congresso: IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference
 - * Ano: 2013
 - * Local: Monterey, California EUA.
 - A Sub-Threshold Halo Implanted MOS Implementation of an Electronic Neuron
 - * Congresso: The International Conference on Microelectronics
 - * Ano: 2013
 - * Local: Beirute, Líbano.
- Publicações em Congressos Internacionais Não Relacionados a esta Tese

- An ultra-low-power first-order asynchronous sigma-delta modulator for biomedical applications.
 - * Congresso: IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference
 - * Ano: 2013
 - * Local: Monterey, California EUA.

APÊNDICE B - Layout

Todo o projeto foi desenvolvido em tecnologia IBM 130 nm utilizando-se dos transistores nFET e pFET (halo-implantados) conectados em forma de matrizes $m \ge n$ como anteriormente mostrado no Capítulo 3.

Além dos dispositivos halo-implantados nFET e pFET, foram utilizados também transistores nFET33 para a implementação das chaves de transmissão dos sinais de *reset*. Tais transistores puderam ser utilizados já que a ordem de grandeza de corrente de passagem pelo canal dos mesmos enquanto chaves fechadas é de pico amperes, possibilitando seu uso como chave de transmissão mesmo fora da região de saturação em inversão forte.

O layout desenvolvido para o neurônio é visto na Fig. 43a na região delimitada pelo retângulo vermelho (localizado na região inferior da figura). Uma outra região demarcada em laranja (área central da figura) mostra o layout desenvolvido para o circuito de acomodação isolado (filtro passa-baixas DTL). As demais regiões do layout mostram circuitos de suporte para outros testes. As regiões ativas não podem ser visualizadas na fotografia da Fig. 43b pois o processo em questão adiciona uma camada opaca de óxido acima da camada de metal 3. Assim, o layout foi posto ao lado para identificação dos circuitos.

A Fig. 44 salienta a região delimitada em vermelho na Fig. 43, que corresponde a um único neurônio. A Fig. 45, por sua vez, apresenta a região delimitada em laranja na Fig. 43, mostrando apenas o circuito de acomodação isolado.



Figura 43 – Neurônio Implementado.



Figura 44 – Modelo de Neurônio Artificial de Izhikevich Implementado com Matrizes de Transistores Halo-Implantados.



Figura 45 – Circuito de Acomodação Isolado (Filtro Passa-Baixas DTL) Implementado com Matrizes de Transistores Halo-Implantados.

APÊNDICE C – Testes de Bancada

C.1 Medidas do Neurônio

As medidas foram realizadas utilizando-se de um analisador de semicondutores Agilent B1500A Semiconductor Device Analizer com três SMUs de alta resolução, denominadas pelo fabricante de HRSMU (High Resolution Source Meter Unit). Duas destas HRSMUs foram utilizadas para polarização de I_{2u} e I_{2v} e a terceira para polarização da corrente I_d mostradas na Fig. 46. As 3 HRSMUs foram configuradas como fontes de corrente. Todas as SMUs foram conectadas a um TestFixture também fabricado pela Agilent com alto isolamento a ruídos ambiente. Cabos triaxiais conectam o TestFixture às HRSMUs presentes no B1500 para colaborar com o isolamento a ruídos.

As outras polarizações foram feitas por tensão auxiliadas por 3 *buffers* externos desenvolvidos com o chip CA3140 da Texas Instruments para se diminuir o efeito de carga dos divisores de tensão para regulação da polarização gerada (vide Fig. 47). Tal amplificador operacional foi utilizado por apresentar entrada diferencial CMOS de altíssima impedância, tornando o efeito de carga do divisor de tensão de 1 M Ω utilizado praticamente desprezível. Offset sistemático do mesmo, bem como outras características como resposta em frequência e ganho, são irrelevantes na operação realizada em que é necessário apenas gerar um valor de tensão contínuo na saída para polarização dos espelhos de corrente.

As saídas dos *buffers* foram conectados capacitores de 1 uF para desacoplar ruídos gerados pelo circuito. Todos os *buffers* foram alimentados por bateria de 9 V de forma assimétrica e todo o aparato foi colocado dentro do *TestFixture* Agilent juntamente com o chip em teste. Todo o ensaio é visualizado na fotografia 48 onde o *chip* desenvolvido foi montado em uma tira de *protoboard* e os *buffers* em outra, para se evitar acoplamento de ruídos.

