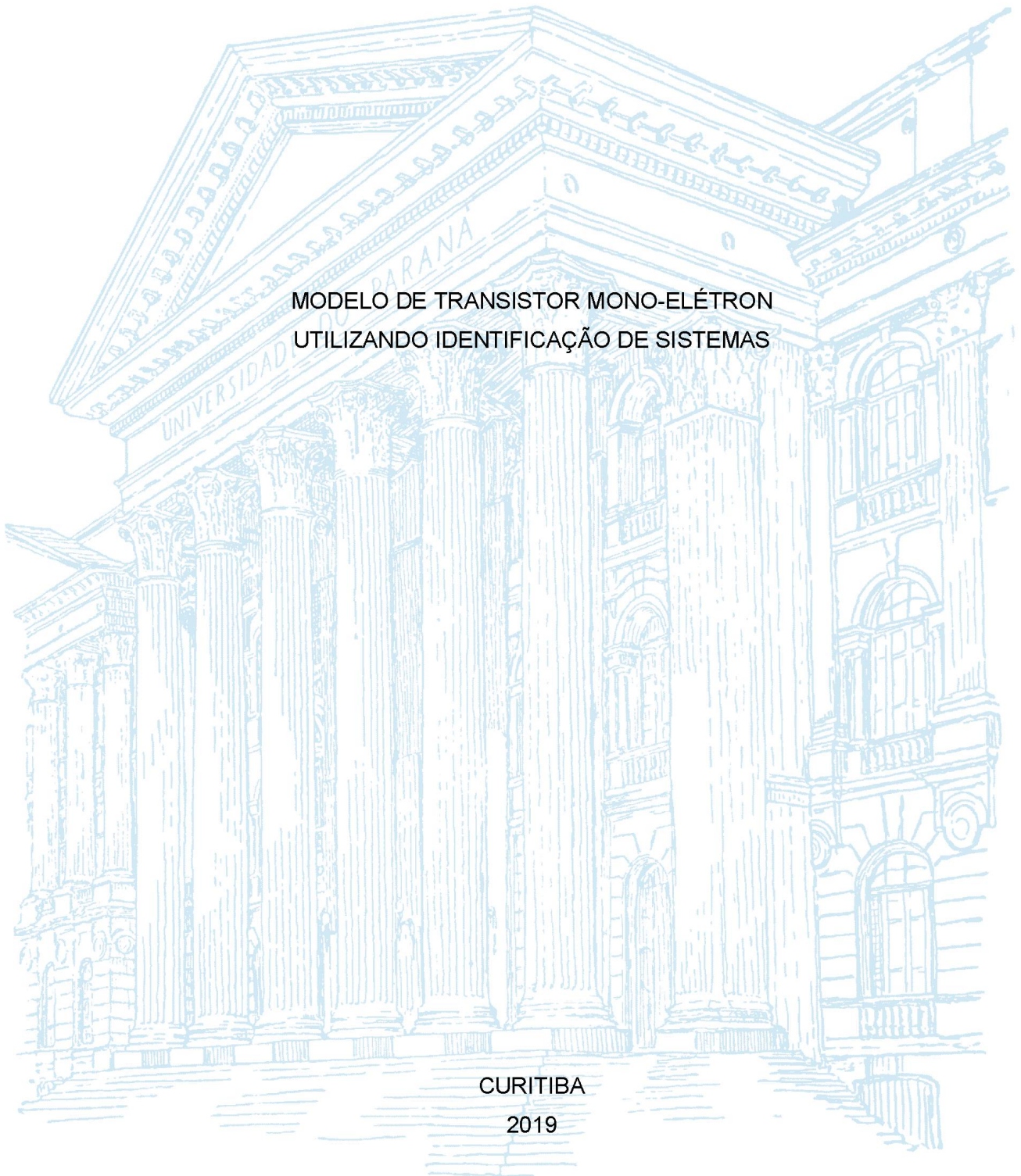


UNIVERSIDADE FEDERAL DO PARANÁ

BEATRIZ DOS SANTOS PÊS

MODELO DE TRANSISTOR MONO-ELÉTRON  
UTILIZANDO IDENTIFICAÇÃO DE SISTEMAS



CURITIBA

2019

BEATRIZ DOS SANTOS PÊS

MODELO DE TRANSISTOR MONO-ELÉTRON  
UTILIZANDO IDENTIFICAÇÃO DE SISTEMAS

Tese apresentada ao curso de Pós-Graduação em Engenharia Elétrica, Setor de Tecnologia, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Engenharia Elétrica – Sistemas Eletrônicos.

Orientador: Prof. Dr. Marlio José do Couto Bonfim

Coorientadora: Profa. Dra. Janaina Gonçalves Guimarães

CURITIBA

2019

Catálogo na Fonte: Sistema de Bibliotecas, UFPR  
Biblioteca de Ciência e Tecnologia

---

- P472m Pês, Beatriz dos Santos  
Modelo de transistor mono-elétron utilizando identificação de sistemas [recurso eletrônico] / Beatriz dos Santos Pês – Curitiba, 2019.
- Tese - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-graduação em Engenharia Elétrica.  
Orientador: Prof. Dr. Marlio José do Couto Bonfim  
Coorientadora: Profa. Dra. Janaina Gonçalves Guimarães
1. Nanoeletrônica. 2. Modelos de Hammerstein-Wiener 3. Identificação de Sistemas. I. Universidade Federal do Paraná. II. Bonfim, Marlio José do Couto. III. Guimarães, Janaina Gonçalves. IV. Título.

CDD: 537.5

---

Bibliotecária: Roseny Rivelini Morciani CRB-9/1585



## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA ELÉTRICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **BEATRIZ DOS SANTOS PÊS** intitulada: **MODELO DE TRANSISTOR MONO-ELÉTRON UTILIZANDO IDENTIFICAÇÃO DE SISTEMAS**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua Aprovação no rito de defesa. A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 05 de Julho de 2019.

MARLIO JOSÉ DO COUTO BONFIM  
Presidente da Banca Examinadora (UFPR)

MARLUS KOEHLER  
Avaliador Externo (UFPR)

GIDEON VILLAR LEANDRO  
Avaliador Interno (UFPR)

VICTOR BAPTISTA FRENCL  
Avaliador Externo (UFPR)

Dedico este trabalho aos meus pais, Altair e Jussara Pês. Meus maiores apoiadores, defensores e incentivadores desde sempre.

## AGRADECIMENTOS

Gostaria de agradecer a todos que, de alguma forma, contribuíram para que esta jornada fosse mais suave.

Ao amigo e parceiro de jornada acadêmica, Elder Oroski. Obrigada pelas discussões e ensinamentos.

A minha família, minha amada família, que compreendeu todas as vezes em que não estive presente e que me apoiou e me amou em todos os momentos.

A minha co-orientadora, profa. Janaina Guimarães, minha grande inspiração, de quem recebi meu diploma e que me acompanha nesta louca jornada acadêmica desde a graduação.

Ao meu orientador, prof. Marlio José do Couto Bonfim, pela sua paciência e serenidade. Por me transmitir tranquilidade em meio a tanta insegurança e ansiedade.

*Happiness is only real when shared.*

Chris MacCandless.

*A mind full of questions and a teacher in my soul, and so it goes.*

(Eddie Vedder, Guaranteed, Into the Wild)

## RESUMO

Diante da necessidade de dispositivos eletrônicos cada vez menores e com maior capacidade de processamento de dados, a nanoeletrônica surge como uma alternativa aos dispositivos tradicionais. Dentre os dispositivos mono-elétron já propostos, o transistor mono-elétron, SET, ocupa lugar de destaque. Entre os principais desafios que dificultam a ampla aplicação do SET na indústria eletrônica encontram-se a falta de técnicas de fabricação consolidadas e a dificuldade que estes dispositivos apresentam para operar em temperatura ambiente. Para que estes e outros desafios sejam vencidos, são necessárias ferramentas de simulação adequadas e modelos precisos. Sendo assim, o objetivo deste trabalho consiste em aplicar conhecimentos relativos à Identificação de Sistemas para efetuar a modelagem caixa preta do transistor mono-elétron. Modelos caixa preta baseiam-se exclusivamente em dados de entrada e saída coletados do sistema a ser modelado. Estes modelos são conhecidos por se adequar com bastante precisão à dinâmica do sistema. Modelos já propostos para o SET baseiam-se em modelagem analítica e requerem conhecimento prévio sobre o sistema. Neste trabalho, foram coletados dados de entrada e saída do transistor mono-elétron por meio de simulações efetuadas com o SIMON, uma das plataformas de simulação de circuitos nanoeletrônicos mais difundidas na literatura. Aplicou-se um sinal persistentemente excitante como tensão de porta do transistor e coletou-se a corrente entre dreno e fonte. De posse destes dados, foram testadas as seguintes classes de modelos: NARX, Hammerstein, Wiener e Hammerstein-Wiener. Como métricas para realizar a comparação entre essas classes, foram utilizados o Erro Médio Quadrático e o tempo médio de simulação, computado para 50 execuções do experimento de validação dos modelos. Além disso, a simplicidade da estrutura matemática foi determinante na escolha da classe de modelos. A classe de Hammerstein-Wiener obteve o melhor desempenho. Deste modo, foram construídos modelos de Hammerstein-Wiener para o SET operando em diferentes condições. Os resultados comprovam a adequação desta representação matemática para a modelagem do transistor.

Palavras-chave: Modelagem Caixa Preta. SET. Eletrônica Mono-Elétron. Identificação de Sistemas. Modelos de Hammerstein-Wiener.

## ABSTRACT

Given the need for smaller electronic devices with great data processing capability, nanoelectronics appears as an alternative to traditional devices. Among the single-electron devices already proposed, the Single-Electron Transistor, SET, has a prominent place. Among the main challenges that prevent SET from being widely applied in the electronics industry are the lack of consolidated manufacturing techniques and the difficulty these devices present to operate at room temperature. In order to overcome these and other challenges, appropriate simulation tools and precise models are needed. Thus, the goal of this work is the black box modeling of the SET. Black box models are based exclusively on input and output data collected from the system being modeled. These models are known to fit very accurately to the dynamics of the system. Models already proposed for SET are based on analytical modeling and require prior knowledge about the system. In this work, the input and output data of the single-electron transistor were collected through simulations performed with SIMON, one of the most widespread nanoelectronic circuit simulation platforms. A persistently exciting signal was applied as the gate voltage of the transistor and the current was collected between drain and source. With these data, the following classes of models were tested: NARX, Hammerstein, Wiener and Hammerstein-Wiener. As metrics to perform the comparison between these classes of models, the MSE and the mean simulation time, computed for 50 runs of the validation experiment of the models, were used. In addition, the simplicity of the mathematical structure of the model was decisive in the choice of the mathematic representation. The Hammerstein-Wiener class achieved the best performance. In this way, Hammerstein-Wiener models were constructed for SET operating under different conditions. The results prove the adequacy of this mathematical representation for modeling the transistor.

Keywords: Black Box Modeling. SET. Single-Electron Electronics. System Identification. Hammerstein-Wiener Models.

## LISTA DE FIGURAS

- FIGURA 1 Princípio do controle individual de elétrons: uma ilha condutora (a) antes e (b) depois da adição de um único elétron. Um elétron em excesso cria um campo elétrico, que impede a adição de mais elétrons à ilha - modificada de (LIKHAREV, 1999). ..... 33
- FIGURA 2 Esquemático e símbolo de uma junção-túnel - modificada de (GUIMARÃES, 2005). ..... 34
- FIGURA 3 Circuito equivalente da junção-túnel. .... 34
- FIGURA 4 Uma região tridimensional delimitada por  $L_x$ ,  $L_y$  e  $L_z$ ,  $\lambda_e$  é o comprimento de onda do elétron. Desta forma, o elétron será livre se  $L_{x,y,z} \gg \lambda_e$  - modificada de (HANSON, 2008). .... 35
- FIGURA 5 Ilha: região constituída de material condutor cercado por barreiras de material isolante, (a) esquemático e (b) ilha isolada por duas junções-túnel - modificada de (HANSON, 2008). .... 36
- FIGURA 6 Tunelamento de uma partícula de energia  $E$  através de uma barreira de potencial  $V_0 > E$ . Os pontos A e B são os pontos críticos clássicos - modificada de (FERRY, 1991). .... 37
- FIGURA 7 Característica I-V com bloqueio de Coulomb evidente - modificada de (HANSON, 2008). .... 41
- FIGURA 8 Transistor mono-elétron.  $V_{DS}$  é a tensão de polarização do transistor,  $V_S$  e  $V_D$  são as tensões nas junções-túnel dos terminais de fonte

e dreno, respectivamente. $C_g$ é o valor da capacitância do capacitor no terminal de porta, $V_{gg}$ é a queda de tensão neste capacitor e $V_g$ é a tensão de controle aplicada ao terminal de porta - modificada de (HANSON, 2008).	43
<b>FIGURA 9</b> SET com $V_g > 0$ . A tensão deprime o <i>gap</i> gerado pelo bloqueio de Coulomb, permitindo, assim, que um elétron entre na ilha. $E_F$ indica a energia do nível de Fermi - modificada de (HANSON, 2008).	44
<b>FIGURA 10</b> Característica $I_{DS}$ - $V_g$ do SET.	45
<b>FIGURA 11</b> Diagrama de um sistema LIT representado por $H(z)$ . As variáveis $u(k)$ e $y(k)$ representam a entrada e a saída, respectivamente.	48
<b>FIGURA 12</b> Teste de homogeneidade: a amplitude da resposta do SET se mantém inalterada, portanto, o transistor não obedece ao princípio da homogeneidade.	49
<b>FIGURA 13</b> Teste de aditividade: $I_{ds1}$ é a resposta do transistor à entrada $V_{gs1}$ e $I_{ds2}$ é a resposta do transistor à entrada $V_{gs2}$ . $I_{ds3}$ é a resposta do transistor à entrada $V_{gs3} = V_{gs1} + V_{gs2}$ .	50
<b>FIGURA 14</b> Teste de aditividade: $I_{ds4} = I_{ds1} + I_{ds2}$ e $I_{ds3} \neq I_{ds4}$ , portanto, o SET não respeita a propriedade da aditividade.	50
<b>FIGURA 15</b> Invariância no tempo: $I_{ds1}$ é a resposta do transistor à $V_{gs1}$ . $I_{ds2}$ é a resposta do transistor à entrada $V_{gs2} = V_{gs1}(t + 250 \text{ ns})$ .	51
<b>FIGURA 16</b> Sinais de teste comumente utilizados em Identificação de Sistemas: sinal PRBS e sinal PRMLS.	55

FIGURA 17 Fluxograma ilustrando os principais passos a serem seguidos na modelagem caixa preta. ....	59
FIGURA 18 Esquemático de um sistema representado por um modelo de Volterra (ROSA, 2009). ....	60
FIGURA 19 Modelo de Hammerstein. Os sinais $u(k)$ , $y(k)$ e $v(k)$ representam a entrada, a saída e o sinal interno, respectivamente. O sinal $v(k)$ é, em geral, desconhecido e sua medição não é possível. ....	63
FIGURA 20 Modelo de Wiener. Os sinais $u(k)$ , $y(k)$ e $v(k)$ representam a entrada, a saída e o sinal interno, respectivamente. O sinal $v(k)$ é, em geral, desconhecido e sua medição não é possível. ....	64
FIGURA 21 Modelo de Hammerstein-Wiener. Os sinais $u(k)$ , $y(k)$ , $v(k)$ e $w(k)$ representam a entrada, a saída e os sinais internos, respectivamente. Os sinais $v(k)$ e $w(k)$ são, em geral, desconhecidos e sua medição não é possível. ....	65
FIGURA 22 Exemplo de característica PWL com $(2 + 2)$ -segmentos, modificada de (VÖRÖS, 2007). ....	66
FIGURA 23 Ilustração das transições possíveis em processo <i>birth-death</i> com número generalizado de estados - modificada de (WASSHUBER, 2001), em que $n$ é o número de elétrons na ilha. ....	71
FIGURA 24 Taxas de tunelamento em uma junção $X$ de interesse. ....	72
FIGURA 25 Fluxograma apresentando os passos para a obtenção da corrente elétrica a partir da EM - modificada de (HASANEEN; WAHAB; AHMED, 2011). ....	73

FIGURA 26 Fluxograma para o método de MC - modificada de (WASSHUBER, 2001). .....	75
FIGURA 27 A carga elementar, $e$ , é espalhada por vários eventos de tunelamento, de acordo com sua probabilidade de ocorrência - modificada de (WASSHUBER, 2001). .....	76
FIGURA 28 Diagrama exibindo as taxas de tunelamento para um elétron através do SET. As letras S, G e D marcam os terminais de fonte, porta e dreno, respectivamente - modificada de (HASANEEN; WAHAB; AHMED, 2011). .....	77
FIGURA 29 Diagrama de um inversor mono-elétron com dois SETs em série. $V_{in}$ é a tensão de entrada, $V_{out}$ é a tensão de saída, $V_B$ é a tensão de polarização, $C_n, R_n$ , para $n = 1, 2, 3, 4$ , são os parâmetros das junções-túnel e $C_L$ é a capacitância de interconexão - modificada de (YU; HWANG; AHN, 1999). .....	80
FIGURA 30 Circuito equivalente para o SET proposto por Yu <i>et al.</i> Os terminais marcados com G, S e D correspondem aos terminais de <i>gate</i> , fonte e dreno, respectivamente - modificada de (YU; HWANG; AHN, 1999). ....	84
FIGURA 31 Circuito equivalente para o SET proposto por Wu e Lin. Os terminais marcados com G, S e D correspondem aos terminais de <i>gate</i> , fonte e dreno, respectivamente - modificada de (WU; LIN, 2003). .....	85
FIGURA 32 Circuito equivalente proposto por Yu <i>et al.</i> Os terminais marcados com G, BG, S e D correspondem aos terminais de <i>gate</i> , <i>backgate</i> , fonte e dreno, respectivamente - modificada de (YU; HWANG; AHN, 2005). ..	87
FIGURA 33 Circuito equivalente proposto por Karimian <i>et al.</i> Os terminais marcados com G, S e D correspondem aos terminais de <i>gate</i> , fonte e dreno, respectivamente - modificada de (KARIMIAN et al., 2009). .....	88

FIGURA 34	Sinais de <i>clock</i> utilizados no modelo de Karimian <i>et al.</i> . Os sinais não possuem sobreposição, cada sinal controla duas chaves de forma simultânea. $T$ é o período dos sinais - modificada de (KARIMIAN <i>et al.</i> , 2009).	88
FIGURA 35	Circuito equivalente proposto por Jain <i>et al.</i> . Os terminais marcados com G, S e D correspondem aos terminais de <i>gate</i> , fonte e dreno, respectivamente. $g_1$ é a fonte de corrente controlada por tensão e $R_3$ foi incluído no modelo para facilitar a medição da corrente no dreno - modificada de (JAIN <i>et al.</i> , 2015).	90
FIGURA 36	Sinal de entrada aplicado ao transistor mono-elétron na etapa de estimação.	116
FIGURA 37	Esquemático da simulação de coleta de dados para a modelagem caixa preta do SET. $V_{GS}$ é o sinal mostrado na Figura 36. O transistor é simétrico, ou seja, as junções-túnel possuem os mesmos parâmetros, $C_j = C_g = 1$ aF e $R_T = 100$ k $\Omega$ . $V_{DS} = 20$ mV é a tensão de polarização do transistor. A saída é a corrente entre dreno e fonte, medida pelo amperímetro.	117
FIGURA 38	Corrente entre dreno e fonte, $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 3 quando a tensão $V_{GS}$ mostrada na Figura 36 é aplicada à sua entrada.	118
FIGURA 39	Esquemático detalhando os passos e métricas que guiaram a escolha da representação matemática para a modelagem caixa preta do SET.	119
FIGURA 40	Sinal de entrada aplicado ao transistor mono-elétron na etapa de validação.	125

FIGURA 41	Corrente entre dreno e fonte, $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 3 quando a tensão $V_{GS}$ mostrada na Figura 40 é aplicada à sua entrada. ....	126
FIGURA 42	Corrente entre dreno e fonte, $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão $V_{GS}$ mostrada na Figura 40 é aplicada à sua entrada a $T = 77$ K. ....	127
FIGURA 43	Corrente entre dreno e fonte, $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão $V_{GS}$ mostrada na Figura 40 é aplicada à sua entrada com $T = 150$ K e $V_{DS} = 20$ mV. ....	128
FIGURA 44	Corrente entre dreno e fonte, $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão $V_{GS}$ mostrada na Figura 40 é aplicada à sua entrada com $T = 150$ K e $V_{DS} = 100$ mV. ....	128
FIGURA 45	Corrente entre dreno e fonte, $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão $V_{GS}$ mostrada na Figura 40 é aplicada à sua entrada com $T = 150$ K e $V_{DS} = 180$ mV. ....	129
FIGURA 46	Corrente entre dreno e fonte, $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão $V_{GS}$ mostrada na Figura 40 é aplicada à sua entrada com $T = 298$ K e $V_{DS} = 20$ mV. ....	129
FIGURA 47	Corrente entre dreno e fonte, $I_{DS}$ , para os diferentes valores de temperatura. ....	130
FIGURA 48	Corrente entre dreno e fonte, $I_{DS}$ , para os diferentes valores de tensão de polarização. ....	130

FIGURA 49	Não linearidades estáticas dos modelos de blocos concatenados. a) PWL de entrada para o modelo de Hammerstein. b) PWL de saída para o modelo de Wiener.c) e (d) PWLs de entrada e de saída, para o modelo de Hammerstein-Wiener, respectivamente. ....	135
FIGURA 50	Resposta dos modelos NARX, de Hammerstein, de Wiener e de Hammerstein-Wiener com estruturas explicitadas na Tabela 8 e nas equações (7.2) e (7.3) e parâmetros expostos na Tabela 9. O transistor mono-elétron é simétrico e possui os seguintes parâmetros: $C_j = C_g = 1$ aF, $R_T = 100$ k $\Omega$ e $V_{DS} = 20$ mV. A simulação foi realizada a $T = 0$ K. ....	136
FIGURA 51	Gráfico tempo médio de simulação $\times$ MSE para os modelos orientados a blocos. ....	137
FIGURA 52	Não linearidades estáticas para os modelos de Hammerstein-Wiener em diferentes condições de operação. (a) PWL de entrada e (b) de saída para $T = 77$ K e $V_{DS} = 20$ mV. (c) PWL de entrada e (d) de saída para $T = 150$ K e $V_{DS} = 20$ mV. ....	140
FIGURA 53	Não linearidades estáticas para os modelos de Hammerstein-Wiener em diferentes condições de operação. (a) PWL de entrada e (b) de saída para $T = 150$ K e $V_{DS} = 100$ mV. (c) PWL de entrada e (d) de saída para $T = 150$ K e $V_{DS} = 180$ mV. ....	141
FIGURA 54	Não linearidades estáticas para os modelos de Hammerstein-Wiener em diferentes condições de operação. (a) PWL de entrada e (b) de saída para $T = 298$ K e $V_{DS} = 20$ mV ....	142
FIGURA 55	Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.4) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes	

parâmetros:  $C_j = C_g = 1$  aF,  $R_T = 100$  k $\Omega$  e  $V_{DS} = 20$  mV. A simulação foi realizada a  $T = 77$  K. O MSE é de  $3,47 \times 10^{-18} A^2$ . ..... 142

FIGURA 56 Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 10$  M $\Omega$  e  $V_{DS} = 20$  mV. A simulação foi realizada a  $T = 150$  K. O MSE é de  $6,79 \times 10^{-23} A^2$ . .. 143

FIGURA 57 Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 10$  M $\Omega$  e  $V_{DS} = 100$  mV. A simulação foi realizada a  $T = 150$  K. O MSE é de  $1,81 \times 10^{-21} A^2$ . .. 143

FIGURA 58 Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 10$  M $\Omega$  e  $V_{DS} = 180$  mV. A simulação foi realizada a  $T = 150$  K. O MSE é de  $2,14 \times 10^{-21} A^2$ . .. 144

FIGURA 59 Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 100$  M $\Omega$  e  $V_{DS} = 20$  mV. A simulação foi realizada a  $T = 298$  K. O MSE é de  $1,82 \times 10^{-24} A^2$ . .. 144

## LISTA DE TABELAS

TABELA 1	Valores dos parâmetros do circuito inversor. ....	83
TABELA 2	Valores dos parâmetros de ajuste de curva utilizados por Yu <i>et al.</i>	85
TABELA 3	Parâmetros do transistor mono-elétron utilizados na coleta de dados. ....	117
TABELA 4	Restrições aplicadas à construção do espaço de regressores candidatos para o modelo NARX. $n_u$ e $n_y$ representam as regressões máximas de entrada e de saída, respectivamente. ....	120
TABELA 5	Matrizes de Regressão para o espaço de regressores candidatos.	120
TABELA 6	Características a serem determinadas na estrutura de modelos de blocos concatenados. $L_{in}$ e $L_{out}$ representam o número de segmentos lineares nas não linearidades de entrada e de saída, respectivamente. $n_a$ , $n_b$ e $n_k$ são a quantidade de polos, zeros e atrasos puros, respectivamente. ....	123
TABELA 7	Configurações utilizadas para a coleta de dados na construção de modelos em diferentes temperaturas de operação e com tensões de polarização variadas. ....	131
TABELA 8	Estrutura dos modelos orientados a blocos para o transistor simétrico com $C_g = C_j = 1$ aF, $R_T = 100$ k $\Omega$ , $V_{DS} = 20$ mV e $T = 0$ K. $L_{in}$ e $L_{out}$ representam o número de segmentos lineares nas não linearidades de	

	entrada e de saída, respectivamente. $n_a$ , $n_b$ e $n_k$ são a quantidade de zeros, polos e atrasos puros, respectivamente. ....	133
TABELA 9	Parâmetros dos blocos dinâmicos lineares para os modelos de Hammerstein, de Wiener e de Hammerstein-Wiener. ....	134
TABELA 10	Desempenho dos modelos com relação ao MSE e ao tempo médio de simulação. ....	137
TABELA 11	Estrutura dos modelos de Hammerstein-Wiener para transistor com parâmetros expostos na Tabela 7. $L_{in}$ e $L_{out}$ representam o número de segmentos lineares nas não linearidades de entrada e de saída, respectivamente. $n_a$ , $n_b$ e $n_k$ são a quantidade de polos, zeros e atrasos puros. ....	138
TABELA 12	Parâmetros dos blocos dinâmicos lineares para os modelos de Hammerstein, de Wiener e de Hammerstein-Wiener em diferentes condições de simulação. ....	139
TABELA 13	MSE obtido pelos modelos de Hammerstein-Wiener em diferentes condições de simulação. ....	139

## LISTA DE SIGLAS

MOS	<i>Metal-Oxide Semiconductor</i>
SET	<i>Single-Electron Transistor</i>
MC	Monte Carlo
EM	Equação Mestra
SPICE	<i>Simulated Program with Integrated Circuits Emphasis</i>
NARX	<i>Nonlinear Auto Regressive with eXogenous input</i>
MSE	<i>Mean Squared Error</i>
FinFET	<i>Fin Field Effect Transistor</i>
VLSI	<i>Very Large Scale Integration</i>
ULSI	<i>Ultra Large Scale Integration</i>
STM	<i>Scanning Tunneling Microscope</i>
SLIT	Sistema Linear e Invariante no Tempo
SISO	<i>Single-Input-Single-Output</i>
PE	Persistência de Excitação ou Persistentemente Excitante
PRBS	<i>Pseudo Random Binary Signal</i>
PRMLS	<i>Pseudo Random Multi Level Signal</i>
MMQ	Método dos Mínimos Quadrados
NARMAX	Nonlinear Auto-Regressive Moving Average with eXogenous input
PWL	<i>PieceWise Linear</i>
ARX	<i>AutoRegressive with eXogenous input</i>
SIMON	<i>Simulation of Nanostructures</i>
KOSEC	<i>Korea Single Electron Circuit Simulator for Spice</i>
SENECA	<i>Single Electron Nanoelectronic Circuit Analyzer</i>
MOSES	<i>Monte Carlo Single Electron Simulator</i>

SECS	<i>Single Electron Circuit Simulator</i>
SEMSIM	<i>Single Electron Multiscale Simulation</i>
CAMSET	<i>Circuit Analysis program including a Model of Single Electron Tunneling</i>
LM	Levenberg-Marquardt
MQ	Mínimos Quadrados
VA	Varável Aleatória
MQO	Método dos mínimos Quadrados Ortogonal
ERR	<i>Error Reduction Ratio</i>
PEM	<i>Prediction Error Methods</i>

## LISTA DE SÍMBOLOS

$\epsilon$	Campo elétrico gerado por uma carga elementar
$e$	Carga elementar, aproximadamente igual a $1,6 \times 10^{-19}$ C
$R_T$	Resistência de tunelamento da junção-túnel
$C_j$	Capacitância da junção-túnel
$E$	Energia da partícula
$p$	Módulo do momento linear
$\lambda$	Comprimento de onda
$h$	Constante de Planck
$\Phi$	Função de onda
$\tau_t$	Tempo entre dois tunelamentos sucessivos
$R_Q$	Resistência quântica
$\Gamma$	Taxa de tunelamento
$\Delta F$	Varição da energia eletrostática livre do sistema
$Q$	Carga Elétrica
$\epsilon$	Permissividade elétrica.
$E_c$	Energia de Carregamento
$k_B$	Constante de Boltzmann
$T_s$	Período de amostragem.
$\Psi$	Matriz de Regressão.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>25</b>
1.1	OBJETIVOS.....	29
1.1.1	OBJETIVO GERAL.....	29
1.1.2	Objetivos Específicos.....	29
1.2	ORGANIZAÇÃO DO TRABALHO.....	30
<b>2</b>	<b>ELETRÔNICA MONO-ELÉTRON.....</b>	<b>31</b>
2.1	DISPOSITIVOS DE TUNELAMENTO MONO-ELÉTRON.....	32
2.1.1	Ilha.....	35
2.1.2	Tunelamento.....	36
2.1.3	Teoria Ortodoxa do Tunelamento Mono-Elétron.....	38
2.1.4	Efeito de Carregamento.....	39
2.1.5	SET.....	43
2.1.5.1	Equações Características.....	44
2.2	SET: OBJETO DE MODELAGEM.....	46
2.2.1	Linearidade.....	46
2.2.2	Invariância no tempo.....	47
2.2.3	SET: sistema não linear e invariante no tempo.....	48
2.3	CONSIDERAÇÕES FINAIS.....	51
<b>3</b>	<b>IDENTIFICAÇÃO DE SISTEMAS.....</b>	<b>52</b>
3.1	MODELAGEM CAIXA PRETA.....	53
3.2	COLETA DE DADOS.....	54

3.3	DETERMINAÇÃO DA ESTRUTURA DO MODELO.....	56
3.4	ESTIMAÇÃO DE PARÂMETROS.....	57
3.5	VALIDAÇÃO DO MODELO.....	58
3.6	REPRESENTAÇÕES NÃO LINEARES.....	59
3.6.1	Série de Volterra.....	60
3.6.2	NARMAX Polinomial.....	61
3.6.3	Modelos de Blocos Concatenados.....	62
3.6.3.1	Modelo de Hammerstein.....	63
3.6.3.2	Modelo de Wiener.....	64
3.6.3.3	Modelo de Hammerstein-Wiener.....	65
3.6.3.4	Não linearidade estática: <i>PieceWise-Linear</i> .....	66
3.7	CONSIDERAÇÕES FINAIS.....	67
<b>4</b>	<b>MODELAGEM CAIXA BRANCA DO SET.....</b>	<b>68</b>
4.1	MÉTODOS DE MODELAGEM E SIMULAÇÃO.....	69
4.1.1	Equação Mestra.....	69
4.1.2	Monte Carlo.....	73
4.2	MODELOS CAIXA BRANCA PARA O SET.....	76
4.2.1	Modelos Analíticos.....	76
4.2.2	Modelos Macro.....	82
4.3	CONSIDERAÇÕES FINAIS.....	91
<b>5</b>	<b>MÉTODOS APLICADOS.....</b>	<b>92</b>
5.1	CONCEITOS INICIAIS.....	92
5.2	MECANISMOS DE BUSCA.....	94
5.2.1	Busca em Linha.....	95
5.2.1.1	Busca Exata em Linha.....	96

5.2.1.2	<i>Backtracking</i> .....	96
5.2.2	Região de Confiança .....	96
5.3	MÉTODO DE NEWTON.....	97
5.4	MÉTODO DE GAUSS-NEWTON .....	99
5.5	LEVENBERG-MARQUARDT .....	100
5.6	MÉTODO DOS MÍNIMOS QUADRADOS.....	102
5.7	MÉTODO DOS MÍNIMOS QUADRADOS ORTOGONAL .....	108
5.8	TAXA DE REDUÇÃO DE ERRO .....	110
5.9	MÉTODOS DE PREDIÇÃO DE ERRO .....	111
5.10	ERRO MÉDIO QUADRÁTICO .....	113
5.11	CONSIDERAÇÕES FINAIS .....	113
<b>6</b>	<b>MODELAGEM CAIXA PRETA DO SET .....</b>	<b>115</b>
6.1	ESCOLHA DA REPRESENTAÇÃO MATEMÁTICA .....	115
6.1.1	Coleta de Dados .....	115
6.1.2	Determinação de Estrutura .....	118
6.1.3	Estimação de Parâmetros .....	123
6.1.4	Validação dos Modelos.....	124
6.2	MODELOS: DIFERENTES CONDIÇÕES DE OPERAÇÃO .....	126
6.3	CONSIDERAÇÕES FINAIS .....	131
<b>7</b>	<b>APRESENTAÇÃO E ANÁLISE DE RESULTADOS.....</b>	<b>132</b>
7.1	ESCOLHA DA REPRESENTAÇÃO MATEMÁTICA .....	132
7.2	MODELOS DE HAMMERSTEIN-WIENER PARA O SET .....	137
7.3	ANÁLISE DE RESULTADOS .....	145
7.4	CONSIDERAÇÕES FINAIS .....	149
<b>8</b>	<b>CONCLUSÕES.....</b>	<b>150</b>

8.1	Discussões .....	150
8.2	Propostas para Trabalhos Futuros .....	151
	<b>REFERÊNCIAS .....</b>	<b>153</b>
	<b>Anexo A - Equação Mestra .....</b>	<b>160</b>
A.1	Método de Hadley e Lientschnig (WASSHUBER, 2001) .....	160
A.2	Cálculo das Probabilidades de Ocupação dos Estados .....	162

# 1 INTRODUÇÃO

Em 1965, Gordon Moore, co-fundador da Intel, enunciou o que ficaria conhecido como a Lei de Moore: a taxa de miniaturização dos componentes eletrônicos dobraria a cada dezoito meses. Durante os primeiros anos após esta previsão, esforços conjuntos de engenheiros e físicos mantiveram a validade da Lei de Moore. No final da década de 90, entretanto, a redução da dimensão característica dos transistores começou a gerar problemas de desempenho.

Microprocessadores atuais possuem transistores com *gate* (terminal no qual aplica-se uma tensão de controle que determina a operação do transistor) em torno de 10 nm. Este avanço significativo se deve ao empenho de pesquisadores impelidos pela promessa de circuitos cada vez mais rápidos e eficientes. O resultado destes esforços são tecnologias inovadoras que mantiveram a validade da Lei de Moore por 50 anos, embora hoje a taxa de miniaturização dobre a cada 30 meses, aproximadamente (TOUMEY, 2016).

Apesar de todos os avanços obtidos até então, a tecnologia *Metal-Oxide-Semi-conductor*, MOS, eventualmente atingirá limites fundamentais e não fornecerá mais avanços em desempenho (GARNER et al., 2007). Devido a estes limites, o progresso demonstrado até então pode ser comprometido. Estudos exploratórios especulam que transistores MOS podem oferecer melhorias até a dimensão de 5 nm, abaixo da qual não mais apresentarão ganhos em desempenho e escala (HAIN, 2015). Conforme o limite dos dispositivos MOS se aproxima, uma nova tecnologia é necessária para levar a indústria eletrônica além deste limite (BOHR, 2002). A nanoeletrônica aparece como uma das alternativas possíveis.

A nanoeletrônica apresenta diversas vantagens que vão além da redução da área. Os dispositivos nanoeletrônicos já propostos têm seu comportamento regido pela física quântica (GOSER et al., 1997). Esta característica permite a condução de corrente através de tunelamento, um fenômeno quântico que possibilita o transporte de

elétrons de forma individual (LIKHAREV, 1999). Deste modo, vários pesquisadores destacam esta peculiaridade dos dispositivos nanoeletrônicos, referindo-se a eles como dispositivos mono-elétron. O transporte de carga via tunelamento fornece melhorias em diversos aspectos fundamentais, tais como dissipação de potência, velocidade de chaveamento e escala (PAUL, 2002).

Algumas características únicas do *Single-Electron Transistor*, SET, fazem dele um dos dispositivos mono-elétron mais estudados (HASANEEN; WAHAB; AHMED, 2011). Dentre estes atributos, pode-se citar a característica oscilatória da corrente de saída em função da tensão de *gate*. Entretanto, o SET ainda precisa superar vários desafios para se tornar o substituto do transistor MOS. Por ser extremamente sensível a variações de carga, o transistor mono-elétron sofre com os efeitos de impurezas, que podem gerar condução em regiões nas quais o transistor deveria estar em bloqueio (ABUTALEB, 2015). Além disso, o ganho de tensão do SET é baixo (LIENTSCHNIG; WEYMANN; HADLEY, 2003). Até pouco tempo atrás, poderia-se citar a operação em temperatura ambiente como outro desafio a ser superado pelos pesquisadores. Contudo, grupos de pesquisa, por meio de avanços nas técnicas de fabricação, conseguiram observar com sucesso as características de bloqueio e oscilatória do SET em temperaturas de até 350 K (SHIN et al., 2010, 2011; LAVIEVILLE et al., 2015).

Desta forma, apesar de promissora, a aplicação direta do SET em circuitos práticos demanda ainda muito esforço em pesquisa. Dentre as alternativas plausíveis que vêm sendo estudadas, destacam-se as implementações híbridas SET-MOS (ABUTALEB, 2015; JAIN et al., 2014; JANA et al., 2013). Fundamentais para tais estudos, sejam relativos a aplicações híbridas ou puras, são simuladores de circuitos e modelos eficientes.

Pode-se definir um sistema como uma estrutura na qual variáveis de diferentes tipos interagem e produzem sinais observáveis (AGUIRRE, 2015; LJUNG, 1999). Um modelo, então, é definido como uma relação entre quantidades observáveis, que permite a predição das propriedades e do comportamento de um objeto (LJUNG, 1999). Deste modo, a modelagem de um sistema pode prover um melhor entendimento de seu funcionamento, além de constituir uma poderosa ferramenta para a previsão de seu comportamento em diversas condições (ASTRÖM; HÄGGLUND, 2011).

Vários modelos para o transistor mono-elétron já foram propostos. Estes modelos diferenciam-se, principalmente, quanto ao método utilizado para sua elaboração. Eles podem ser desenvolvidos, por exemplo, através do método de Monte Carlo, MC,

que considera os tunelamentos possíveis, calcula suas probabilidades de ocorrência e as usa para determinar os eventos relevantes (WASSHUBER, 2001). Apesar de ser a forma mais precisa para simular circuitos mono-elétron, o método de Monte Carlo não é indicado para circuitos com um número grande de junções-túnel, devido ao alto tempo de execução da simulação (RADWAN et al., 2015). Além disso, plataformas que utilizam este método não são capazes de simular circuitos híbridos SET-MOS (JAIN et al., 2015).

A fim de obter modelos que possam ser simulados de forma mais rápida, vários pesquisadores propuseram modelos analíticos do SET, baseados na Equação Mestre, EM, do tunelamento mono-elétron (MAHAPATRA et al., 2004). A EM utiliza a teoria ortodoxa, explicitada na Seção 2.1.3, para computar a taxa de tunelamento de cada junção e as probabilidades de ocupação de cada estado (WASSHUBER, 2001). A solução da EM fornece as probabilidades de ocupação dos estados do SET, que são, posteriormente, aplicadas ao cálculo da corrente no transistor (RADWAN et al., 2015). Os modelos desenvolvidos através desta técnica são menos precisos que os simulados através de MC. Contudo, a técnica da EM permite ao projetista determinar qual a precisão necessária em seu modelo, por meio do cálculo do número de estados a ser considerado (WASSHUBER, 2001). Quanto mais estados forem considerados, mais preciso e mais lento será o modelo. Em 2003, Lientschnig e Hadley empregaram uma relação de recursão para determinar as probabilidades dos estados (LIENTSCHNIG; WEYMANN; HADLEY, 2003). Esta relação é amplamente utilizada na elaboração de modelos analíticos para o SET.

Há, ainda, a técnica de modelagem através de modelos macro em *Simulated Program with Integrated Circuits Emphasis*, SPICE. Neste tipo de modelo, o comportamento do transistor mono-elétron é emulado por circuitos equivalentes. Estes circuitos consistem de componentes microeletrônicos convencionais, tais como fontes de corrente e tensão, diodos e resistores. O aspecto mais vantajoso desta aproximação é sua compatibilidade com qualquer plataforma capaz de simular circuitos eletrônicos convencionais (JAIN et al., 2015).

Desta forma, pode-se dizer que os modelos já propostos encaixam-se em um destes três métodos: MC, EM ou modelos macro em SPICE. Nenhum deles utiliza a Identificação de Sistemas, cujo propósito é justamente a criação de modelos.