Desta forma, polarizando-se os espelhos Espelho_pFET1, Espelho_pFET2 e ConjEspelhos pelas tensões geradas pelos *buffers* obteve-se as correntes necessárias para polarização dos parâmetros do circuito neurônio.

Um outro *buffer* foi utilizado para geração da tensão de alimentação de 250 mV do neurônio visto na Fig. 46 e desta forma, evitar que ruídos oriundos do meio ambiente e acoplados por cabos não protegidos oriundos de fontes de alimentação interferissem nas medições.

Por último, o Espelho_pFET3 presente na Fig. 46 foi polarizado diretamente com



Figura 46 – Esquemático de Ligação para teste de Bancada - Circuito Neurônio.



Figura 47 – Buffer Utilizado para Geração de Polarizações dos Espelhos por Tensão - Circuito Neurônio.



Figura 48 – Ensaio para Medições do Neurônio.

um gerador de funções com cabo coaxial para se diminuir efeitos de ruídos ambiente. Este foi necessário para gerar o estímulo I_{in} visualizado na Fig. 26 presente no modelo.

C.2 Medidas da Acomodação com Filtro Passa-Baixas

As medições do circuito de acomodação isolado, funcionando como filtro passabaixas em topologia DTL, foram feitas com os mesmos equipamentos do circuito neurônio completo mostrado na Fig. 46. No entanto, como eram menos pontos a serem polarizados, não foi necessária a utilização de *buffers* como mostrados na Fig. 47. Assim, o ensaio foi realizado como mostrado na Fig. 49.

No ensaio, 3 HRSMUs e um gerador de funções foram utilizados. As HRSMUs polarizavam diretamente em corrente enquanto o gerador gerava as correntes I_v através do espelho de corrente auxiliar Espelho_pFET3. É importante ressaltar que todos os espelhos auxiliares utilizados para conversão de tensão em corrente para polarização dos nós do circuito foram caracterizados previamente (curva V_{SG} vs. I_D) a fim de se evitar erros de medições. Dessa forma, era sabido anteriormente a tensão V_{SG} exata que geraria as correntes desejadas.



Figura 49 – Esquemático de Ligação para teste de Bancada - Circuito Neurônio.

APÊNDICE D – *Reset* Preciso da Acomodação

A Fig. 50 apresenta um circuito preciso para o reset da acomodação que torna possível calcular analiticamente o valor da corrente I_d de reset da mesma. O circuito destacado em verde composto pelos transistores M_{19} , M_{20} , M_{21} é similar ao circuito para reset da membrana mostrado na Fig. 34 e a análise é a mesma feita no Capítulo 4, Sessão 4.3.2. Já o circuito em destaque azul, composto pelos transistores M_{22} , M_{23} e M_{24} funciona como espelho da corrente I_u para que esta possa ser somada à corrente I_d no dreno do transistor M_{21} . Este circuito não foi implementado devido à necessidade de acréscimo de 6 matrizes (M_{19} , M_{20} , M_{21} , M_{22} , M_{23} e M_{24} , 4 matrizes nMOS e 2 pMOS, respectivamente), o que custaria maior área ao circuito. Visto que a intenção futura da topologia é a implementação de redes de neurônios com milhares de células, este custo não foi assumido já que a implementação realizada apresentou resultados satisfatórios (como visto no Capítulo 4).



Figura 50 – Circuito de Reset da Acomodação Preciso.

No caso da utilização deste novo circuito no projeto, os parâmetros convertidos para DTL seriam os mesmos da Tabela 1, com exceção do parâmetro d convertido para I_d . Desta forma, os novos parâmetros para o circuito seriam os da Tabela 3.

	Conversão de Parâmetros									
Padrões Neurais	a	b	I_{1u} /	I_{2u} [pA]	\mathbf{c} / I_c	[pA]	d / <i>I</i> _d	[pA]	$I_{dc} = 40$	+ 100b [pA]
	Modelo	Modelo	DTL		Modelo	$\mathbf{D}\mathbf{L}\mathbf{\Gamma}$	Modelo	$\mathbf{D}\mathbf{L}$	Modelo	DTL
Disparo Regular (RS)	0,02	0,2	10	60	-65	35	8	20	60	60
Rajada Intrínseca (IB)	0,02	0,2	10	60	-55	45	4	10	60	60
Chattering (CH)	0,02	0,2	10	60	-50	50	2	5	60	60
Disparo Rápido (FS)	0,1	0,2	50	300	-65	35	2	5	60	60
Disparo Sub-Limiar (LTS)	0,02	0,25	$12,\!5$	62,5	-65	35	2	5	65	65

Tabela 3 – Conversão de Parâmetros - Do Modelo Izhikevich para Circuito DTL.