A Identificação de Sistemas pode ser definida como a arte de construir modelos matemáticos que descrevam a dinâmica dos sistemas a partir de seus dados de

entrada e de saída (LJUNG, 1999). Um experimento de Identificação de Sistemas consiste de excitar as entradas com um conjunto “rico” de dados, como o ruído branco, e medir sua resposta amostrada. Assim, o algoritmo de identificação é capaz de produzir uma equação diferencial (sistemas de tempo contínuo) ou a diferenças (caso a natureza do sistema seja discreta ou contínua amostrada) relativa à resposta do sistema (ELKAIN, 2009).

Os sistemas dinâmicos podem ser separados em duas grandes classes: os sistemas lineares e os não lineares. A diferença entre estas duas classes é que a primeira agrupa sistemas que possuem as propriedades de aditividade e homogeneidade. Os sistemas não lineares não possuem tais propriedades (DORF; BISHOP, 2001). Atualmente, a obtenção de modelos para sistemas lineares não representa um desafio à área de Identificação de Sistemas. Contudo, encontrar modelos para sistemas não lineares pode ser uma tarefa de alta complexidade (OROSKI, 2015).

Existe uma gama de modelos que podem ser utilizados na identificação de um sistema não linear. Dentre eles, pode-se citar os modelos de Hammerstein e Wiener, que combinam modelos dinâmicos lineares com não linearidades estáticas (AGUIRRE, 2015), as séries de Volterra, que fornecem estimativas precisas e possuem alto custo computacional (NEMETH; KOLLAR, 2002) e os modelos *Nonlinear Auto Regressive with exogenous input*, NARX, que são precisos e menos custosos que os modelos de Volterra, devido à realimentação da saída (OROSKI, 2015).

Um modelo construído através da Identificação de Sistemas ajusta-se à dinâmica real observada na saída do sistema. Portanto, pode-se dizer que, uma vez que se obtenha um bom modelo, ele tende a ser bastante preciso (AGUIRRE, 2015). Outra grande vantagem específica da aplicação de Identificação de Sistemas à construção de um modelo de SET refere-se à plataforma de simulação. Visto que o modelo final é matemático, ele pode ser simulado nos mais diversos *softwares*. Isto permite ao projetista realizar análises significativas sobre o comportamento do transistor e pode fornecer dados valiosos em seus estudos.

Dado o exposto, evidencia-se a necessidade de uma nova tecnologia que permita a continuidade do progresso da indústria eletrônica. A nanoeletrônica é, atualmente, uma das alternativas mais estudadas. Sendo assim, estudos aprofundados sobre os dispositivos mono-elétron são necessários, a fim de identificar e buscar soluções para os desafios que impedem sua aplicação prática. Neste trabalho, é proposta a construção de um modelo de transistor mono-elétron por meio de conheci-

mentos relativos à área de Identificação de Sistemas. Este modelo será implementado na plataforma MATLAB. Amplamente utilizado em diversas áreas de pesquisa, o MATLAB é uma ferramenta repleta de recursos, que possibilita ao pesquisador avaliar os mais diversos aspectos de seu objeto de estudo.

## 1.1 OBJETIVOS

Este trabalho busca apresentar o desenvolvimento de um novo modelo matemático para o transistor mono-elétron. A finalidade desta proposta é fornecer um modelo conciso, com precisão satisfatória e cuja execução possa ser realizada em diversas plataformas de simulação. Desta forma, pode-se concluir que o desenvolvimento de tal modelo contribui positivamente para que análises necessárias à implementação física do transistor mono-elétron sejam realizadas.

### 1.1.1 OBJETIVO GERAL

O objetivo geral deste trabalho é desenvolver um modelo de transistor mono-elétron baseado em técnicas e conhecimentos inerentes à Identificação de Sistemas.

### 1.1.2 Objetivos Específicos

Para que o objetivo geral citado acima possa ser atingido, alguns passos intermediários, aqui denominados de objetivos específicos, devem ser atingidos:

1. Realizar o levantamento dos dados de saída do transistor mono-elétron, quando um conjunto de dados persistentemente excitante é aplicado em sua entrada;
2. Determinar a classe de modelos não lineares que melhor se adequa ao comportamento do transistor. Para tanto, a métrica utilizada é o *Mean Squared Error*, MSE, além do tempo médio demandado para a simulação da representação matemática;
3. Analisar a possibilidade de otimizar os parâmetros do modelo;
4. Validar o modelo obtido através de *benchmarkings* utilizados na literatura: a aplicação de um conjunto de dados de validação, cálculo do MSE e comparação do tempo médio de simulação em relação ao método de Monte Carlo;

5. Expandir o modelo obtido, considerando diferentes configurações de operação do transistor.

## 1.2 ORGANIZAÇÃO DO TRABALHO

Este trabalho foi dividido em oito capítulos, incluindo a introdução. O conteúdo de cada um destes capítulos será brevemente explanado a seguir.

Conceitos gerais sobre eletrônica mono-elétron são apresentados no Capítulo 2. Buscou-se enfatizar aspectos fundamentais sobre o transistor mono-elétron, objeto de modelagem deste trabalho. Neste capítulo, o transistor mono-elétron foi posicionado dentro da classificação de sistemas lineares, não lineares, variantes e invariantes no tempo. Esta classificação é imprescindível à determinação da melhor representação matemática para o SET.

O Capítulo 3 versa sobre a Identificação de Sistemas. O foco deste capítulo é a modelagem caixa preta, suas ideias e técnicas principais. As representações não lineares consideradas durante o desenvolvimento do modelo caixa preta do SET são expostas neste capítulo.

No Capítulo 4, a modelagem caixa branca do SET, como é encontrada na literatura, é apresentada. Além disso, este capítulo apresenta uma revisão sobre os principais modelos caixa branca já propostos para o transistor mono-elétron.

No Capítulo 5, são apresentados conceitos referentes a métodos de otimização determinística que foram aplicados no desenvolvimento do modelo. Este capítulo também revisa os principais métodos utilizados na concepção de modelos matemáticos com base na Identificação de Sistemas.

A metodologia seguida neste trabalho para a obtenção de um modelo caixa preta para o transistor mono-elétron é detalhada no Capítulo 6.

Os principais resultados são apresentados e analisados no Capítulo 7. Por fim, as principais conclusões e trabalhos futuros são expostos no Capítulo 8.

## 2 ELETRÔNICA MONO-ELÉTRON

Nas últimas cinco décadas, a tecnologia MOS impulsionou o desenvolvimento da indústria eletrônica. Uma grande vantagem dos transistores MOS é a facilidade com que suas dimensões podem ser reduzidas (BOHR, 2002). A redução do comprimento do *gate*, ou porta, proporciona diversas vantagens. Quanto menor o comprimento do *gate*, maior a frequência em que o transistor é capaz de chavear e, portanto, o processamento de dados será mais rápido. Além disso, com dimensões menores, mais transistores cabem em um mesmo *chip*, o que leva a processos de fabricação mais eficientes e econômicos (RAJA; AGRAWAL; BUSHNELL, 2004). Contudo, estudos comprovam que os dispositivos MOS perdem desempenho quando o terminal de porta é muito reduzido. Mark Bohr apontou um aumento de três ordens de grandeza na corrente de fuga para uma redução no comprimento da porta de  $0,5 \mu\text{m}$  para 22 nm (BOHR, 2002).

Além da perda de desempenho dos transistores MOS, problemas devidos à proximidade entre os terminais do transistor foram detectados. Dentre estes, pode-se citar a indução de correntes indesejadas. Desta forma, diversas técnicas de fabricação foram desenvolvidas a fim de solucionar tais problemas. Em 2012, a Intel lançou a família *Ivy Bridge*, inicialmente com 22 nm. Esta tecnologia utiliza transistores tridimensionais, o que diminui os problemas causados pela distância reduzida entre componentes no mesmo *chip*. Recentemente, a Samsung anunciou o processador *Exynos 9820*, que utiliza a tecnologia *Fin Field Effect Transistor*, FinFET, com portas de 8 nm.

Uma vez que dimensões tão reduzidas já foram alcançadas, poderia-se considerar que a indústria eletrônica já é nanoeletrônica. Contudo, a simples redução das dimensões não é a principal característica da nanoeletrônica. Conforme o tamanho dos dispositivos é reduzido, diversos fenômenos puramente quânticos surgem. Estes fenômenos são fontes de erros para dispositivos que têm seu comportamento regido pela física clássica. Dentre estes fenômenos pode-se citar correntes causadas

por tunelamento, ver Seção 2.1.2, e condução balística, i.e., livre de colisões. Estes fenômenos alteram drasticamente o desempenho do dispositivo (HANSON, 2008).

O desenvolvimento da eletrônica mono-elétron foi impulsionado pela procura de soluções que proporcionassem dispositivos eletrônicos mais rápidos e eficientes. Pode-se dizer que a eletrônica mono-elétron consiste em manipular, de forma controlada, elétrons individuais através de uma barreira (JAIN et al., 2015). Para que isto seja possível, as dimensões dos dispositivos devem ser nanométricas. Sendo assim, além da eliminação das dificuldades relativas à redução da escala, um novo paradigma para o transporte de cargas é utilizado: o tunelamento. Devido ao fato de a corrente elétrica consistir, idealmente, de elétrons individuais, o consumo de potência dos dispositivos mono-elétron é bastante reduzido.

Dentre os diversos dispositivos mono-elétron já propostos, o SET ocupa uma posição de destaque. Ele figura entre os candidatos mais promissores para implementações em *Very Large Scale Integration*, VLSI, e *Ultra Large Scale Integration*, ULSI (JAIN; SARKAR, 2004). A relevância do SET para o campo da eletrônica mono-elétron é refletida no grande volume de pesquisas envolvendo este dispositivo. Como exemplos de trabalhos que relatam implementações puramente nanoeletrônicas, pode-se citar as propostas de portas inversoras (ZARDALIDIS; KARAFYLLIDIS, 2006), conversor analógico-digital (CHAOHONG et al., 2006) e arranjos lógicos reconfiguráveis para a realização de portas lógicas (GEROUSIS; GREPIOTIS, 2013). Há, também, vários trabalhos que analisam o desempenho de arquiteturas híbridas SET-MOS. Como exemplo, citam-se as análises feitas por Jana *et al.* (JANA et al., 2013) e por Jain *et al.* (JAIN et al., 2014).

Neste trabalho, será proposto um modelo de SET desenvolvido através de ferramentas da área de Identificação de Sistemas. As seções seguintes, portanto, dedicam-se a explicar conceitos necessários para o entendimento da operação do SET.

## 2.1 DISPOSITIVOS DE TUNELAMENTO MONO-ELÉTRON

A manipulação de elétrons de forma individual, em circuitos de estado sólido, só foi implementada na década de 1980, quando técnicas de fabricação confiáveis surgiram. Uma vez que estas técnicas começaram a apresentar resultados satisfatórios, um novo campo se desenvolveu: a eletrônica mono-elétron (LIKHAREV, 1999).

A Figura 1 ilustra o conceito básico da eletrônica mono-elétron. Considere uma ilha (Seção 2.1.1), material condutor cercado de material isolante, de dimensões reduzidas e eletricamente neutra (possui o mesmo número,  $m$ , de portadores de cargas positivas e negativas). Uma força externa,  $F$ , faz com que um elétron carregue a ilha com uma carga elementar,  $-e$ . No escopo deste trabalho, considera-se que esta injeção de carga se deve ao tunelamento (Seção 2.1.2). Esta carga injetada na ilha gera um campo elétrico,  $\epsilon$ , que repele outros elétrons.

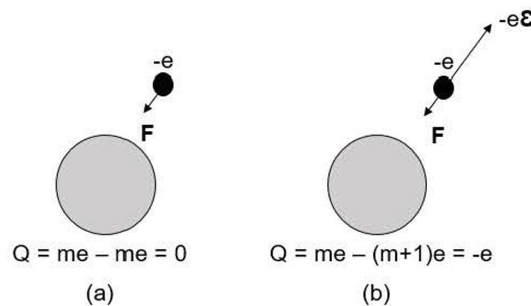


Figura 1: Princípio do controle individual de elétrons: uma ilha condutora (a) antes e (b) depois da adição de um único elétron. Um elétron em excesso cria um campo elétrico, que impede a adição de mais elétrons à ilha - modificada de (LIKHAREV, 1999).

Apesar do pequeno valor da carga elementar,  $e \approx 1,6 \times 10^{-19}$  C, o campo elétrico  $\epsilon$ , para uma ilha de dimensões nanométricas, é intenso. Por exemplo, na superfície de uma esfera com raio de 10 nm no vácuo, o campo elétrico vale aproximadamente 140 kV/cm (LIKHAREV, 1999). Logo, pode-se concluir que, para uma ilha de dimensões reduzidas, o campo elétrico gerado por um único elétron em excesso é capaz de influenciar as probabilidades de tunelamento (GUIMARÃES, 2005).

A corrente elétrica através de uma esfera condutora macroscópica, como a da Figura 1, é dada pela quantidade de carga transferida por unidade de tempo, conforme (2.1). Em dispositivos mono-elétron, a quantidade de carga transferida,  $dQ$ , é quantizada. Sendo assim, idealmente, o dispositivo seria capaz de controlar o fluxo de um único elétron. Na prática, este controle não é tão preciso, de modo que um pequeno fluxo de elétrons passa através do dispositivo quando há condução (GUIMARÃES, 2005).

$$I = \frac{dQ}{dt}, \quad (2.1)$$

em que,  $I$  é a corrente elétrica,  $Q$  é a carga elétrica e  $t$  é o tempo.

Para que haja controle da quantidade de elétrons fluindo durante a condução, é necessário um dispositivo através do qual os elétrons possam fluir de forma discreta.

Este dispositivo, conhecido como junção-túnel, é apresentado na Figura 2. É importante ressaltar, no entanto, que não basta o dispositivo possuir a construção ilustrada na Figura 2 para ser classificado como junção-túnel.

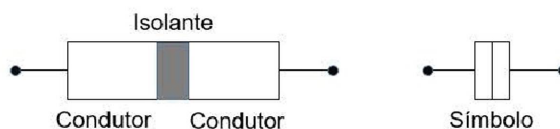


Figura 2: Esquemático e símbolo de uma junção-túnel - modificada de (GUIMARÃES, 2005).

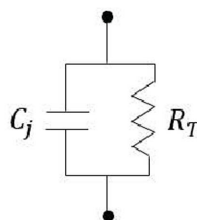


Figura 3: Circuito equivalente da junção-túnel.

A junção-túnel é caracterizada por dois parâmetros macroscópicos: (i) sua resistência de tunelamento,  $R_T$ , e (ii) sua capacitância,  $C_j$ . A resistência de tunelamento é uma quantidade fenomenológica definida quando uma tensão é imposta aos dois eletrodos da junção. A capacitância é uma característica inerente da junção-túnel, que pode ser vista como um capacitor cuja camada de dielétrico é fina o suficiente para permitir que elétrons tunelem através dela (GRABERT; DEVORET, 1992). Desta forma, o circuito equivalente de uma junção-túnel é o de um capacitor que permite fuga de corrente, i.e., um capacitor em paralelo com um resistor (HEINZEL, 2007), conforme ilustrado na Figura 3.

Estes parâmetros são de extrema importância no funcionamento dos dispositivos mono-elétron. A influência de  $R_T$  e  $C_j$  no comportamento destes dispositivos será discutida na Seção 2.1.4.

### 2.1.1 Ilha

Em um transistor mono-elétron, a ilha é a região na qual os elétrons ficam confinados. Nesta seção, os conceitos necessários para a caracterização de uma ilha são apresentados.

Um elétron é considerado livre quando se encontra em uma região do espaço na qual todas as dimensões são substancialmente maiores que o seu comprimento de onda, conforme demonstrado na Figura 4.

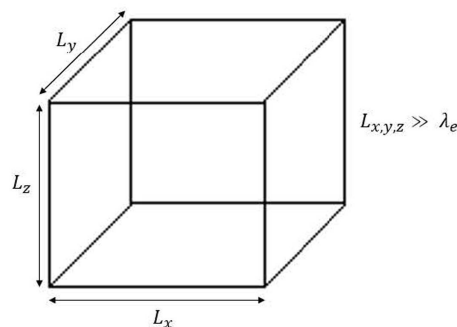


Figura 4: Uma região tridimensional delimitada por  $L_x$ ,  $L_y$  e  $L_z$ ,  $\lambda_e$  é o comprimento de onda do elétron. Desta forma, o elétron será livre se  $L_{x,y,z} \gg \lambda_e$  - modificada de (HANSON, 2008).

Sendo assim, confinar um elétron é equivalente a fazer uma ou mais das dimensões que caracterizam a região ser aproximadamente igual ao comprimento de onda do elétron (HANSON, 2008). Por exemplo, se  $L_x \leq \lambda_e$  e  $L_{y,z} \gg \lambda_e$ , a região é denominada poço quântico. Se  $L_{x,y} \leq \lambda_e$  e  $L_z \gg \lambda_e$ , tem-se um fio quântico. Finalmente, se  $L_{x,y,z} \leq \lambda_e$ , a região é um ponto quântico.

Um aspecto importante a ser considerado diz respeito à interação do ponto quântico com o meio exterior. Uma ilha nada mais é do que um ponto quântico isolado do meio exterior através de barreiras de potencial. Neste trabalho, a situação de interesse consiste em conectar o ponto quântico a meios condutores. Esta conexão deve garantir que a única maneira de elétrons entrarem ou saírem do ponto quântico seja através de tunelamento (HANSON, 2008), ou seja, o ponto quântico é isolado por duas junções túnel, conforme ilustra a Figura 5.

É válido ressaltar que as barreiras isolantes devem ser de fina espessura, a fim de que os elétrons possam tunelar através delas. No caso ideal, apenas um elétron tunela por vez. Deste modo, a variação de carga na ilha é exatamente igual à carga elementar.

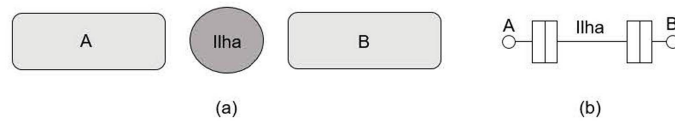


Figura 5: Ilha: região constituída de material condutor cercado por barreiras de material isolante, (a) esquemático e (b) ilha isolada por duas junções-túnel - modificada de (HANSON, 2008).

## 2.1.2 Tunelamento

O campo do tunelamento mono-elétron compreende os fenômenos nos quais o tunelamento de uma carga microscópica leva a efeitos observáveis na escala macroscópica (GRABERT; DEVORET, 1992). Na década de 1920, o entendimento do efeito túnel desafiava os cientistas (MERZBACHE, 2002). Desde então, diversas aplicações práticas que envolvem este fenômeno quântico foram desenvolvidas. Um dos exemplos mais famosos consiste na caracterização da superfície de materiais utilizando o microscópio de tunelamento com varredura, *Scanning Tunneling Microscope* - STM. Esta seção dedica-se a explicar o tunelamento mono-elétron, mecanismo de transporte de carga nos dispositivos utilizados neste trabalho.

Em 1923, Louis de Broglie sugeriu que partículas com energia,  $E$ , e momento linear,  $p$ , deveriam possuir propriedades ondulatórias. Desta forma, cada partícula possuiria um comprimento de onda associado. O momento linear e o comprimento de onda se relacionariam segundo (2.2). Este comprimento de onda também é chamado de comprimento de onda de de Broglie, e é dado por:

$$\lambda = \frac{h}{mv}, \quad (2.2)$$

em que  $\lambda$  é o comprimento de onda em metros,  $h = 6,626 \times 10^{-34}$  kg(m<sup>2</sup>/s) é a constante de Planck,  $m$  é a massa da partícula, em kg, e  $v$  é o módulo de sua velocidade, em m/s. O momento linear é dado por  $p = mv$ , em kg(m/s).

A sugestão de de Broglie foi comprovada posteriormente. Em 1926, Clinton Davission e Lester Germer observaram padrões em um experimento de difração de elétrons em um cristal. Estes padrões de difração já haviam sido observados em experimentos com fótons. A Equação (2.2) deixa claro que a massa da partícula deve ser pequena para que sua natureza ondulatória seja observável em um dado sistema. Esta relação se deve ao valor reduzido de  $h$ . Considerando a escala nanométrica, o

comprimento de onda,  $\lambda$ , deve ser da ordem de nanômetros para que o comportamento ondulatório da partícula possa ser observado. Comprimentos de onda desta magnitude podem ser obtidos para elétrons e outras partículas subatômicas, que possuem pequena massa. Quando  $\lambda$  é muito pequeno em comparação com a distância na qual a energia potencial muda, as partículas quânticas comportam-se de maneira clássica (HANSON, 2008).

Também na década de 1920, inspirado pelo trabalho de Louis de Broglie, Erwin Schrödinger desenvolveu a equação de onda. Esta equação, conhecida como equação de Schrödinger, é satisfeita por uma função de onda,  $\Phi$ . Partículas quânticas podem ser associadas a uma função de onda (BORN, 1927). O quadrado do módulo desta função,  $|\Phi|^2$ , resulta na probabilidade de se encontrar a partícula em uma dada região.

Em termos gerais, o tunelamento é um fenômeno quântico que permite o transporte de partículas por regiões classicamente proibidas, ou seja, regiões cuja energia potencial é maior do que a energia total da partícula (GUIMARÃES, 2005). A Figura 6 ilustra o caso em que uma partícula com energia  $E$  incide, pela esquerda, em uma barreira de potencial.

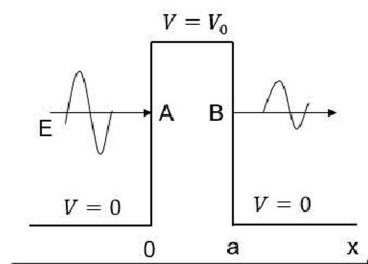


Figura 6: Tunelamento de uma partícula de energia  $E$  através de uma barreira de potencial  $V_0 > E$ . Os pontos A e B são os pontos críticos clássicos - modificada de (FERRY, 1991).

Para este caso geral de tunelamento, duas condições são de interesse: (i) quando a energia da partícula,  $E$ , é maior do que a energia da barreira,  $E > V_0$ , e (ii) quando  $E < V_0$ . Para a física clássica, estas são condições de análise simples. Para  $E > V_0$ , a partícula atravessará a barreira. Para  $E < V_0$ , a partícula não atravessará a barreira.

A mecânica quântica impõe condições matemáticas de continuidade à função de onda e sua derivada na fronteira entre duas regiões com energias potenciais dife-

rentes: ambas devem ser contínuas. Uma vez que  $\varphi$  é contínua na fronteira e que  $|\varphi|^2$  representa a probabilidade de se encontrar uma partícula em uma dada região, existe uma probabilidade não nula de encontrar a partícula do outro lado da barreira (GUIMARÃES, 2005).

Existem diversos efeitos que podem tornar complexa a análise do tunelamento (GRABERT; DEVORET, 1992). Contudo, há uma teoria que descreve o tunelamento mono-elétron de forma bastante simplificada: a teoria ortodoxa do tunelamento mono-elétron. Apesar de ser bastante simples, esta teoria descreve, com precisão satisfatória, a maioria dos dados obtidos em experimentos com dispositivos mono-elétron (LIKHAREV, 1999). A próxima seção dedica-se a explicar os postulados e as principais conclusões desta teoria.

### 2.1.3 Teoria Ortodoxa do Tunelamento Mono-Elétron

A teoria ortodoxa do tunelamento mono-elétron foi proposta, para uma junção metal-óxido-metal, por Kulik e Shekhter (KULIK; SHEKHTER, 1975). Posteriormente, Averin e Likharev propuseram uma generalização da teoria (AVERIN; LIKHAREV, 1991). Os principais postulados desta teoria são:

- A quantização da energia no interior dos condutores é desprezada. Desta forma, o espectro de energia dos elétrons é considerado contínuo. Esta suposição é válida sempre que a energia cinética do elétron for muito menor do que a energia de carregamento, dada pela Equação (2.11).
- O tempo,  $\tau_t$ , entre dois tunelamentos sucessivos através da barreira é negligenciável, quando comparado com outras escalas temporais do sistema. O valor de  $\tau_t$ , para dispositivos práticos, é da ordem de  $10^{-15}$  s. Assim, esta suposição é bastante razoável.
- Co-tunelamentos<sup>1</sup> são ignorados. A resistência de tunelamento da junção,  $R_T$ , deve ser muito maior do que a resistência quântica,  $R_Q$ , para que esta suposição seja verdadeira,  $R_Q$  é conceituada na Seção 2.1.4.

O principal resultado da teoria ortodoxa é descrever o tunelamento mono-elétron como um evento aleatório. Este evento possui uma determinada taxa de

<sup>1</sup>Co-tunelamentos podem ser resumidamente conceituados como dois ou mais tunelamentos simultâneos, ou seja, é um fenômeno quântico no qual  $p$  elétrons tunelam através de  $p$  barreiras de potencial (GRABERT; DEVORET, 1992).

ocorrência,  $\Gamma$ , i.e., probabilidade por unidade de tempo, que depende apenas da redução da energia eletrostática livre do sistema,  $\Delta F$ , causada pelo tunelamento (LIKHA-REV, 1999).

#### 2.1.4 Efeito de Carregamento

Conforme visto na Seção 2.1.1, a eletrônica mono-elétron baseia-se na manipulação individual de elétrons. Para que esta manipulação seja possível, é necessário que o efeito de carregamento seja predominante na operação do dispositivo (GEERLIGS; MOJUI, 1991). Este efeito ocorre quando um elétron carrega um material condutor, isolado eletricamente, de dimensões muito pequenas. As dimensões reduzidas fazem com que o potencial eletrostático varie significativamente, mesmo que apenas um elétron entre no material. Para uma ilha de dimensões nanométricas, com capacitância de 1 aF, esta variação é de aproximadamente 160 mV, ou seja, consideravelmente superior ao ruído térmico à temperatura ambiente, que é de 26 mV (ONO et al., 2005). Capacitâncias de 1 aF foram obtidas experimentalmente através de litografia por feixe de elétrons (PASHKIN; NAKAMURA; TSAI, 2000) e através do uso de uma molécula de fulereno,  $C_{60}$ , como ilha (PARK et al., 2000), por exemplo.

O aumento no potencial eletrostático impede que outros elétrons tunelem para dentro do material, dando origem ao chamado bloqueio de Coulomb. Os primeiros estudos sobre o bloqueio de Coulomb foram realizados em sistemas formados por partículas metálicas. Estes estudos demonstraram que o fluxo de corrente através do material poderia ser suprimido, se a capacitância da partícula fosse suficientemente pequena (HANSON, 2008). A tensão e a carga, em um capacitor, relacionam-se conforme (2.3).

$$Q = C_j V, \quad (2.3)$$

em que,  $Q$  é a carga em cada placa condutora ( $+Q$  para a placa positiva e  $-Q$ , para a negativa),  $C_j$  é a capacitância da junção e  $V$  é a tensão entre as placas.

A energia eletrostática armazenada em um capacitor é dada por (2.4) e representa a energia necessária para separar as cargas através de aplicação de potencial nas placas do capacitor.

$$E = \frac{1}{2}C_j V^2 = \frac{Q^2}{2C_j}. \quad (2.4)$$

A capacitância de um capacitor de placas paralelas depende da permissividade elétrica do material entre as placas,  $\epsilon$ , da distância que separa as placas,  $d$ , e da área das placas,  $A$ , conforme (2.5).

$$C_j = \frac{\epsilon A}{d}. \quad (2.5)$$

De acordo com a Equação (2.5), existem duas formas de se obter um valor pequeno de capacitância: (i) aumentando a distância,  $d$ , entre as placas ou (ii) reduzindo a área das placas. Uma vez que os elétrons devem tunelar através das barreiras, a camada isolante que separa as placas deve ter espessura reduzida. Assim, a primeira opção é descartada. Deste modo, a redução da área das placas é fundamental para que o bloqueio de Coulomb possa ser observado.

Considere um capacitor de dimensões nanométricas cuja energia eletrostática inicial,  $E_i$ , é dada por (2.4). Se um elétron tunelar através da camada isolante do capacitor, ele fará a carga em cada placa variar de  $\pm e$ . Deste modo, a energia eletrostática armazenada, após o tunelamento será (GRABERT; DEVORET, 1992):

$$E_f = \frac{(Q - e)^2}{2C_j}. \quad (2.6)$$

A variação na energia eletrostática é dada pela diferença entre as energias inicial e final,

$$\Delta E = E_i - E_f = \frac{e(Q - \frac{e}{2})}{C_j}. \quad (2.7)$$

Considerando a temperatura  $T = 0$  K, para que o tunelamento ocorra, a variação de energia definida em (2.7) deve ser positiva,  $\Delta E > 0$  (HANSON, 2008), ou seja, a carga e a tensão devem satisfazer (2.8) e (2.9), respectivamente.

$$Q > \frac{e}{2} \quad (2.8)$$

$$V > \frac{e}{2C_j} \quad (2.9)$$

Uma vez que o tunelamento pode ocorrer em ambas as direções, tem-se que, no intervalo dado por (2.10), não há fluxo de corrente (GRABERT; DEVORET, 1992).

$$-\frac{e}{2C_j} < V < \frac{e}{2C_j}, \quad (2.10)$$

ou seja, só haverá tunelamento quando uma tensão suficientemente grande ( $|V| > \frac{|e|}{2C_j}$ ) for imposta ao capacitor. Este efeito é chamado de bloqueio de Coulomb, sua característica é ilustrada na Figura 7 (HANSON, 2008).

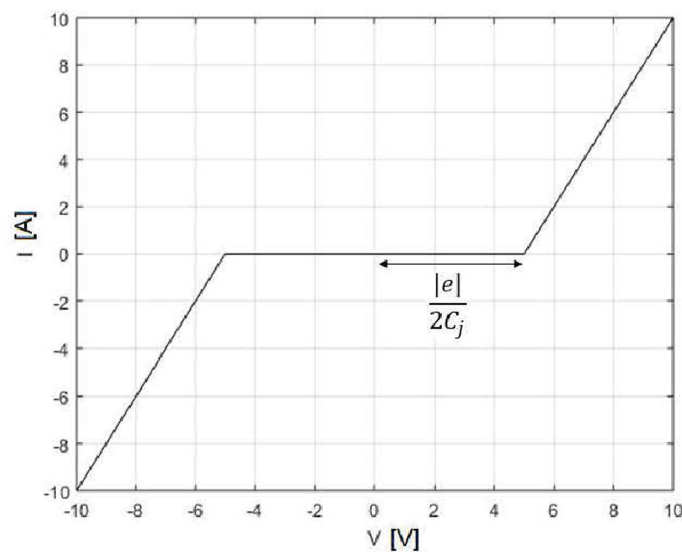


Figura 7: Característica I-V com bloqueio de Coulomb evidente - modificada de (HANSON, 2008).

Embora a análise acima seja feita a 0 K, para aplicações práticas de dispositivos mono-elétron, o bloqueio de Coulomb deve ser evidente em temperaturas mais altas. Para estas temperaturas, há energia térmica disponível. Desta forma, deve-se garantir que os elétrons não sejam transportados devido à energia térmica. Para tanto, a energia de carregamento,  $E_c$ , dada por (2.11) e definida como a energia necessária para adicionar uma carga ao capacitor, deve ser superior à energia térmica disponível.

$$E_c = \frac{e^2}{2C_j}. \quad (2.11)$$

A condição expressa em (2.12) é, assim, o primeiro requisito para que o efeito de carregamento seja predominante na operação dos dispositivos mono-elétron, quando  $T \neq 0$  K.

$$E_c \gg k_B T, \quad (2.12)$$

em que,  $k_B$  é a constante de Boltzmann ( $1,38064852 \times 10^{-23} \text{m}^2 \text{kg s}^{-2} \text{K}^{-1}$ ) e  $T$  é a temperatura, em K.

Substituindo (2.11) em (2.12) e rearranjando os termos, obtém-se (2.13). A condição, escrita nesta forma, apresenta uma restrição ao valor da capacitância da junção-túnel.

$$C_j \ll \frac{e^2}{k_B T}. \quad (2.13)$$

A segunda condição diz respeito ao outro parâmetro macroscópico que descreve a junção-túnel, sua resistência de tunelamento. Este requisito tem origem no princípio da incerteza de Heisenberg. Este princípio foi proposto em 1927 e impõe restrições à precisão com que se pode efetuar medidas simultâneas em um sistema subatômico, no qual o mecanismo de medição influencia o resultado da medição (CHIBENI, 2004). O princípio da incerteza de Heisenberg pode ser enunciado conforme (2.14).

$$\Delta E \Delta t \geq \hbar, \quad (2.14)$$

em que,  $\Delta E$  é a incerteza quanto à energia,  $\Delta t$  é a incerteza quanto ao tempo e  $\hbar$  é a constante de Planck.

A incerteza relativa ao tempo,  $\Delta t$ , será dada pelo tempo característico para flutuações de carga. Este, por sua vez, é dado pela escala de tempo relacionada à junção, de acordo com (2.15) (GRABERT; DEVORET, 1992).

$$\Delta t \approx R_T C_j, \quad (2.15)$$

em que,  $R_T$  é a resistência de tunelamento da junção e  $C_j$  é a capacitância da junção.

A incerteza relativa à energia é dada pela variação de energia associada a um elétron em excesso, conforme (2.11).

A substituição das equações (2.15) e (2.11) em (2.14) resulta em um valor mínimo para a resistência de tunelamento da junção, conforme (2.16).

$$R_T \geq \frac{h}{e^2}. \quad (2.16)$$

Assim, o valor mínimo para  $R_T$  é de, aproximadamente,  $25,8 \text{ k}\Omega$ . A quantidade  $\frac{h}{e^2}$  é frequentemente chamada de resistência quântica,  $R_Q$ . Desta forma, a condição expressa por (2.16) também pode ser escrita como  $R_T \geq R_Q$ .

Essencialmente, a condição expressa em (2.16) garante que a função de onda de um elétron em excesso está localizada na ilha (GRABERT; DEVORET, 1992). A próxima seção explica o funcionamento do transistor mono-elétron, dispositivo a ser modelado neste trabalho.

### 2.1.5 SET

O transistor mono-elétron consiste de duas junções-túnel em série, formando uma ilha. O potencial eletrostático nesta ilha é controlado através do terminal de *gate* (HEINZEL, 2007). Neste terminal, uma tensão externa é aplicada à ilha por meio de um capacitor. A Figura 8 ilustra a configuração do SET.

O efeito do *gate* é facilmente visualizado através do diagrama de bandas de energia. Uma tensão de *gate* positiva,  $V_g > 0$ , deprime o nível de Fermi (nível energético mais alto ocupado a  $T = 0 \text{ K}$ ) na ilha. A Figura 9 exemplifica o caso em que a tensão de *gate* deprime o *gap* gerado pelo bloqueio de Coulomb, de modo que ele fique alinhado com os níveis de Fermi à esquerda e à direita da ilha.

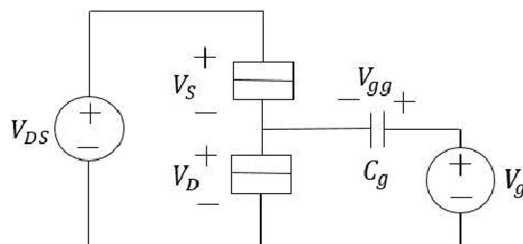


Figura 8: Transistor mono-elétron.  $V_{DS}$  é a tensão de polarização do transistor,  $V_S$  e  $V_D$  são as tensões nas junções-túnel dos terminais de fonte e dreno, respectivamente.  $C_g$  é o valor da capacitância do capacitor no terminal de porta,  $V_{gg}$  é a queda de tensão neste capacitor e  $V_g$  é a tensão de controle aplicada ao terminal de porta - modificada de (HANSON, 2008).

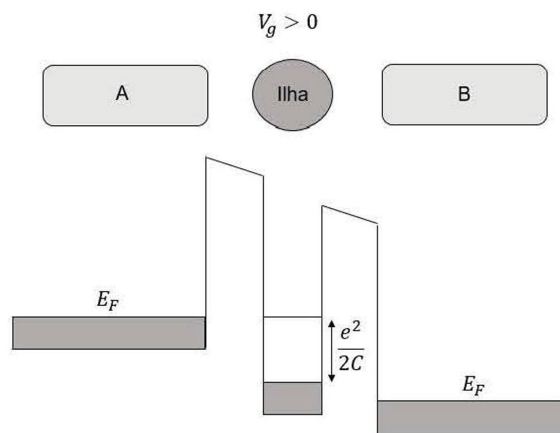


Figura 9: SET com  $V_g > 0$ . A tensão deprime o *gap* gerado pelo bloqueio de Coulomb, permitindo, assim, que um elétron entre na ilha.  $E_F$  indica a energia do nível de Fermi - modificada de (HANSON, 2008).

Conforme discutido na Seção 2.1.4, para que haja tunelamento através da junção-túnel, a tensão entre os terminais da junção deve satisfazer (2.9), portanto  $V > e/2C_j$ . No transistor mono-elétron, o fluxo de corrente é estabelecido através da ilha, cuja tensão é controlada no terminal de *gate*. Desta forma, mantendo-se  $V_{DS}$  constante e aumentando  $V_g$  de forma gradual, a partir de  $V_g = 0$ , um elétron terá condições de passar pela ilha quando  $V_g = e/2C_g$ . Quando este elétron tunelar para fora da ilha, não haverá fluxo de corrente até que  $V_g = 3e/2C_g$ , dando condições de outro elétron entrar na ilha. Consequentemente, apenas em valores múltiplos da tensão  $V_g$  tais que,

$$\frac{e}{2C_g}, \frac{3e}{2C_g}, \frac{5e}{2C_g}, \dots, \frac{(2n+1)e}{2C_g}, \quad (2.17)$$

em que  $n \in \mathbb{N}$ , haverá fluxo de corrente na ilha (GRABERT; DEVORET, 1992). Este comportamento fica evidente na característica  $I_{DS}-V_g$  do SET, exemplificada na Figura 10, este efeito é denominado Oscilações de Coulomb.

#### 2.1.5.1 Equações Características

As equações para as cargas nas junções e para a energia eletrostática armazenada, em um SET, são apresentadas a seguir. Analisando a Figura 8, pode-se concluir que a carga,  $Q$ , na ilha é dada por:

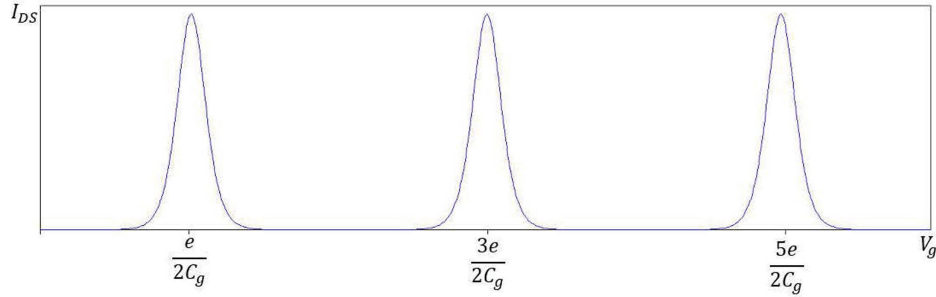


Figura 10: Característica  $I_{DS}$ - $V_g$  do SET.

$$Q = Q_D - Q_S - Q_g, \quad (2.18)$$

em que,  $Q_D$  é a carga na junção-túnel de dreno,  $Q_S$  é a carga na junção-túnel de fonte e  $Q_g$  é a carga no capacitor de *gate*. Aplicando a lei de Kirchhoff para as tensões nas duas malhas do circuito mostrado na Figura 8, tem-se:

$$V_{DS} = V_S + V_D = \frac{Q_S}{C_S} + \frac{Q_D}{C_D}, \quad (2.19)$$

$$V_g = V_{gg} + V_D = \frac{Q_g}{C_g} + \frac{Q_D}{C_D}, \quad (2.20)$$

em que,  $C_S$ ,  $C_D$  e  $C_g$  são as capacitâncias da junção de fonte, de dreno e do capacitor de *gate*, respectivamente.

Manipulando as equações (2.18)-(2.20), pode-se chegar às seguintes expressões para as cargas nas junções e no capacitor de *gate*:

$$Q_S = C_S \frac{V_{DS}(C_D + C_g) - C_g V_g - Q}{C_{ilha}}, \quad (2.21)$$

$$Q_D = C_D \frac{C_S V_{DS} + C_g V_g + Q}{C_{ilha}}, \quad (2.22)$$

$$Q_g = C_g \frac{V_g(C_D + C_S) - C_S V_{DS} - Q}{C_{ilha}}, \quad (2.23)$$

em que,  $C_{ilha} = C_S + C_D + C_g$ . Utilizando as equações (2.21)-(2.23), a energia armazenada nos capacitores,  $E_{arm}$ , será dada por:

$$E_{arm} = \frac{1}{2C_{ilha}} [C_g C_S (V_{DS} - V_g)^2 + C_S C_D V_{DS}^2 + C_D C_g V_D^2 + Q^2]. \quad (2.24)$$

A corrente no transistor mono-elétron pode ser calculada considerando-se a taxa de tunelamento através de cada junção, conforme (2.25) (GRABERT; DEVORET, 1992). A taxa de tunelamento representa a probabilidade de tunelamento por unidade de tempo. O cálculo desta taxa, dada pela equação (4.1), é detalhado no capítulo 4.

$$I_{DS}(t) = e \frac{\Gamma_S^+ \Gamma_D^-}{\Gamma_S^+ + \Gamma_D^-}, \quad (2.25)$$

em que,  $\Gamma_S^+$  representa a taxa de tunelamento da fonte para a ilha e  $\Gamma_D^-$  representa a taxa de tunelamento da ilha para o dreno. Assim, o sinal positivo indica que a carga está tunelando para dentro da ilha e o sinal negativo indica que ela está saindo da ilha. A próxima seção dedica-se a apresentar uma análise do transistor mono-elétron sob a ótica da área de Controle.