APÊNDICE E – Código MATLAB®

Para a automatização da análise de Monte Carlo apresentada no Capítulo 4, os dados de simulação retornados pelo Spectre® foram processados pelos *scripts* MATLAB® a seguir, identificados individualmente para cada tipo de padrão de disparo e rajada de potenciais de ação.

Scripts individuais para cada tipo de padrão de disparo do neurônio são necessários devido às variações do intervalo dos disparos para cada padrão, bem como a adaptação em frequência de alguns.

E.1 RS

```
%Trata os dados do cadence (peaks) para transforma-los num
%readble input para o histogram
delta_{RS} = 20e - 3;
x_RS = 0:2:50;
delta_FS = 1e-3;
x_FS = 0:2:100;
delta_{IB} = 10e - 3;
x_{IB} = 0:2:65;
delta LTS = 5e-3;
x\_LTS = 0:2:70;
delta_CH = 15e - 3;
x_CH = 0:2:70;
delta_CH_short = 0.3e-3;
x_CHLong = 0:2:35;
dataRS = csvread('All_Variant_100-runs/RS.csv',1,0);
[num_linhas num_colunas] = size(dataRS);
diferenca_RS = zeros(num_columas/2,1);
 {\it for iteracao=1:2:num\_colunas} \quad \% Varre \ to \ das \ as \ 100 \ iteracoes \\
   maior_pico = 0;
   maior_tempo=0;
   ponteiro_tempo = num_linhas;
   indice_tempo=0;
   indice tempol = 0;
   indice_tempo2 = 0;
   indice_tempo3 = 0;
```

```
indice_menor_tempo = 0;
    segundo_maior_tempo = 0;
    %Verifica o maior valor de pico da leitura
    for picos = 1:1:num_linhas
        if dataRS(picos, iteracao+1) > maior_pico
                maior_pico = dataRS(picos, iteracao+1);
        end
    \mathbf{end}
   \%Verifica o maior tempo dos picos
    for indice_tempo=ponteiro_tempo:-1:1
        if ( dataRS(indice_tempo,iteracao+1) >= maior_pico*0.9
                && dataRS(indice_tempo, iteracao+1) <= maior_pico*1.1)
            maior_tempo = dataRS(indice_tempo, iteracao);
            break:
        end
    \mathbf{end}
    % Verifica o segundo maior tempo evitando os picos intermediarios
    %proximos ao pico principal
    controle = maior_tempo - delta_RS;
    %Acha o indice da linha que contem o tempo do tempo maximo - 1.5ms
    for indice_tempo2=indice_tempo:-1:1
        if (dataRS(indice_tempo2, iteracao) < controle)
            indice_menor_tempo = indice_tempo2;
            break:
        \mathbf{end}
    end
    \% Verifica o segundo maior tempo dos picos
    for indice_tempo3=indice_menor_tempo:-1:1
        if ( dataRS(indice_tempo, iteracao+1) >= maior_pico*0.9
                && dataRS(indice_tempo, iteracao+1) <= maior_pico*1.1)
            segundo_maior_tempo = dataRS(indice_tempo3, iteracao);
            break;
        else
            segundo\_maior\_tempo = 0;
        \mathbf{end}
    end
    if segundo_maior_tempo == 0
        diferenca_RS((iteracao+1)/2, 1) = 0;
    else
        diferenca_RS(((teracao+1)/2, 1) = (maior_tempo - segundo_maior_tempo)*1000;
    end
\mathbf{end}
figure
%title('Histograms');
subplot (3,2,1);
hist(diferenca_RS, x_RS);
ylabel('Samples');
\mathbf{xlabel}('Time_{\sqcup}[ms]');
title('RS');
%hist(differenca_RS);
```