## 2.2 SET: OBJETO DE MODELAGEM

Pode-se separar as técnicas de identificação de sistemas em grandes grupos, como se segue: (i) aplicáveis a sistemas lineares e a sistemas não lineares e (ii) aplicáveis a sistemas variantes e invariantes no tempo. Uma vez que o objetivo desta tese consiste em construir um modelo para o transistor mono-elétron, nas próximas seções, será apresentada a classificação do SET como um sistema não linear e invariante no tempo.

Dentre os modelos teóricos, os **Sistemas Lineares e Invariantes no Tempo**, SLIT, constituem uma das classes mais abordadas na literatura (ISERMANN; MÜNCHHOF, 2011). Para tratar estes sistemas, primeiramente é necessário definir os conceitos de linearidade e invariância no tempo.

### 2.2.1 Linearidade

Sistemas lineares apresentam relação linear entre causa e efeito. Para se classificar um sistema como linear, ele deve obedecer ao princípio da superposição (DORF; BISHOP, 2001).

O princípio da superposição agrupa duas propriedades: a homogeneidade e a aditividade. Para a descrição destas propriedades, considere um sistema com entrada,  $u(k)$ , e saída,  $y(k)$ . Sendo  $k \in \mathbb{Z}$  e  $u(k), y(k) \in \mathbb{R}$  de natureza discreta ou contínua amostrada. A propriedade da homogeneidade diz que:

- Se, dada uma entrada  $u(k)$ , obtém-se uma saída  $y(k)$ ;

Então, uma entrada  $au(k)$  produz uma saída,  $ay(k)$ , em que  $a \in \mathbb{R}$ .

A propriedade da aditividade diz que:

- Se as entradas  $u_1(k), u_2(k), \dots, u_n(k)$  produzem as saídas  $y_1(k), y_2(k), \dots, y_n(k)$ , respectivamente;

Então, uma entrada  $u(k) = u_1(k) + u_2(k) + \dots + u_n(k)$  produz uma saída,  $y(k) = y_1(k) + y_2(k) + \dots + y_n(k)$ .

Generalizando, pode-se dizer que, em um sistema linear, quando aplicada uma combinação linear de entradas  $\sum_{k=0}^{N-1} c_k u(k)$ , obtém-se uma combinação linear de saídas  $\sum_{k=0}^{N-1} c_k y(k)$ , em que  $c_k$  são escalares reais e  $u(k)$  e  $y(k)$  são sinais discretos no tempo.

Deste modo, se um sistema possui as propriedades de homogeneidade e aditividade, ele respeita o princípio da superposição e, portanto, pode ser classificado como linear (MONTEIRO, 2003).

## 2.2.2 Invariância no tempo

A classificação de um sistema como invariante no tempo pode ser feita como se segue:

- Se, dada uma entrada  $u(k)$ , obtém-se uma saída  $y(k)$ ;

Então, dada uma entrada  $u(k - \tau)$ , obtém-se uma saída  $y(k - \tau)$ , sendo  $k$  e  $\tau \in \mathbb{Z}$ ,  $u(k)$  e  $y(k) \in \mathbb{R}$ .

A fusão das propriedades de linearidade e de invariância no tempo é característica da classe de sistemas LIT.

Na Figura 11 é mostrado um sistema LIT, modelado por uma função de transferência  $H(z)$ . Note que o sistema da figura possui apenas uma entrada,  $x(t) \in \mathbb{R}$ , e apenas uma saída,  $y(t) \in \mathbb{R}$  para  $t \in \mathbb{R}^+$ , sendo, portanto, um sistema **Single-Input-Single-Output**, SISO.

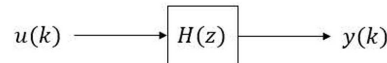


Figura 11: Diagrama de um sistema LIT representado por  $H(z)$ . As variáveis  $u(k)$  e  $y(k)$  representam a entrada e a saída, respectivamente.

Para sistemas discretos, que são de interesse no escopo deste trabalho, as funções de transferência podem ser definidas como uma relação entre os sinais de saída,  $y(k) \in \mathbb{R}$ , e entrada,  $u(k) \in \mathbb{R}$ , para  $k \in \mathbb{Z}$ , ambos no domínio da frequência.

$$H(z) = \frac{Y(z)}{U(z)} = \frac{\mathcal{Z}\{y(k)\}}{\mathcal{Z}\{u(k)\}}, \quad (2.26)$$

sendo a transformada  $\mathcal{Z}$  definida por:

$$\mathcal{Z}\{u(k)\} = \sum_{-\infty}^{\infty} u(k)z^{-k}, \quad (2.27)$$

em que  $u(k)$  é o sinal a ser transformado e a variável  $z$  é definida como  $z = re^{jw}$ , sendo  $j$  a unidade imaginária,  $r \in \mathbb{R}$  e  $w$  a frequência angular, em radianos.

### 2.2.3 SET: sistema não linear e invariante no tempo

Nesta seção, a caracterização do SET como um sistema não linear e invariante no tempo é detalhada. Todos os gráficos apresentados nesta seção foram obtidos por meio de simulação efetuada no SIMON.

Iniciando pela não linearidade do sistema, como visto na Seção 2.2.1, deve-se checar se o sistema possui as propriedades de homogeneidade e aditividade. Para checar a homogeneidade, foram aplicadas ao transistor as entradas  $V_{gs1}$  e  $V_{gs2}$ , em que  $V_{gs2} = 2V_{gs1}$ . Conforme pode-se notar na Figura 12, a amplitude da resposta do transistor,  $I_{ds}$ , não é alterada.

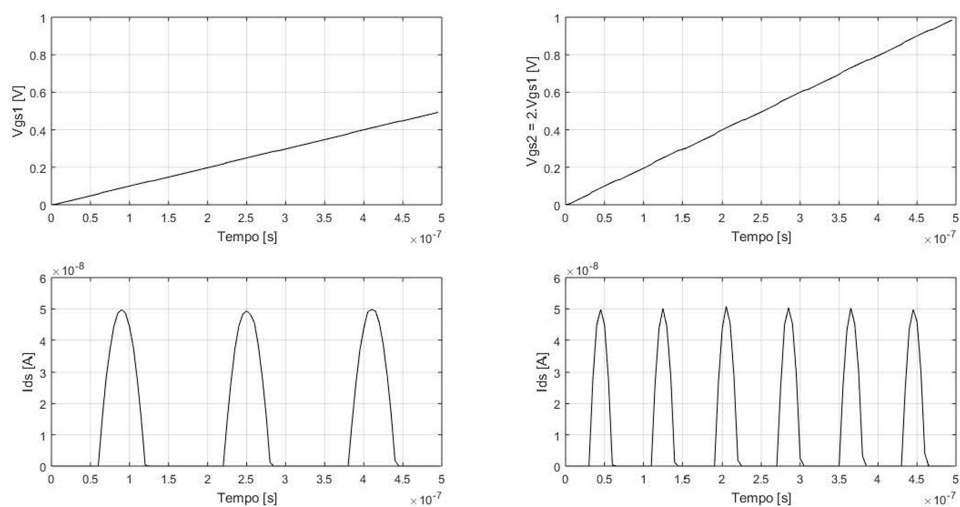


Figura 12: Teste de homogeneidade: a amplitude da resposta do SET se mantém inalterada, portanto, o transistor não obedece ao princípio da homogeneidade.

Para checar se o SET possui a propriedade da aditividade, foram aplicadas três entradas distintas ao transistor:  $V_{gs1}$ ,  $V_{gs2}$  e  $V_{gs3} = V_{gs1} + V_{gs2}$ . As respostas para as três entradas são ilustradas na Figura 13. Para maior clareza, considere a Figura 14, na qual pode-se ver que a resposta  $I_{ds3}$ , para a entrada  $V_{gs3}$ , é diferente de  $I_{ds4}$ , que representa  $I_{ds1} + I_{ds2}$ . Deste modo, o SET não obedece ao princípio da aditividade.

Uma vez que o transistor mono-elétron não possui as propriedades de homogeneidade e de aditividade, ele também não obedece ao princípio da superposição. Deste modo, conclui-se que o SET é um sistema não linear.

Para determinar se o transistor mono-elétron é variante ou invariante no tempo, foram aplicadas duas entradas:  $V_{gs1}$  e  $V_{gs2} = V_{gs1}(t + 250 \text{ ns})$ . Para ser considerado invariante no tempo, o transistor deve fornecer a mesma resposta aos dois sinais, sendo a segunda deslocada de 250 ns em relação à primeira. A Figura 15 mostra os resultados obtidos.

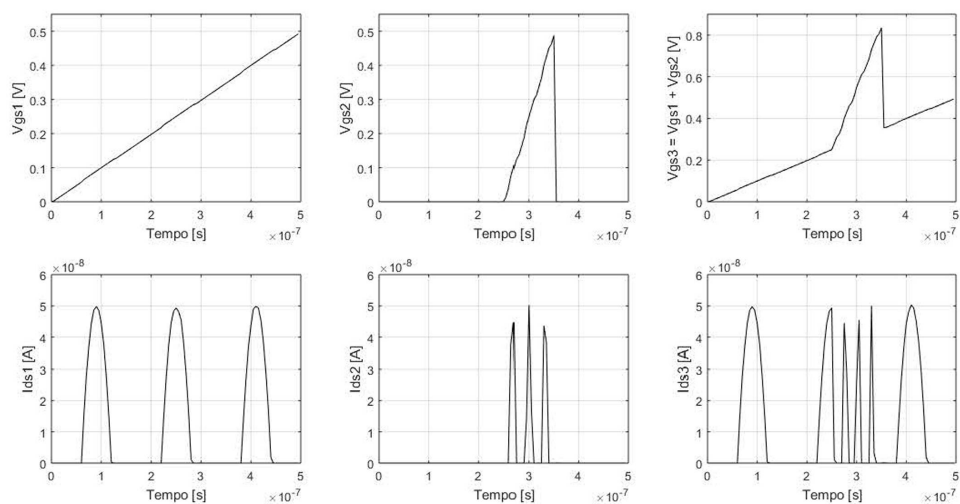


Figura 13: Teste de aditividade:  $I_{ds1}$  é a resposta do transistor à entrada  $V_{gs1}$  e  $I_{ds2}$  é a resposta do transistor à entrada  $V_{gs2}$ .  $I_{ds3}$  é a resposta do transistor à entrada  $V_{gs3} = V_{gs1} + V_{gs2}$ .

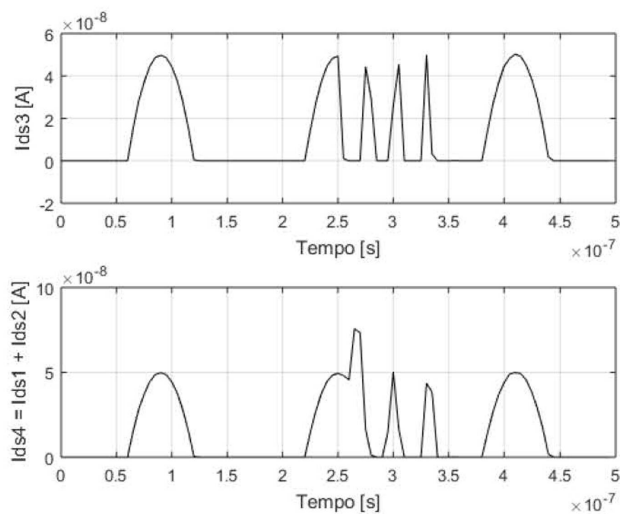


Figura 14: Teste de aditividade:  $I_{ds4} = I_{ds1} + I_{ds2}$  e  $I_{ds3} \neq I_{ds4}$ , portanto, o SET não respeita a propriedade da aditividade.

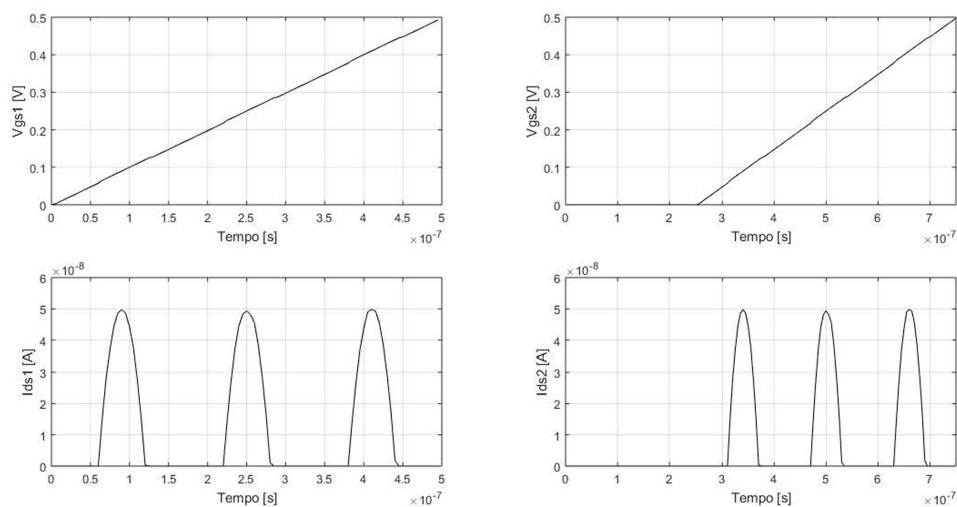


Figura 15: Invariância no tempo:  $I_{ds1}$  é a resposta do transistor à  $V_{gs1}$ .  $I_{ds2}$  é a resposta do transistor à entrada  $V_{gs2} = V_{gs1}(t + 250 \text{ ns})$ .

Para a saída  $I_{ds1}$ , as oscilações de Coulomb têm início em  $t = 60 \text{ ns}$ . Já  $I_{ds2}$  começa a apresentar oscilações de Coulomb em  $t = 310 \text{ ns}$ , exatamente 250 ns deslocada no tempo em relação à  $I_{ds1}$ . Desta forma, conclui-se que o SET possui a propriedade da invariância no tempo, ou seja, o transistor mono-elétron, quando analisado através de simulação, consiste em um sistema não linear e invariante no tempo. Sendo assim, o próximo capítulo dedica-se a apresentar algumas representações não lineares que foram consideradas neste trabalho.

## 2.3 CONSIDERAÇÕES FINAIS

Neste capítulo, os principais conceitos referentes à eletrônica mono-elétron e ao SET foram apresentados. O funcionamento e as equações características do transistor mono-elétron foram explicitados. Por fim, o transistor mono-elétron, objeto de modelagem deste trabalho, foi classificado como um sistema não linear e invariante no tempo

### 3 IDENTIFICAÇÃO DE SISTEMAS

Nas ciências exatas, modelar matematicamente um sistema é uma tarefa cotidiana. Especialmente na Engenharia, em que se busca prever comportamentos específicos, a modelagem de sistemas ganha uma dimensão bastante relevante.

Modelar um sistema pode ou não exigir conhecimento prévio sobre ele. É comum, desta forma, dividir as técnicas de modelagem em três grandes grupos, a saber:

- **Modelagem Caixa Branca:** baseada nas leis físicas e conceitos que regem o sistema;
- **Modelagem Caixa Preta:** baseada exclusivamente nos dados de entrada e saída do sistema;
- **Modelagem Caixa Cinza:** funde as duas metodologias anteriores. Ou seja, faz uso de conhecimento prévio sobre o sistema (estrutura ou leis físicas) e de medidas de sinais de entrada e saída.

A modelagem caixa branca é frequentemente chamada de modelagem física. As modelagens caixa preta e caixa cinza podem ser agrupadas dentro da Identificação de Sistemas. Esta área busca obter modelos puramente matemáticos que se baseiam, em algum nível, em dados coletados de experimentos realizados com o objeto de modelagem (AGUIRRE, 2015). Entretanto, é comum na literatura que a Identificação de Sistemas se confunda com a modelagem caixa preta.

A modelagem caixa preta é preferível em situações em que a modelagem caixa branca é muito complexa ou, até mesmo, impossível. Para a construção de modelos caixa branca, é necessário um vasto conhecimento sobre as leis físicas que regem o sistema. Em situações práticas, no entanto, raramente se possui tanto conhecimento sobre o sistema a ser modelado. Além disso, as leis que regem sistemas

reais podem apresentar complexidade razoável. Desta forma, na modelagem caixa branca costuma-se fazer suposições que simplificam a dinâmica do sistema e reduzem a precisão do modelo (ISERMANN; MÜNCHHOF, 2011).

Existem diversos modelos analíticos, elaborados através de modelagem caixa branca, para o transistor mono-elétron. Estes modelos serão revisados no Capítulo 4. Neste trabalho, propõe-se o uso da modelagem caixa preta para a obtenção de um modelo matemático para o SET. Desta forma, as próximas seções dedicam-se a explicar a metodologia envolvida na modelagem caixa preta.

Conforme visto na Seção 2.2.3, o transistor mono-elétron é um sistema não linear. Deste modo, a seção final deste capítulo mostra algumas representações não lineares, consideradas durante a construção do modelo caixa preta para o SET.

### 3.1 MODELAGEM CAIXA PRETA

A modelagem caixa preta obtém modelos experimentais para os sistemas sob estudo (ISERMANN; MÜNCHHOF, 2011). A denominação experimental deriva da metodologia utilizada para o desenvolvimento do modelo. Os modelos caixa preta são construídos a fim de explicar, pelo menos em parte, as relações de causa e efeito presentes nos dados de entrada e saída coletados do sistema (AGUIRRE, 2015). Pode-se dizer, então, que os parâmetros de um modelo caixa preta não fornecem interpretação física sobre o comportamento do sistema.

Segundo Coelho (COELHO, 2016), a obtenção de um modelo através da abordagem caixa preta deve seguir quatro etapas, a saber:

1. Medição das variáveis de entrada e saída do sistema (coleta de dados);
2. Determinação da estrutura do modelo;
3. Estimação dos parâmetros;
4. Validação do modelo.

Estas quatro etapas descrevem o chamado experimento de Identificação de Sistemas. Cada uma delas exige que o projetista faça algumas escolhas que podem estar bem fundamentadas na literatura e outras que podem ser subjetivas. Nas próximas seções, cada uma destas etapas será discutida.

## 3.2 COLETA DE DADOS

No processo de identificação, o sistema pode ser excitado por sinais operacionais de entrada ou por sinais criados artificialmente, os sinais de teste (ISERMANN; MÜNCHHOF, 2011). Na prática, os sinais de teste são muito utilizados, pois são concebidos para excitar as diversas dinâmicas e não linearidades do sistema.

O sinal escolhido para alimentar o sistema na etapa de coleta de dados tem duas funções principais. A primeira diz respeito a excitar as características dinâmicas e estáticas do sistema, na faixa de frequências desejada. A segunda, refere-se à excitação das características não lineares do sistema. Esta excitação está atrelada à variação de amplitude presente no sinal de excitação. Desta forma, a escolha do sinal de teste adequado é o primeiro passo para um experimento bem sucedido (ALVES, 2009).

Segundo Aguirre (AGUIRRE, 2015), o sistema a ser modelado possui os modos dominantes de sua potência espectral em uma faixa finita de frequências. Do ponto de vista deste sistema, uma entrada será considerada branca, se contiver energia suficiente na faixa de frequências dominante. Desta forma, exigir que uma entrada seja branca equivale a exigir que esta entrada excite a dinâmica do sistema. Cabe enfatizar que as dinâmicas que não forem excitadas não aparecerão no modelo.

Em sistemas reais, a faixa de frequência dominante nem sempre é conhecida. Sendo assim, o sinal de entrada utilizado na identificação deve ter potência espectral em diversas faixas de frequência. Para estimar um número,  $n$ , de parâmetros, o sinal deve possuir potência espectral em  $n$  faixas de frequência. Esta é a definição de Persistência de Excitação, PE, de ordem  $n$  (AGUIRRE, 2015).

Dois tipos de sinais muito utilizados em identificação são mostrados na Figura 16. O sinal do tipo *Pseudo Random Binary Signal*, PRBS, é bastante utilizado na identificação de sistemas lineares, uma vez que sua amplitude varia apenas entre dois valores pré-determinados. Já o sinal *Pseudo Random Multi Level Signal*, PRMLS, costuma ser aplicado na identificação de sistemas não lineares.

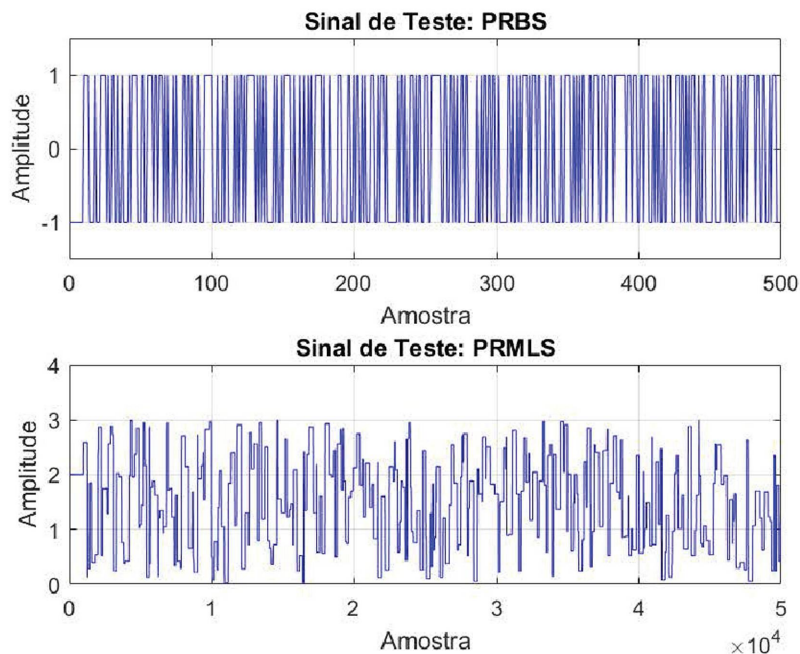


Figura 16: Sinais de teste comumente utilizados em Identificação de Sistemas: sinal PRBS e sinal PRMLS.

Ainda no que concerne à etapa de coleta de dados, a escolha do período de amostragem,  $T_s$ , exerce influência nos passos seguintes do experimento de identificação. A escolha inadequada de  $T_s$  pode causar a subamostragem ou a superamostragem do sinal. No primeiro caso, as amostras coletadas serão insuficientes para que as dinâmicas desejadas sejam modeladas. No segundo caso, as matrizes envolvidas nas etapas de determinação de estrutura, Seção 3.3, e estimação de parâmetros, Seção 3.4, podem ficar mal condicionadas, o que dificulta e pode até impossibilitar a obtenção do modelo.

Na prática, pode-se utilizar uma frequência de amostragem,  $f_s = 1/T_s$ , entre cinco e dez vezes maior que a maior frequência de interesse. Caso a suposição de superamostragem do sinal possa ser feita, existe um procedimento que auxilia na redução do número de amostras, através da decimação do sinal superamostrado. Este procedimento foi proposto por Aguirre e usa o grau de correlação entre medidas consecutivas para determinar se uma observação pode ou não ser descartada (AGUIRRE, 1995).

### 3.3 DETERMINAÇÃO DA ESTRUTURA DO MODELO

Existe uma variada gama de representações matemáticas possíveis para sistemas reais ou simulados. Neste trabalho, são de interesse as representações não lineares, algumas destas serão apresentadas na Seção 3.6. Uma vez escolhida a representação matemática, o projetista deve determinar a estrutura do modelo.

A escolha da representação matemática pode levar em consideração conhecimento prévio sobre o sistema. Por exemplo, pode-se optar por um modelo de Wiener, apresentado na Seção 3.6.3.2, para representar sistemas cuja dinâmica varie com o ponto de operação (AGUIRRE; COELHO; CORRÊA, 2005). Deve-se, também, considerar a finalidade do modelo e as ferramentas disponíveis para sua obtenção.

A determinação da estrutura, ou ordem, do modelo pode ser uma tarefa desafiadora. Caso se considere uma ordem maior do que o necessário, a estimação de parâmetros pode se tornar um problema numérico mal condicionado. Caso se considere uma ordem menor, o modelo pode deixar de representar dinâmicas importantes do sistema (AGUIRRE, 2015).

Para modelos lineares, a determinação da estrutura resume-se a encontrar o número de polos, zeros e atrasos puros de tempo (AGUIRRE, 2015). Caso a ordem do modelo seja maior do que a ordem do sistema, pode haver cancelamento de polos e zeros em sua função de transferência. Este fato pode ser utilizado na determinação da ordem adequada do modelo linear, bastando, para tanto, monitorar a posição das singularidades do modelo conforme a ordem é aumentada (BOOM; ENDEN, 1974).

Na identificação de sistemas não lineares, a sobreparametrização, elevado número de termos, causa a presença de termos redundantes no modelo. Estes termos redundantes, além de aumentar o custo computacional do modelo, podem fazer com que regimes dinâmicos espúrios estejam presentes. A sobreparametrização também pode tornar o modelo instável (AGUIRRE; BILLINGS, 1995).

Desta forma, termos redundantes em modelos lineares podem auxiliar na determinação da ordem adequada. Já, em modelos não lineares, a redundância entre os termos prejudica a modelagem da dinâmica original do sistema.

Diversas técnicas, heurísticas e determinísticas, foram propostas a fim de auxiliar na escolha da ordem de um determinado modelo. Entretanto, a determinação de

estrutura ainda é uma escolha subjetiva, que faz diversos autores considerarem esta etapa como a arte envolvida na Identificação de Sistemas (AGUIRRE, 2015).

### 3.4 ESTIMAÇÃO DE PARÂMETROS

Após a determinação da estrutura do modelo, deve-se partir para a estimação dos parâmetros dessa estrutura.

Em um problema de identificação de sistemas, estão disponíveis dados de entrada,  $u \in \mathbb{R}$ , e de saída,  $y \in \mathbb{R}$ . Estes dados relacionam-se da seguinte maneira,  $y = f(u)$ , em que  $f : \mathbb{R} \rightarrow \mathbb{R}$ . A função  $f(\cdot)$ , nos casos de interesse para este trabalho, é caracterizada por um vetor de parâmetros  $\theta$ . Assim, pode-se escrever  $y = f(u, \theta)$ . A estimação de parâmetros consiste em estimar o vetor  $\theta$  a partir dos conjuntos de medidas  $u$  e  $y$ .

A estimação dos parâmetros pode ser vista como um problema de otimização (AGUIRRE, 2015). Deste modo, deve-se optar por um algoritmo para realizar esta etapa. Algumas considerações devem ser feitas no processo de escolha do algoritmo que estimará os parâmetros de  $f(\cdot)$ . Especialmente, deve-se ter em mente o escopo de aplicação do método. O Método dos Mínimos Quadrados, MMQ, é um dos mais utilizados em identificação. No entanto, em sua forma clássica, este método apresenta resultados bons apenas para funções lineares nos parâmetros e para dados livres de ruído (AGUIRRE, 2015). Na prática, a maioria dos dados disponíveis não atende estes critérios.

Existem diversos algoritmos capazes de determinar, de forma bastante precisa, os parâmetros de um modelo. Conhecendo-se as características dos dados coletados e respeitando-se o escopo de aplicação do algoritmo, a estimação de parâmetros tende a ser uma das etapas mais simples do processo de identificação (AGUIRRE, 2015).

A estimação dos parâmetros pode ser feita de modo *on-line* ou *off-line*. No primeiro caso, os dados de entrada e saída são processados conforme ficam disponíveis, ou seja, não há armazenamento destes dados. No segundo caso, os dados são armazenados e processados em outro momento (ISERMANN; MÜNCHHOF, 2011).

O Capítulo 5 apresentará alguns algoritmos determinísticos que podem ser aplicados na estimação de parâmetros.

### 3.5 VALIDAÇÃO DO MODELO

A etapa final de um processo de modelagem caixa preta é a validação do modelo obtido. Esta etapa determinará se o modelo é bem sucedido em representar as características de interesse do sistema original. É importante mencionar que, para ser considerado válido, o modelo precisa representar, com a precisão desejada, as características fundamentais do sistema, dada a aplicação visada.

No caso de sistemas lineares, testes de correlação são utilizados para detectar dinâmicas não modeladas. Estes testes são aplicados nos resíduos, o erro que o modelo comete ao tentar explicar a medida real,  $y$ , através da estimativa,  $\hat{y}$  (LJUNG, 1999).

Para sistemas não lineares, existem testes que envolvem funções de correlação mais complexas (BILLINGS; VOON, 1986). Métodos baseados em correlação, entretanto, garantem apenas que não existem correlações não modeladas nos resíduos. Deste modo, modelos validados através destes métodos podem não seguir a dinâmica original do sistema (AGUIRRE; BILLINGS, 1995).

Para verificar se o modelo é capaz de recuperar as características dinâmicas do sistema usa-se a chamada validação dinâmica. Este método consiste em aplicar ao modelo e ao sistema um sinal persistentemente excitante, diferente do utilizado na etapa de estimação, e mensurar a qualidade da resposta do modelo. Para tanto, utilizam-se métricas como o MSE.

A validação dinâmica ganha especial destaque quando objetiva-se construir um modelo capaz de prever o comportamento do sistema em diferentes situações. A Figura 17 ilustra, em um fluxograma, as principais etapas envolvidas na modelagem caixa preta.

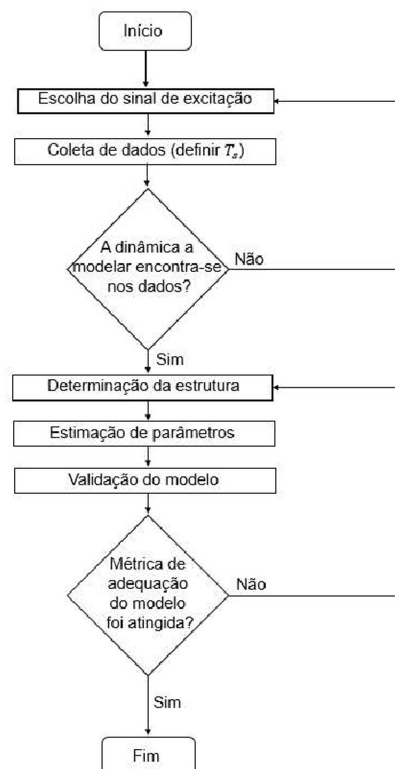


Figura 17: Fluxograma ilustrando os principais passos a serem seguidos na modelagem caixa preta.

### 3.6 REPRESENTAÇÕES NÃO LINEARES

A absoluta maioria dos sistemas dinâmicos reais é não linear. Em alguns casos, sob limites de aplicação bem definidos, aproximações lineares são capazes de fornecer bons resultados (AGUIRRE, 2015). Entretanto, a modelagem linear de um sistema não linear é, em geral, inadequada. A melhoria da precisão do modelo é um argumento a favor do uso de representações não lineares. Também deve-se considerar que representações não lineares são capazes de produzir regimes dinâmicos que modelos lineares não conseguem representar, o que é de fundamental importância em aplicações de previsão de comportamento, por exemplo.

Existe uma grande variedade de representações não lineares que podem ser aplicadas à Identificação de Sistemas. O objetivo desta seção é apresentar, de forma resumida, as representações consideradas durante a elaboração do modelo caixa preta para o transistor mono-elétron.

### 3.6.1 Série de Volterra

Considere a Equação (3.1), conhecida como equação de convolução:

$$y(k) = \sum_{\tau_1}^{\infty} h_1(\tau_1)u(k - \tau_1), \quad (3.1)$$

em que  $k, \tau_1 \in \mathbb{Z}$  e  $u, h_1 \in \mathbb{R}$ , sendo  $h_1$  a resposta ao impulso.

A série de Volterra generaliza o conceito de convolução para os sistemas não lineares (ISERMANN; MÜNCHHOF, 2011). Um sistema não linear pode ter sua saída  $y(k)$  relacionada com sua entrada  $u(k)$  através da série de Volterra, definida como (VOLTERRA, 1959):

$$y(k) = \sum_{\tau_1}^{\infty} \sum_{\tau_2}^{\infty} \cdots \sum_{\tau_j}^{\infty} h_j(\tau_1, \dots, \tau_j) \prod_{i=1}^j u(k - \tau_i), \quad (3.2)$$

em que  $i, j \in \mathbb{N}$  e as funções  $h_j \in \mathbb{R}$  são generalizações não lineares da resposta ao impulso  $h_1$ . As funções  $h_j$  são conhecidas como núcleos, ou *kernels*, de Volterra. Para um sistema linear, faz-se  $j = 1$  em (3.2) e obtém-se (3.1).

A Figura 18 ilustra o diagrama de blocos de um modelo de Volterra, considerando os operadores  $H_1, H_2, \dots, H_n$  descritos pela Equação (3.4).

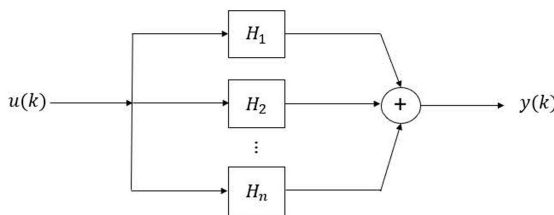


Figura 18: Esquemático de um sistema representado por um modelo de Volterra (ROSA, 2009).

A saída do sistema, considerando-se a Figura 18, pode ser escrita da seguinte forma:

$$y(k) = H_1[u(k)] + H_2[u(k)] + \cdots + H_n[u(k)], \quad (3.3)$$

em que,

$$H_n[u(k)] = \sum_{\tau_1}^{\infty} \sum_{\tau_2}^{\infty} \cdots \sum_{\tau_n}^{\infty} h_n(\tau_1, \dots, \tau_n) u(k - \tau_1) \cdots u(k - \tau_n). \quad (3.4)$$

Uma vez que o termo  $H_n[u(k)]$  em (3.3) inclui contribuições de grau  $n$  para a entrada  $u(k)$ , tem-se que o modelo de Volterra é uma generalização não linear polinomial do modelo linear em (3.1) (ROSA; CAMPELLO; AMARAL, 2009).

Dentre as principais vantagens do modelo de Volterra são frequentemente citadas a capacidade de representação e o fato de o modelo não precisar de realimentação da saída. A falta desta realimentação, no entanto, é a causa da principal desvantagem do modelo: o alto número de termos necessários para que a aproximação atinja precisão adequada. Uma vez que o número de termos é elevado, o número de parâmetros a serem determinados também é alto. Séries de Volterra não são capazes de modelar comportamentos como histerese e caos (NEMETH; KOLLAR, 2002).

Na próxima seção, serão apresentados os modelos NARMAX polinomiais, que podem ser vistos como uma generalização da Série de Volterra.

### 3.6.2 NARMAX Polinomial

Os modelos *Nonlinear AutoRegressive Moving Average with eXogenous input*, NARMAX, descrevem um sistema a partir de estimativas dos sinais de entrada, de saída e do ruído (AGUIRRE, 2015):

$$y(k) = f^l[u(k - \tau_d), \dots, u(k - n_u), y(k - 1), \dots, y(k - n_y), e(k - 1), \dots, e(k - n_e)] + e(k), \quad (3.5)$$

em que a sequência  $e(k) \in \mathbb{R}$  indica todos os efeitos que não podem ser bem representados por  $f^l[\cdot]$ ,  $e(k)$  é chamado de ruído.  $u(k)$ ,  $y(k) \in \mathbb{R}$  são os sinais de entrada e saída e  $\tau_d \in \mathbb{Z}$  é o tempo morto ou atraso puro de tempo,  $n_u$ ,  $n_y$  e  $n_e \in \mathbb{Z}$  são os atrasos máximos, ou ordem de regressão, da entrada, da saída e do ruído, respectivamente, e  $l \in \mathbb{N}$  é o grau de não linearidade da função  $f^l[\cdot]$ .

O modelo apresentado em (3.5) é uma versão geral para todos os modelos da classe NARMAX. A determinação da função  $f^l[\cdot]$  é uma das dificuldades inerentes à obtenção do modelo. Em modelos NARMAX polinomiais, a função  $f^l[\cdot]$  é polinomial.

Os modelos NARMAX possuem alta capacidade de representação (BILLINGS, 2013). O problema de dimensionalidade, relacionado ao alto número de termos ne-

cessários ao modelo, não é tão crítico nos modelos NARMAX quanto nos modelos de Volterra. Este fato se deve à inclusão de termos de saída e de ruído na modelagem. A inclusão do ruído na modelagem ajuda a evitar polarização na etapa de estimação de parâmetros, este problema será discutido na Seção 5.6.

Os modelos de Volterra podem ser considerados casos particulares dos modelos NARMAX, em que os termos relativos à saída e ao ruído foram suprimidos, e a modelagem é obtida apenas com regressores de entrada (AGUIRRE, 2015).

Os modelos NARX, são, como a série de Volterra, um caso particular dos modelos NARMAX. Enquanto na série de Volterra os regressores de saída e de ruído são retirados do modelo, nos modelos NARX apenas os regressores do ruído são retirados. A equação geral da saída de um sistema estimada por um modelo NARX é expressa em (3.6) (AGUIRRE, 2015).

$$y(k) = \sum_{m=0}^l \sum_{p=0}^m \sum_{\tau_1, \tau_m}^{n_y, n_u} c_{p,m}(\tau_1, \dots, \tau_m) \prod_{j=1}^p y(k - \tau_j) \prod_{i=1}^m u(k - \tau_{p+i}) + e(k), \quad (3.6)$$

em que  $c_{p,m} \in \mathbb{R}$  são os coeficientes do modelo,  $\tau_i \in \mathbb{N}^+$  e

$$\sum_{\tau_1, \tau_m}^{n_y, n_u} = \sum_{\tau_1=1}^{n_y} \cdots \sum_{\tau_p=1}^{n_y} \sum_{\tau_{p+1}=\tau_d}^{n_u} \cdots \sum_{\tau_{p+m}=\tau_d}^{n_u} . \quad (3.7)$$

A Equação (3.6) expressa que os termos de um modelo NARX podem possuir graus de não linearidade  $m$ , tal que  $1 \leq m \leq l$ . Desta forma, cada termo de grau  $m$  pode conter um fator de grau  $p$  do tipo  $y(k - \tau_i)$  e um fator de grau  $(m - p)$  do tipo  $u(k - \tau_i)$ . Em (3.7), o limite superior será  $n_y$ , caso o somatório se refira a regressores de saída. Da mesma forma, o limite superior será  $n_u$ , caso o somatório se refira a regressores de entrada.

### 3.6.3 Modelos de Blocos Concatenados

É comum que se busque a maior simplicidade possível ao se desenvolver um modelo para um dado sistema (ASTRÔM; HÄGGLUND, 2011). Isto se deve à facilidade de utilização do modelo e à procura pela minimização do custo computacional.

Na modelagem de sistemas não lineares, uma das abordagens mais simples consiste em concatenar um modelo dinâmico e uma não linearidade estática (AGUIRRE, 2015; ISERMANN; MÜNCHHOF, 2011).

Nesta seção, serão apresentados os modelos baseados em blocos concatenados de Hammerstein, Wiener e Hammerstein-Wiener. Como os modelos de blocos concatenados elaborados neste trabalho fazem uso da não linearidade *PieceWise Linear*, PWL, esta função é apresentada na Seção 3.6.3.4.

### 3.6.3.1 Modelo de Hammerstein

Nos modelos de Hammerstein, um sistema é representado através de uma não linearidade estática,  $f_H^l(\cdot)$ , seguida por um modelo dinâmico linear (AGUIRRE, 2015). Este modelo dinâmico linear é do tipo *AutoRegressive with eXogenous input*, ARX. Os modelos ARX são versões lineares dos modelos NARX, vistos na Seção 3.6. Desta forma, pode-se representar um modelo ARX através da Equação (3.8).

$$y(k) = \sum_{i=\tau_d}^{n_u} b_i u(k-i) + \sum_{j=1}^{n_y} a_j y(k-j), \quad (3.8)$$

em que  $b_i$  e  $a_j \in \mathbb{R}$  são os parâmetros do modelo.

A representação em diagrama de blocos de um modelo de Hammerstein pode ser observada na Figura 19.

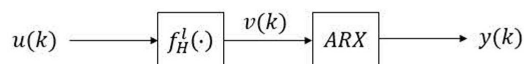


Figura 19: Modelo de Hammerstein. Os sinais  $u(k)$ ,  $y(k)$  e  $v(k)$  representam a entrada, a saída e o sinal interno, respectivamente. O sinal  $v(k)$  é, em geral, desconhecido e sua medição não é possível.

Pode-se considerar que os modelos de Hammerstein são casos particulares dos modelos NARX em que a função não linear é aplicada apenas à entrada, conforme (3.9).

$$y(k) = f_H^l[b_0 u(k - \tau_d) + \dots + b_{n_u} u(k - n_u)] + a_1 y(k - 1) + \dots + a_{n_y} y(k - n_y). \quad (3.9)$$

A determinação da função  $f_H^l(\cdot)$  faz parte da escolha da estrutura do modelo. Uma vez que se conheça esta função, o sinal interno  $v(k)$  pode ser obtido através da aplicação da não linearidade ao sinal de entrada,  $v(k) = f_H^l[u(k)]$ . O modelo ARX pode ser obtido, a partir do conhecimento dos sinais  $v(k)$  e  $y(k)$ , através de técnicas lineares (AGUIRRE; COELHO; CORRÊA, 2005).