E.2 CH

```
%Trata os dados do cadence (peaks) para transforma-los num
%readble input para o histogram
delta_{RS} = 20e - 3;
x_RS = 0:2:50;
delta_FS = 1e-3;
x_FS = 0:2:100;
delta_IB = 10e-3;
x_{IB} = 0:2:65;
delta LTS = 5e-3;
x_LTS = 0:2:70;
delta_CH = 15e-3;
x_CH = 0:2:70;
delta\_CH\_short = 0.3e-3;
x\_CHLong = 0:2:35;
dataCH = csvread('All_Variant_100-runs/CH_30p.csv',1,0);
[num_linhas num_colunas] = size(dataCH);
diferenca_CH = \mathbf{zeros}(\operatorname{num_columas}/2, 1);
for iteracao=1:2:num colunas %Varre todas as 100 iteracoes
   maior_pico = 0;
   maior\_tempo=0;
   ponteiro_tempo = num_linhas;
   indice_tempo=0;
   indice_tempo1 = 0;
   indice_tempo2 = 0;
   indice_tempo3 = 0;
   indice_menor_tempo = 0;
   segundo_maior_tempo = 0;
   %Verifica o maior valor de pico da leitura
   for picos=1:1:num_linhas
       if dataCH(picos, iteracao+1) > maior_pico
               maior_pico = dataCH(picos, iteracao+1);
       \mathbf{end}
   \mathbf{end}
   \%Verifica o maior tempo dos picos
   for indice_tempo=ponteiro_tempo:-1:1
       if ( dataCH(indice_tempo,iteracao+1) >= maior_pico*0.9
              && dataCH(indice_tempo,iteracao+1) <= maior_pico*1.1)
           maior_tempo = dataCH(indice_tempo, iteracao);
           break:
       end
```

```
\mathbf{end}
   % Verifica o segundo maior tempo evitando os picos intermediarios
   %proximos ao pico principal
    controle = maior_tempo - delta_CH;
   %Acha o indice da linha que contem o tempo do tempo maximo - 1.5ms
    for indice_tempo2=indice_tempo:-1:1
        if (dataCH(indice_tempo2,iteracao) < controle)
           indice_menor_tempo = indice_tempo2;
           break;
       \mathbf{end}
    end
   %Verifica o segundo maior tempo dos picos
    for indice_tempo3=indice_menor_tempo:-1:1
        if ( dataCH(indice_tempo,iteracao+1) >= maior_pico*0.9
               && dataCH(indice_tempo,iteracao+1) <= maior_pico*1.1)
           segundo_maior_tempo = dataCH(indice_tempo3, iteracao);
           break;
        else
            segundo_maior_tempo = 0;
        \mathbf{end}
   \mathbf{end}
    {\bf if} \ segundo\_maior\_tempo == 0
        diferenca_CH((iteracao+1)/2, 1) = 0;
    else
        diferenca_CH(((teracao+1)/2, 1) = (maior_tempo - segundo_maior_tempo)*1000;
   end
end
%title('Histograms');
subplot (3,2,5);
hist(diferenca_CH, x_CH);
ylabel('Samples');
xlabel('Time_{\Box}[ms]');
```

MATTERTANTATION FINITATION CHEETERTANTATION CHEETERTANTATION CHEETERTANTA

E.3 IB

title('CH');