### 3.6.3.2 Modelo de Wiener

De forma dual aos modelos de Hammerstein, os modelos de Wiener são representações não lineares em que o modelo dinâmico linear, ARX, é seguido por uma não linearidade estática,  $f_W^l(\cdot)$  (AGUIRRE, 2015). A representação em diagrama de blocos de um modelo de Wiener pode ser observada na Figura 20.

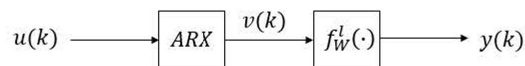


Figura 20: Modelo de Wiener. Os sinais  $u(k)$ ,  $y(k)$  e  $v(k)$  representam a entrada, a saída e o sinal interno, respectivamente. O sinal  $v(k)$  é, em geral, desconhecido e sua medição não é possível.

Da mesma forma que os modelos de Hammerstein, os modelos de Wiener podem ser considerados casos particulares de modelos NARX. Entretanto, nos modelos de Wiener, a não linearidade estática age na saída do modelo ARX. É comum, então, expressar o modelo de Wiener através da Equação (3.10), em que a função inversa da não linearidade,  $g_W = (f_W^l)^{-1}$ , sendo que  $g_W : \mathbb{R} \rightarrow \mathbb{R}$ , age nos valores passados da saída do sistema.

$$y(k) = f_W^l[b_0 u(k - \tau_d) + \dots + b_{n_u} u(k - n_u) + g_W(a_1 y(k - 1) + \dots + a_{n_y} y(k - n_y))], \quad (3.10)$$

em que considerou-se que  $f_W^l(\cdot)$  é uma função inversível.

Analisando a Figura 20, nota-se que a função estática  $f_W^l(\cdot)$  expressa a relação estática entre o sinal interno  $v(k)$  e a saída  $y(k)$ . A relação estática entre a entrada  $u(k)$  e a saída  $y(k)$  também pode ser expressa diretamente por  $f_W^l(\cdot)$ , desde que o modelo ARX tenha ganho unitário (AGUIRRE; COELHO; CORRÊA, 2005). Considerando o modelo ARX com ganho unitário e o sistema em estado estacionário, pode-se usar a

função não linear para traçar um gráfico  $\bar{y} - \bar{u} = \bar{v}$ , sendo que  $\bar{y}$ ,  $\bar{u}$  e  $\bar{v}$  indicam os sinais em estado estacionário. Deste modo, pode-se usar regressão linear para se obter  $\bar{v}$  e, por fim, usando  $\bar{v} = g_W(\bar{y})$  encontra-se uma estimativa de  $(f_W^l)^{-1}$ .

### 3.6.3.3 Modelo de Hammerstein-Wiener

Como fusão dos modelos de Hammerstein e de Wiener, surge o modelo de Hammerstein-Wiener. Este modelo é composto por duas não linearidades estáticas permeadas por um sistema dinâmico linear (ISERMANN; MÜNCHHOF, 2011).

Observando as figuras 19 e 20, pode parecer, intuitivamente, que as representações de Wiener e de Hammerstein são equivalentes. Entretanto, é possível provar analiticamente que os autovalores da matriz Jacobiana da função de saída dos modelos de Hammerstein são constantes, enquanto que, nos modelos de Wiener, estes autovalores são dependentes do ponto de operação (AGUIRRE; COELHO; CORRÊA, 2005).

A representação em diagrama de blocos de um modelo de Hammerstein-Wiener pode ser observada na Figura 21.

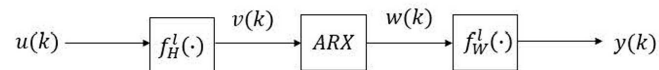


Figura 21: Modelo de Hammerstein-Wiener. Os sinais  $u(k)$ ,  $y(k)$ ,  $v(k)$  e  $w(k)$  representam a entrada, a saída e os sinais internos, respectivamente. Os sinais  $v(k)$  e  $w(k)$  são, em geral, desconhecidos e sua medição não é possível.

Para expressar o modelo de Hammerstein-Wiener, considere as equações (3.11) e (3.12).

$$v(k) = f_H[u(k)]. \quad (3.11)$$

$$w(k) = g_W[y(k)]. \quad (3.12)$$

Assim, a saída de um sistema expresso por um modelo de Hammerstein-Wiener pode ser escrita como se segue,

$$y(k) = f_W[g_W(a_1 y(k-1) + \dots + a_{n_y} y(k-n_y)) + f_H(b_0 u(k-\tau_d) + \dots + b_{n_u} u(k-n_u))], \quad (3.13)$$

em que o grau de não linearidade  $l$  foi omitido por simplicidade.

A próxima seção dedica-se a apresentar a não linearidade estática que foi aplicada nos modelos orientados a blocos desenvolvidos neste trabalho.

#### 3.6.3.4 Não linearidade estática: *PieceWise-Linear*

A característica linear por partes é encontrada com frequência em sistemas que operam com ganhos diferentes para valores diferentes de entrada (VÖRÖS, 2007). A Figura 22 apresenta um exemplo desta característica. A Equação (3.14) expõe sua modelagem matemática.

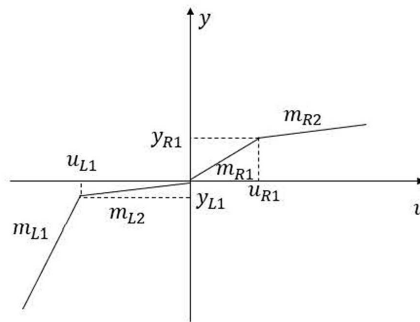


Figura 22: Exemplo de característica PWL com  $(2 + 2)$ -segmentos, modificada de (VÖRÖS, 2007).

$$y = \begin{cases} m_{R1}u, & \text{se } 0 \leq u \leq u_{R1}, \\ m_{R2}(u - u_{R1}) + m_{R1}u_{R1}, & \text{se } u > u_{R1}, \\ m_{L2}u, & \text{se } u_{L1} \leq u < 0, \\ m_{L1}(u - u_{L1}) + m_{L2}u_{L1}, & \text{se } u < u_{L1}, \end{cases} \quad (3.14)$$

em que  $m_{L1}$ ,  $m_{L2}$ ,  $m_{R1}$  e  $m_{R2}$  são os coeficientes angulares dos segmentos lineares,  $u_{L1}$  é a constante para entradas negativas e  $u_{R1}$  é a constante para entradas positivas (TAO; TIAN, 1998).

### 3.7 CONSIDERAÇÕES FINAIS

Neste capítulo, foram abordados os conceitos de modelo e de sistema, além dos conceitos iniciais necessários à Identificação de Sistemas. O procedimento padrão para a obtenção de um modelo caixa preta foi dividido nas etapas de coleta de dados, determinação de estrutura, estimação de parâmetros e validação. Apresentou-se, também, uma revisão geral de cada uma destas etapas e das considerações que o projetista deve fazer para desempenhar satisfatoriamente cada passo e obter um bom modelo. A não-linearidade estática do tipo PWL, que será aplicada nos modelos de blocos concatenados, foi brevemente explanada. Por fim, foram apresentadas algumas representações não lineares. As representações aqui resumidas buscam familiarizar o leitor com as classes de modelos que serão testadas para o SET. Os principais conceitos envolvendo os modelos de Volterra, NARMAX polinomiais, NARX polinomiais, de Hammerstein, de Wiener e de Hammerstein-Wiener foram apresentados.

## 4 MODELAGEM CAIXA BRANCA DO SET

Para que a implementação física de um circuito seja desenvolvida, ferramentas de simulação precisas e rápidas são fundamentais (WASSHUBER, 2001). Desta forma, diversos métodos e ferramentas de simulação para circuitos nanoeletrônicos foram propostos. Dentre os de maior destaque, pode-se citar: *SIMmulation Of Nanostructures*, SIMON (WASSHUBER; KSOINA; SELBERHERR, 1997), *KOrea Single Electron Circuit Simulator for Spice*, KOSEC (YU et al., 2000), *Single Electron NanoElectronic Circuit Analyzer*, SENECA (FONSECA; KOROTKOV; LIKHAREV, 1995), *MOnte Carlo Single Electron Simulator*, MOSES (CHEN, 1996), *Single Electron Circuit Simulator*, SECS (ZARDALIDIS; KARAFYLLIDIS, 2008) e *Single Electron Multiscale SIMulation*, SEMSIM (ALLEC; KNOBEL; SHANG, 2008).

Além de ferramentas de simulação adequadas, modelos precisos e capazes de prever o comportamento do dispositivo físico são fundamentais.

Conforme discutido nos capítulos 1 e 2, as características únicas do transistor mono-elétron fazem dele objeto de estudo constante. Contudo, o SET possui algumas desvantagens que dificultam sua aplicação em circuitos práticos. Dentre estas, pode-se citar o baixo ganho de tensão e a alta impedância de saída.

O ganho de tensão em um SET é definido pela razão entre a capacitância de *gate* e a capacitância da junção (KIRIHARA; NAKAZATO; WAGNER, 1999). Como estas capacitâncias costumam ter valores similares, o ganho de tensão do transistor fica próximo, ou até abaixo, da unidade.

A alta impedância de saída se deve às resistências de tunelamento, ver Seção 2.1.4. Estas resistências também fazem com que a corrente conduzida pelo SET seja extremamente baixa. Isto pode ser vantajoso quando se considera a potência dissipada, mas também pode causar atrasos na saída de um circuito (LIENTSCHNIG; WEYMANN; HADLEY, 2003).

Uma maneira de superar estas desvantagens é combinar o SET com transistores MOS. Diversas propostas e implementações destes circuitos híbridos já foram realizadas (OHATA; TORIUMI; UCHIDA, 1997; MAHAPATRA et al., 2003; GEROUSIS; GREPIOTIS, 2013; JANA et al., 2013). Para estes circuitos, no entanto, deve-se tomar cuidado com a tensão de polarização do SET. Estudos mostram que, para  $|V_{DS}| > \frac{e}{C_{ilha}}$  (ver Seção 2.1.5.1), o bloqueio de Coulomb não é mais observado (RADWAN et al., 2015). Uma vez que os transistores MOS são, em geral, polarizados em tensões maiores, em circuitos híbridos,  $V_{DS}$  pode assumir valores maiores que  $e/C_{ilha}$  (JAIN; NAMERIAKPAM; SARKAR, 2015).

Deste modo, para que as peculiaridades de operação e desempenho do SET sejam estudadas, modelos precisos e eficientes são necessários. Neste capítulo, os principais métodos de simulação e modelagem de circuitos nanoeletrônicos serão apresentados. Posteriormente, a aplicação destes métodos é exemplificada por meio de uma revisão de modelos caixa branca já propostos para o SET.

## 4.1 MÉTODOS DE MODELAGEM E SIMULAÇÃO

Os dois principais métodos de modelagem e simulação para dispositivos nanoeletrônicos são a **Equação Mestra**, EM, e o método de **Monte Carlo**, MC. As próximas seções dedicam-se a explicar, de forma resumida, as principais considerações envolvidas nestas metodologias.

### 4.1.1 Equação Mestra

Conforme visto na seção 2.1.3, a teoria ortodoxa do tunelamento mono-elétron é capaz de calcular a taxa de tunelamento para um elétron através de uma barreira de potencial. Em um circuito mono-elétron, no entanto, deseja-se estabelecer uma corrente através da junção-túnel ou da ilha. Esta corrente é formada por diversos tunelamentos sucessivos.

A EM auxilia no cálculo da corrente através da junção, ou através da ilha em um SET, utilizando conceitos de mecânica estatística e de processos estocásticos (WASSHUBER, 2001). Para tanto, são consideradas as probabilidades de ocupação dos estados e as taxas de tunelamento para cada transição possível.

As seguintes suposições são feitas no algoritmo da EM:

- A taxa de tunelamento de um elétron depende apenas do estado atual do sistema, ou seja, ela não é afetada pelo conhecimento de estados anteriores do sistema;
- O sistema evolui em saltos, i.e., os estados são discretos.

A primeira suposição é a chamada propriedade markoviana e descreve a característica de um sistema no qual não há efeito de memória (PAPOULIS; PILLAI, 2002), ou seja, o estado futuro do processo depende apenas de seu estado atual.

A fim de reduzir o custo computacional da simulação, o algoritmo da EM considera que apenas transições entre estados vizinhos são possíveis. É válido apontar que, na EM, o tempo de simulação é linearmente dependente do número de estados considerados (RADWAN et al., 2015). Processos em que ocorrem somente transições entre estados vizinhos são conhecidos como processos *birth-death* (WASSHUBER, 2001; PAPOULIS; PILLAI, 2002). Esta simplificação assume que apenas um elétron tunela por vez. Desta forma, co-tunelamentos são inerentemente desconsiderados (HASANEEN; WAHAB; AHMED, 2011).

Da teoria ortodoxa que descreve o tunelamento mono-elétron, Seção 2.1.3, a taxa de tunelamento de um elétron através de uma barreira de potencial é dada pela Equação (4.1).

$$\Gamma = -\frac{\Delta F}{e^2 R_T \left[ 1 - e^{\frac{\Delta F}{k_B T}} \right]}, \quad (4.1)$$

na qual  $\Delta F$  é a variação total na energia eletrostática do sistema (causada pelo tunelamento),  $R_T$  é a resistência de tunelamento da junção, em Ohm,  $k_B$  é a constante de Boltzmann e  $T$  é a temperatura, dada em Kelvin.

Pode-se, então, escrever a EM conforme (4.2) (WASSHUBER, 2001). Os termos  $\Gamma_{ij}$  e  $\Gamma_{ji}$  expressam a taxa de transição do estado  $j$  para o estado  $i$  e vice-versa.  $P_i(t)$  é a probabilidade de ocupação do estado  $i$  no instante  $t$ .

$$\frac{\partial P_i(t)}{\partial t} = \sum_{j \neq i} [\Gamma_{ij} P_j(t) - \Gamma_{ji} P_i(t)], \quad (4.2)$$

em que  $i$  e  $j \in \mathbb{Z}$ . Os estados  $i$  e  $j$  são caracterizados pelo número de elétrons na ilha.

A Figura 23 ilustra o diagrama de transição para o processo *birth-death* descrito pela Equação (4.2).

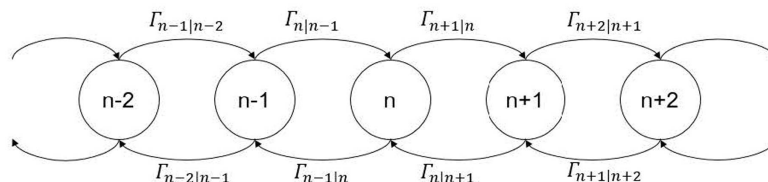


Figura 23: Ilustração das transições possíveis em processo *birth-death* com número generalizado de estados - modificada de (WASSHUBER, 2001), em que  $n$  é o número de elétrons na ilha.

A aplicação da Equação (4.1) no cálculo das taxas de tunelamento, exige que  $\Delta F$  seja conhecida. Existem três formas bastante difundidas para o cálculo de  $\Delta F$ : o método de Hadley e Lientschnig (HADLEY; LIENTSCHNIG; LAI, 2002), o método da tensão crítica (GEERLIGS et al., 1990) e através do cálculo da variação da energia livre de Gibbs, também conhecido como abordagem termodinâmica. O método de Hadley e Lientschnig é o mais frequente na literatura que se refere à modelagem de transistores mono-elétron. Sendo assim, ele será explanado no Anexo A. Para maiores informações sobre os demais métodos, consulte (WASSHUBER, 2001).

Uma vez que as taxas de tunelamento tenham sido computadas, a solução da EM ainda depende do cálculo das probabilidades de ocupação dos estados. Este cálculo aplica a forma matricial da Equação (4.2), os passos seguidos para que se obtenha as probabilidades de ocupação dos estados são descritos no Anexo A.

Finalmente, após computar as probabilidades de ocupação dos estados e as taxas de tunelamento, a corrente será dada por:

$$I = e \sum_{n=-\infty}^{+\infty} P(n) \Gamma_X(n), \quad (4.3)$$

em que  $I$  é a corrente média fluindo no componente,  $e$  é a carga elementar,  $P(n)$  é a probabilidade de ocupação do estado  $n$  e  $\Gamma_X(n)$  é a taxa de tunelamento total para o estado  $n$  na junção  $X$  de interesse. A taxa de tunelamento total, na junção de interesse  $X$ , é definida em (4.4).

$$\Gamma_X(n) = \Gamma_X^+(n) - \Gamma_X^-(n), \quad (4.4)$$

em que,  $\Gamma_X^+(n)$  é a taxa de tunelamento através da junção da esquerda para a direita e  $\Gamma_X^-(n)$  é a taxa de tunelamento através da junção da direita para a esquerda. A Figura 24 ilustra a ideia.

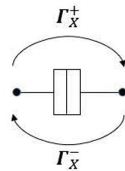


Figura 24: Taxas de tunelamento em uma junção  $X$  de interesse.

O primeiro passo para solucionar (4.2) é determinar o conjunto de estados relevantes. Existe um *trade off* entre precisão e custo computacional nesta determinação. Quanto maior o número de estados considerados, mais precisa e mais lenta será a simulação.

Deste modo, uma maneira frequente de dar início ao cálculo é considerar um estado inicial, geralmente  $n = 0$ , a partir do qual as taxas de todas as transições possíveis serão calculadas. Estas taxas são, então, utilizadas como estimativa inicial das probabilidades de estado. Outra prática comum é estabelecer um valor de limiar para  $\Gamma$ . Estados com taxas de transição abaixo deste limiar são desconsiderados, uma vez que possuem baixa probabilidade de ser atingidos. Este processo é iterativo e, para cada iteração, novos estados podem ser adicionados ao conjunto considerado (WASSHUBER, 2001).

A Figura 25 apresenta um fluxograma que detalha os passos a serem seguidos no algoritmo da Equação Mestra.

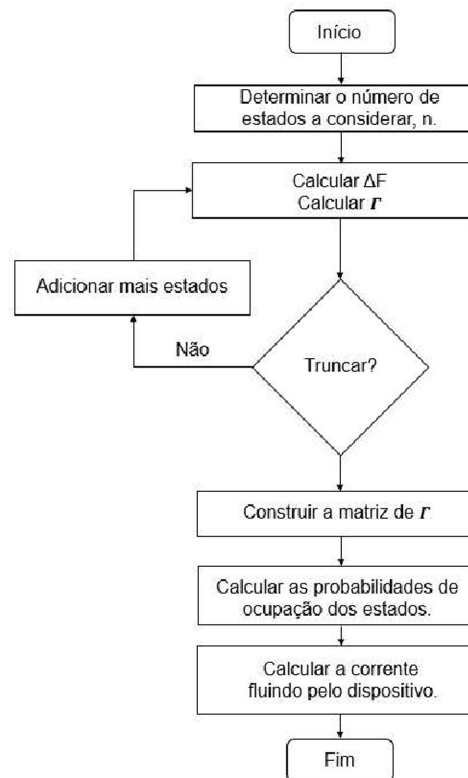


Figura 25: Fluxograma apresentando os passos para a obtenção da corrente elétrica a partir da EM - modificada de (HASANEEN; WAHAB; AHMED, 2011).

#### 4.1.2 Monte Carlo

Nomeado em alusão ao cassino homônimo em Mônaco, o método de MC utiliza amostragem aleatória para obter resultados numéricos. Este método aplica as leis da probabilidade e da estatística às ciências naturais. A essência do método é utilizar distribuições de números aleatórios a fim de obter amostras que aproximem o comportamento do sistema como um todo (ANDERSON, 1986).

O método de MC foi apresentado à comunidade científica pelos matemáticos Stanislaw Ulam e Nicholas Metropolis, em um artigo publicado em 1949 (METROPOLIS; ULAM, 1949). Em 1946, enquanto jogava Paciência, Ulam tentou utilizar análise combinatória para determinar a probabilidade de sucesso no jogo. Enquanto tentava chegar a uma solução analítica, ele pensou em simplesmente realizar, por exemplo, 100 experimentos e contar o número de tentativas bem sucedidas. Posteriormente, o método de MC seria utilizado por Stanislaw Ulam e John von Neumann para simular o problema da difusão aleatória de nêutrons durante a fissão nuclear.

Em nanoeletrônica, o método de MC é utilizado no *software* SIMON, considerado um dos mais precisos atualmente (RADWAN et al., 2015). Conforme será visto na Seção 4.2, os resultados das simulações feitas no SIMON são frequentemente utilizados como *benchmark* em artigos que propõem novas formas de simulação para circuitos mono-elétron. A explicação dada a seguir, sobre o método de MC, baseia-se no livro de Christoph Wasshuber, desenvolvedor do SIMON (WASSHUBER, 2001).