%hist(diferenca_RS);

```
%Trata os dados do cadence (peaks) para transforma-los num
%readble input para o histogram
delta_RS = 20e-3;
x_RS = 0:2:50;
delta_FS = 1e-3;
x_FS = 0:2:100;
delta_IB = 10e-3;
x_IB = 0:2:65;
```

```
delta LTS = 5e-3;
x_LTS = 0:2:70;
delta_CH = 15e-3:
x_CH = 0:2:70;
delta_CH_short = 0.3e-3;
x_CHLong = 0:2:35;
MAARKANTAU COUTUURAARKANTAU COUTUURAARKANTAU COUTUURAARKANTAU COUTUURAARKANTAU COUTUURAARKANTA
dataIB = csvread('All_Variant_100-runs/IB.csv',1,0);
[num_linhas num_colunas] = size(dataIB);
diferenca_IB = \mathbf{zeros}(\operatorname{num\_colunas}/2, 1);
for iteracao=1:2:num_colunas %Varre todas as 100 iteracoes
    maior pico = 0;
    maior_tempo=0;
    ponteiro_tempo = num_linhas;
    indice\_tempo=0;
    indice_tempo1 = 0;
    indice_tempo2 = 0;
    indice_tempo3 = 0;
    indice_menor_tempo = 0;
    segundo_maior_tempo = 0;
    %Verifica o maior valor de pico da leitura
    for picos = 1:1:num_ linhas
        if dataIB(picos, iteracao+1) > maior_pico
                maior_pico = dataIB(picos, iteracao+1);
        \mathbf{end}
    end
    \%Verifica o maior tempo dos picos
    {\bf for} \ indice\_tempo=ponteiro\_tempo:-1:1
        if ( dataIB(indice_tempo,iteracao+1) >= maior_pico*0.9
                && dataIB(indice_tempo,iteracao+1) <= maior_pico*1.1)
            maior_tempo = dataIB(indice_tempo, iteracao);
            break:
        end
    end
    % Verifica o segundo maior tempo evitando os picos intermediarios
    %proximos ao pico principal
    controle = maior_tempo - delta_IB;
    %Acha o indice da linha que contem o tempo do tempo maximo - 1.5ms
    for indice_tempo2=indice_tempo:-1:1
        if (dataIB(indice_tempo2, iteracao) < controle)
            indice_menor_tempo = indice_tempo2;
            break:
        end
    end
    %Verifica o segundo maior tempo dos picos
    {\bf for} \ indice\_tempo3=indice\_menor\_tempo:-1:1
        if ( dataIB(indice_tempo, iteracao+1) >= maior_pico*0.9
```

```
&& dataIB(indice_tempo,iteracao+1) <= maior_pico*1.1)
           segundo_maior_tempo = dataIB(indice_tempo3, iteracao);
           break;
       else
           segundo_maior_tempo = 0;
       end
   \mathbf{end}
   if segundo_maior_tempo == 0
       diferenca_IB((iteracao+1)/2, 1) = 0;
   else
       diferenca_IB((iteracao+1)/2 ,1) = (maior_tempo - segundo_maior_tempo)*1000;
   \mathbf{end}
\mathbf{end}
%title('Histograms');
subplot (3,2,3);
hist(diferenca_IB, x_IB);
ylabel('Samples');
xlabel('Time_{\Box}[ms]');
title('IB');
%hist(diferenca_RS);
がとしてもちたたたとしてもしてもちたたたとなら IBさともただとうとしてもちたたたたたただとしていたもたたたたたただとうとしてもたたたたた
```

E.4 FS

% Trata os dados do cadence (peaks) para transforma-los num % readble input para o histogram delta RS = 20e-3; $x_RS = 0:2:50;$ delta_FS = 1e-3; $x_FS = 0:2:100;$ delta_IB = 10e-3; $x_{IB} = 0:2:65;$ delta_LTS = 5e-3; $x_LTS = 0:2:70;$ $delta_CH = 15e-3;$ $x_CH = 0:2:70;$ $delta_CH_short = 0.3e-3;$ $x_CHLong = 0:2:35;$ dataFS = csvread('All_Variant_100-runs/FS_20p.csv',1,0); [num_linhas num_colunas] = **size**(dataFS); diferenca_FS = $\mathbf{zeros}(\operatorname{num_colunas}/2, 1);$

```
for iteracao=1:2:num_colunas %Varre todas as 100 iteracoes
    maior_pico = 0;
    maior_tempo=0;
    ponteiro_tempo = num_linhas;
    indice_tempo=0;
    indice_tempo1 = 0;
    indice_tempo2 = 0;
    indice_tempo3 = 0;
    indice_menor_tempo = 0;
    segundo\_maior\_tempo = 0;
    %Verifica o maior valor de pico da leitura
    for picos = 1:1:num\_linhas
        if dataFS(picos, iteracao+1) > maior_pico
                maior_pico = dataFS(picos, iteracao + 1);
        end
    end
    % Verifica o maior tempo dos picos
    for indice_tempo=ponteiro_tempo:-1:1
        if ( dataFS(indice_tempo, iteracao+1) >= maior_pico*0.9
                && dataFS(indice_tempo,iteracao+1) <= maior_pico*1.1)
            maior_tempo = dataFS(indice_tempo,iteracao);
            break;
        \mathbf{end}
    end
    \%Verifica o segundo maior tempo evitando os picos intermediarios
    %proximos ao pico principal
    controle = maior_tempo - delta_FS;
    %Acha o indice da linha que contem o tempo do tempo maximo - 1.5ms
    for indice_tempo2=indice_tempo:-1:1
        if (dataFS(indice_tempo2,iteracao) < controle)
            indice_menor_tempo = indice_tempo2;
            break;
        \mathbf{end}
    end
    \%Verifica o segundo maior tempo dos picos
    for indice_tempo3=indice_menor_tempo:-1:1
        if ( dataFS(indice_tempo,iteracao+1) >= maior_pico*0.9
                && dataFS(indice_tempo, iteracao+1) <= maior_pico*1.1)
            segundo_maior_tempo = dataFS(indice_tempo3, iteracao);
            break;
        else
            segundo_maior_tempo = 0;
        end
    end
    if segundo_maior_tempo == 0
        diferenca_FS((iteracao+1)/2, 1) = 0;
    else
        diferenca_FS((iteracao+1)/2, 1) = (maior_tempo - segundo_maior_tempo)*1000;
    \mathbf{end}
end
```

```
subplot(3,2,2);
hist(diferenca_FS, x_FS);
ylabel('Samples');
xlabel('Time_[ms]');
title('FS');
%hist(diferenca_RS);
```

E.5 LTS

```
%Trata os dados do cadence (peaks) para transforma-los num
\ensuremath{\mathscr{D}}\xspace{readble} input para o histogram
delta_{RS} = 20e - 3;
x_RS = 0:2:50;
delta_FS = 1e-3;
x_FS = 0:2:100;
delta_IB = 10e - 3;
x_{IB} = 0:2:65;
delta_LTS = 5e-3;
x\_LTS = 0:2:70;
delta_CH = 15e - 3;
x_CH = 0:2:70;
delta_CH_short = 0.3e-3;
x\_CHLong = 0:2:35;
dataLTS = csvread('All_Variant_100-runs/LTS.csv',1,0);
[num_linhas num_colunas] = size(dataLTS);
diferenca_LTS = \mathbf{zeros}(\operatorname{num\_colunas}/2, 1);
for iteracao=1:2:num_colunas %Varre todas as 100 iteracoes
   maior_pico = 0;
   maior_tempo=0;
   ponteiro_tempo = num_linhas;
   indice_tempo=0;
   indice_tempo1 = 0;
   indice_tempo2 = 0;
   indice_tempo3 = 0;
   indice_menor_tempo = 0;
   segundo_maior_tempo = 0;
   %Verifica o maior valor de pico da leitura
   for picos=1:1:num_linhas
       if dataLTS(picos, iteracao+1) > maior_pico
               maior_pico = dataLTS(picos,iteracao+1);
       end
```

```
\mathbf{end}
   % Verifica o maior tempo dos picos
    for indice_tempo=ponteiro_tempo:-1:1
        if ( dataLTS(indice_tempo, iteracao+1) >= maior_pico*0.9
               && dataLTS(indice_tempo,iteracao+1) <= maior_pico*1.1)
            maior_tempo = dataLTS(indice_tempo, iteracao);
            break;
        \mathbf{end}
   end
   \% Verifica o segundo maior tempo evitando os picos intermediarios
   %proximos ao pico principal
    controle = maior_tempo - delta_LTS;
   %Acha o indice da linha que contem o tempo do tempo maximo - 1.5ms
   for indice_tempo2=indice_tempo:-1:1
        if (dataLTS(indice_tempo2, iteracao) < controle)
           indice_menor_tempo = indice_tempo2;
           break;
        end
   end
   \%Verifica o segundo maior tempo dos picos
    for indice_tempo3=indice_menor_tempo:-1:1
        if ( dataLTS(indice_tempo,iteracao+1) >= maior_pico*0.9
               && dataLTS(indice_tempo,iteracao+1) <= maior_pico*1.1)
            segundo_maior_tempo = dataLTS(indice_tempo3, iteracao);
           break;
        else
           segundo_maior_tempo = 0;
        end
   end
    if segundo_maior_tempo == 0
        diferenca_LTS(((iteracao+1)/2, 1) = 0;
    else
        diferenca_LTS ((iteracao+1)/2 ,1) = (maior_tempo - segundo_maior_tempo)*1000;
   \mathbf{end}
end
%title('Histograms');
subplot (3,2,4);
hist(diferenca_LTS, x_LTS);
ylabel('Samples');
xlabel('Time_{\Box}[ms]');
title('LTS');
%hist(differenca_RS);
```