No SIMON, a junção-túnel é descrita de forma estocástica através de processos de Poisson. Um processo de Poisson caracteriza-se pela distribuição exponencial do tempo entre eventos. Desta forma, a probabilidade de um tunelamento ocorrer a partir do estado inicial no instante  $\tau$ , e não antes, é dada pela Equação (4.5).

$$P_0(\tau) = e^{-\Gamma\tau}, \quad (4.5)$$

em que  $\Gamma$  é a taxa de tunelamento descrita na Equação (4.1).

A aplicação de MC exige um gerador de números aleatórios uniformemente distribuídos no intervalo  $[0, 1]$ . Neste caso, para construir números aleatórios com distribuição descrita por  $P_0(\tau)$  a partir de números com distribuição uniforme, faz-se o inverso da distribuição  $P_0(\tau)$  (ECKHARDT, 1987):

$$\tau = -\frac{\ln(r)}{\Gamma}, \quad (4.6)$$

sendo  $r$  um número aleatório com distribuição uniforme no intervalo  $[0, 1]$ .

A simulação através de MC começa a partir de uma lista com os eventos de tunelamento possíveis e suas respectivas taxas. Os tempos de tunelamento,  $\tau_i$ , são computados para todos os eventos de acordo com (4.6). O evento com o menor valor de  $\tau$  ocorre primeiro e, portanto, é escolhido como vencedor. A corrente é calculada diretamente através da razão entre carga transportada e intervalo de tempo, conforme a Equação (2.1).

Este procedimento é repetido até que um dos limites definidos seja atingido: tempo, precisão ou número máximo de eventos. Desta forma, o comportamento macroscópico do circuito é simulado. A Figura 26 ilustra em um fluxograma os passos da simulação de Monte Carlo efetuada no SIMON.

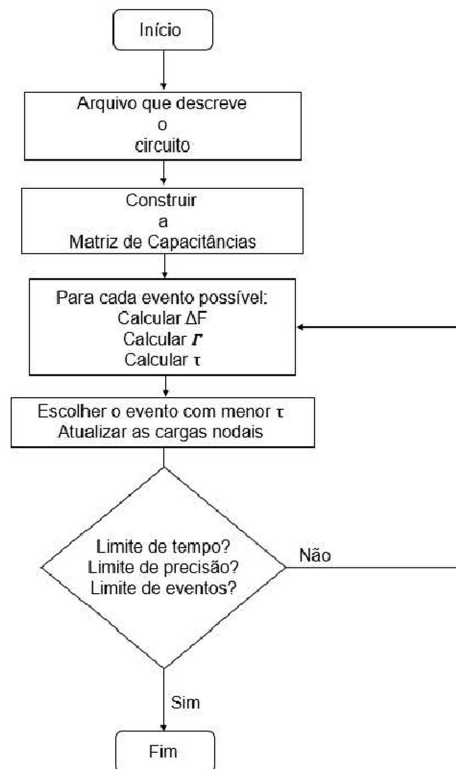


Figura 26: Fluxograma para o método de MC - modificada de (WASSHUBER, 2001).

Na simulação de um circuito nanoeletrônico, o método de MC simula o transporte de elétrons através de tunelamentos, considerando sua natureza estocástica. A corrente é dada pela transferência total de carga no tempo. Desta forma, se um evento de tunelamento nunca for selecionado, devido à sua natureza rara, ele não contribuirá para a corrente. Esta particularidade acaba gerando imprecisões nos resultados da simulação. A solução, proposta por Wasshuber, consiste em calcular a corrente de forma diferente: em vez de considerar apenas o evento vencedor, todos os eventos contribuem para a corrente de acordo com sua probabilidade de ocorrência.

Uma vez que as taxas de tunelamento,  $\Gamma$ , já são calculadas no algoritmo de MC, a probabilidade de ocorrência de tunelamento do estado  $k$  para o estado  $j$  será dada por (4.7).

$$P_{jk} = \frac{\Gamma_{jk}}{\sum_i \Gamma_{ik}}, \quad (4.7)$$

em que  $P_{jk}$  é a probabilidade de transição do estado  $k$  para o estado  $j$  e  $\Gamma_{jk}$  é a taxa de tunelamento do estado  $k$  para o estado  $j$ .

Quanto ao transporte de carga no circuito, esta alteração no cálculo da corrente equivale a “espalhar” a carga por todos os eventos de tunelamento, em vez de considerar que apenas o evento vencedor transporta toda a carga. A Figura 27 ilustra esta ideia.

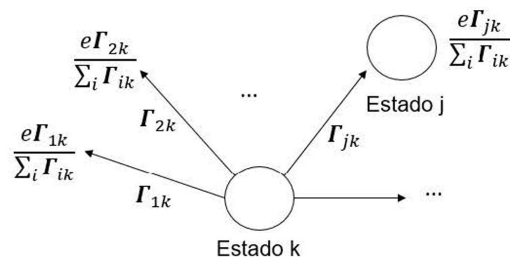


Figura 27: A carga elementar,  $e$ , é espalhada por vários eventos de tunelamento, de acordo com sua probabilidade de ocorrência - modificada de (WASSHUBER, 2001).

## 4.2 MODELOS CAIXA BRANCA PARA O SET

Os modelos já propostos para o transistor mono-elétron podem ser divididos em dois grandes grupos: modelos analíticos e modelos macro para SPICE. Os primeiros baseiam-se na EM do tunelamento mono-elétron. Já os modelos macro fazem uso de dispositivos convencionais para construir circuitos equivalentes para o SET.

### 4.2.1 Modelos Analíticos

Conforme discutido na Seção 4.1.1, o primeiro passo para resolver a EM é determinar o número de estados de interesse. Este número determinará a precisão e o esforço computacional demandado pelo modelo. Desta forma, pode-se considerar que a modelagem analítica do SET busca encontrar um número de estados ótimo, que forneça precisão e seja pouco custoso computacionalmente.

A Figura 28 ilustra um transistor mono-elétron e as taxas de tunelamento de interesse.

Conforme ilustrado na Figura 28, existem quatro eventos de tunelamento possíveis em um SET: uma carga,  $e$ , pode tunelar da fonte para ilha ( $\Gamma_S^+$ ), da ilha para a fonte

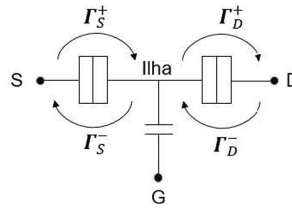


Figura 28: Diagrama exibindo as taxas de tunelamento para um elétron através do SET. As letras S, G e D marcam os terminais de fonte, porta e dreno, respectivamente - modificada de (HASANEEN; WAHAB; AHMED, 2011).

( $\Gamma_S^-$ ), da ilha para o dreno ( $\Gamma_D^+$ ) ou do dreno para a ilha ( $\Gamma_D^-$ ). A corrente no SET é dada por (4.3), que pode ser reescrita como:

$$I_{SET} = e \sum_n P(n) (\Gamma_D^+ - \Gamma_D^-), \quad (4.8)$$

em que foi considerada a corrente através da junção-túnel de dreno. Como a corrente que percorre a junção de dreno é igual à que percorre a junção da fonte, pode-se escrever a equação utilizando quaisquer das taxas.

O desenvolvimento de modelos analíticos para o SET se deu pela necessidade de estudar e simular suas características. A técnica de simulação numérica, baseada no método de MC, era pouco eficiente na simulação de circuitos grandes (UCHIDA et al., 2000).

Os primeiros modelos analíticos foram propostos no início da década de 90 (GLAZMAN; SHEKHTER, 1989; BEENAKKER, 1991). Estes modelos buscavam a reprodução de características intrínsecas do SET, como o bloqueio de Coulomb. Contudo, embora pioneiros, eles são bastante simples e incapazes de reproduzir o comportamento do transistor em condições de interesse para o desenvolvimento de circuitos integrados. Especificamente, nos primeiros modelos analíticos, a tensão  $V_{DS}$  é considerada nula, o que não ocorre na maioria das aplicações do SET (UCHIDA et al., 2000). Quanto maior for a tensão de polarização do transistor, mais estados devem ser considerados para que o modelo apresente um grau aceitável de precisão. Ou seja, quanto maior for  $V_{DS}$ , mais equações devem ser resolvidas e maior é a complexidade do modelo.

Uchida *et al.* propuseram, pela primeira vez, um modelo analítico em que  $V_{DS}$  não era nula (UCHIDA et al., 2000). Para desenvolver seu trabalho, os autores fizeram as seguintes suposições:

- Os terminais de fonte e dreno do SET são conectados a fontes de tensão ou a capacitores cuja capacitância é muito maior que a capacitância da ilha do transistor;
- As junções-túnel de fonte e dreno possuem parâmetros idênticos;
- Para cada valor de  $V_{GS}$ , os dois números mais prováveis de elétrons na ilha são considerados (0 ou 1).

A primeira suposição garante que a característica do SET seja afetada por componentes vizinhos apenas através das tensões nodais dos terminais do transistor. Esta suposição é assumida como verdadeira em todos os modelos propostos. A segunda suposição simplifica as equações relativas ao modelo, restringindo sua validade a dispositivos simétricos. A terceira suposição também faz com que o modelo seja simplificado. Além disso, ela impõe uma restrição ao intervalo máximo de valores para  $V_{DS}$  no qual a validade do modelo se mantém:

$$|V_{DS}| \leq \frac{e}{C_{ilha}}, \quad (4.9)$$

$$C_{ilha} = C_D + C_S + C_{g1} + C_{g2}, \quad (4.10)$$

em que  $e$  é a carga elementar e  $C_D$ ,  $C_S$ ,  $C_{g1}$  e  $C_{g2}$  são as capacitâncias das junções de dreno, de fonte e dos capacitores de *gate*, respectivamente. Para o modelo de Uchida,  $C_{g2} = 0$  F, uma vez que os autores consideraram dispositivos com apenas um *gate*.

Para validar a funcionalidade de seu modelo, Uchida comparou as curvas  $I_{DS}$ - $V_{GS}$  obtidas através de seu modelo e através do simulador *Circuit Analysis program including a Model of Single Electron Tunneling*, CAMSET (KIRIHARA; NAKAZATO; WAGNER, 1999). No CAMSET, o número de estados considerado na EM depende da temperatura e da tensão  $V_{DS}$ . Sendo assim, ao comparar seus resultados com o deste simulador, Uchida buscou demonstrar que um modelo mais simples, que considera apenas dois estados, é capaz de reproduzir o comportamento do transistor. Os resultados comprovam que, para uma temperatura máxima de 20 K e  $V_{DS} = 20$  mV, o modelo de Uchida possui precisão muito próxima à do CAMSET.

Wang e Porod propuseram um modelo analítico para o SET que considera cargas de *background*<sup>1</sup> e pode ser aplicado a dispositivos assimétricos (WANG; POROD, 2001).

Os autores levantaram a curva  $I_{DS}-V_{GS}$  de seu modelo e compararam com simulações feitas com o SIMON. Os resultados obtidos mostram que o modelo de Wang e Porod subestima o valor da corrente. Deste modo, os autores propõem que seja aplicado um processo de interpolação a fim de compensar o valor da corrente. Além de este ser um método pouco representativo do comportamento do transistor, os autores não fornecem maiores informações sobre como esta interpolação deveria ser executada.

Um dos modelos mais bem sucedidos para o SET foi proposto em 2002 e batizado de MIB, devido às iniciais de seus desenvolvedores (MAHAPATRA; IONESCU; BANERJEE, 2002). O modelo MIB considera fluxo unidirecional de corrente, ou seja, na Figura 28, as taxas  $\Gamma_S^-$  e  $\Gamma_D^-$  são nulas.

O modelo MIB efetua uma comparação entre as tensões de polarização e da ilha do transistor. A partir desta comparação, determina-se a ocorrência ou não de tunelamento e calculam-se as correntes. A corrente total entre dreno e fonte é composta pela soma de dois termos: a corrente de tunelamento e a corrente devido à excitação térmica dos elétrons.

Deste modo, os autores não aplicam a resolução da EM. Isto faz com que o MIB seja um modelo mais simples que o de Uchida, em termos de tempo computacional demandado para a simulação. A validação da funcionalidade foi feita através do levantamento da curva  $I_{DS}-V_{GS}$  e da característica transiente de um inversor mono-elétron implementado através do MIB, do modelo de Uchida e de simulação com o SIMON. O MIB apresenta maior proximidade com as curvas resultantes do SIMON. As simulações foram realizadas a uma temperatura de 5 K. É válido ressaltar que o cálculo das taxas de tunelamento no modelo MIB é feito pela Equação (4.1), considerando  $T = 0$  K. Desta forma os autores desacoplam as correntes de tunelamento e de excitação térmica. A Figura 29 ilustra o circuito de um inversor mono-elétron.

<sup>1</sup>As cargas de *background* podem ser induzidas por impurezas decorrentes do processo de fabricação. Como o SET é bastante sensível, estas cargas podem ocasionar erros na operação do transistor, levando-o a conduzir quando deveria estar bloqueado e vice-versa. Este é um dos maiores obstáculos à aplicação prática do SET em circuitos integrados (LIENTSCHNIG; WEYMANN; HADLEY, 2003; ABUTALEB, 2015).

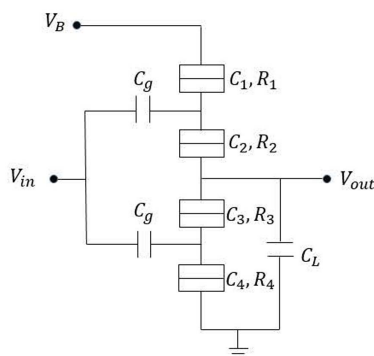


Figura 29: Diagrama de um inversor mono-elétron com dois SETs em série.  $V_{in}$  é a tensão de entrada,  $V_{out}$  é a tensão de saída,  $V_B$  é a tensão de polarização,  $C_n, R_n$ , para  $n = 1, 2, 3, 4$ , são os parâmetros das junções-túnel e  $C_L$  é a capacitância de interconexão - modificada de (YU; HWANG; AHN, 1999).

Ainda em 2002, Mahapatra melhorou a primeira versão do MIB, considerando a dependência da temperatura nas taxas de tunelamento, conforme a teoria ortodoxa (MAHAPATRA et al., 2002). Esta nova versão também foi validada em comparação com as curvas obtidas pelo SIMON. Os resultados comprovam que esta abordagem fornece aproximações mais precisas que a abordagem anterior. Mahapatra *et al.* ressaltam que esta versão do MIB pode modelar dispositivos assimétricos, com um ou dois *gates* e considera os efeitos de cargas de *background*.

O próximo modelo de destaque foi apresentado por Lientschnig (LIENTSCHNIG; WEYMANN; HADLEY, 2003). Baseado na teoria ortodoxa, o modelo calcula a corrente média através do SET em função de  $V_{DS}$ ,  $V_{GS}$  e da temperatura,  $T$ . O grande diferencial desta proposta é o fato de ela não resolver a EM para determinar a probabilidade de ocupação dos estados. Lientschnig *et al.* fazem uso da relação recursiva dada por (4.11). Este procedimento faz com que menos esforço computacional seja demandado na simulação, tornando-a mais eficiente. Sendo  $\sum_n P(n) = 1$ , as probabilidades de ocupação dos estados podem ser calculadas.

$$P(n) = P(n-1) \left( \frac{\Gamma_S^+(n-1) + \Gamma_D^-(n-1)}{\Gamma_S^-(n) + \Gamma_D^+(n)} \right). \quad (4.11)$$

As taxas de tunelamento são calculadas por (4.1) e a corrente é dada por (4.8). Estas equações foram implementadas em SPICE, a fim de que a simulação de circuitos híbridos fosse efetuada.

Os autores compararam as curvas  $I_{DS} - V_{GS}$  e  $I_{DS} - V_{DS}$  de seu modelo com as obtidas através do SIMON. Os resultados apresentam boa concordância com o SIMON, classificando o modelo de Lientschnig como preciso e eficiente.

Inokawa *et al.* estenderam o modelo de Uchida *et al.* a dispositivos assimétricos e com dois *gates* (INOKAWA; TAKAHASHI, 2003). Os autores, no entanto, consideram apenas dois estados na resolução da EM. Ou seja, embora a aplicabilidade do modelo tenha sido ampliada, não houve ganho em precisão ou no intervalo de  $V_{DS}$  no qual o modelo é válido.

Buscando aumentar a aplicabilidade do modelo MIB, Mahapatra *et al.* propuseram mais algumas alterações no modelo original (MAHAPATRA *et al.*, 2004). Nestas alterações, o objetivo era aumentar o intervalo de  $V_{DS}$  no qual o modelo mantém a validade, tornando o MIB mais adequado para a simulação de circuitos híbridos. O novo MIB usa a relação recursiva, expressa em (4.11), para calcular as probabilidades de ocupação dos estados. O modelo novo é aplicável para valores de  $V_{DS} \leq 3e/C_{ilha}$ . Os autores aproveitaram a periodicidade da corrente  $I_{DS}$ , ilustrada na Figura 10, e efetuaram os cálculos em apenas um período, minimizando o esforço computacional. Apesar de buscarem aumentar a precisão de seu modelo, Mahapatra *et al.* demonstraram a funcionalidade de sua proposta em temperaturas de até 15 K, não se aproximando da temperatura ambiente, como fizeram Lientschnig *et al.*

Pruvost *et al.* desenvolveram um modelo capaz de adaptar o número de estados necessários conforme a temperatura e o valor de  $V_{DS}$  (PRUVOST; MIZUTA; ODA, 2008). A proposta envolve a definição de uma região de cálculo. Esta região é delimitada no plano  $V_{DS}-V_{GS}$ , também conhecido como diagrama de estabilidade do SET. Considerando a periodicidade de  $I_{DS}$  em função de  $V_{GS}$ , esta região tem largura igual a  $e/C_g$ .

Os autores levantaram as curvas características do SET e compararam seus resultados com os obtidos por Uchida (UCHIDA *et al.*, 2000) e por Mahapatra (MAHAPATRA *et al.*, 2004) para temperaturas variando de 10 a 300 K. A 300 K, o modelo de Pruvost é o único a apresentar bloqueio de Coulomb.

As propostas mais recentes de modelos analíticos buscam aumentar a representatividade dos modelos utilizando o menor número possível de estados. Ou seja, busca-se a melhor alternativa para o *trade-off* entre precisão e esforço computacional. O modelo desenvolvido por Hasaneen *et al.* aplica os métodos de cálculo propostos por Lientschnig. A fim de obter um intervalo de validade em que  $|V_{DS}| \leq 4e/C_{ilha}$ , os

autores utilizaram 10 estados e obtiveram resultados bastante precisos quando comparados com os fornecidos pelo SIMON, a  $T = 77$  K.

Radwan *et al.* combinaram a metodologia proposta por Lientschnig para o cálculo das probabilidades de estado com a de Pruvost *et al.* para encontrar o número de estados que melhor representa cada temperatura e polarização (RADWAN *et al.*, 2015).

Os autores compararam as curvas características obtidas com o seu modelo e com o SIMON. De fato, Radwan *et al.* utilizaram como métrica o erro quadrático médio para mostrar que os modelos de Hasaneen e o MIB não conseguem fornecer a mesma precisão que o seu, quando comparados com o SIMON. O modelo de Hasaneen utiliza 10 estados para conseguir um erro de 0,02 %. Radwan obtém este valor com apenas quatro estados.

A próxima seção dedica-se a apresentar a revisão relativa aos modelos macro para o SET. Esta técnica de modelagem é, como será visto, bastante frequente na literatura.

#### 4.2.2 Modelos Macro

As técnicas de modelagem e simulação de circuitos mono-elétron apresentadas anteriormente, EM e MC, são computacionalmente custosas. Além disso, as ferramentas de simulação que empregam tais técnicas, em geral, não permitem a análise de circuitos híbridos SET-MOS. Desta forma, a possibilidade de se desenvolver um modelo compacto para o SET começou a ser considerada pelos pesquisadores no final dos anos 90. Os principais requisitos consistiam em simulação eficiente, no que concerne ao tempo computacional demandado, e na utilização de uma plataforma de simulação abrangente, como o SPICE.

Em 1999, Yu *et al.* (YU; HWANG; AHN, 1999) desenvolveram a base teórica para a modelagem macro do SET. Os autores também propuseram o primeiro modelo macro de um SET, que viria a ser aprimorado por diversos grupos de pesquisa. Segundo Yu *et al.*, simuladores compactos como o SPICE descrevem o comportamento de um circuito a partir de sua topologia. Considerando transistores, portanto, há duas suposições básicas para este tipo de simulação:

- Uma vez que um transistor isolado tenha sido parametrizado, seus parâmetros são válidos em todo o circuito;
- A característica I-V de um dispositivo é afetada apenas pelas alterações das tensões terminais de transistores vizinhos. A interação entre dispositivos adjacentes, em geral, é desprezada.

Para transistores mono-elétron, a princípio, a segunda suposição não é válida. A corrente terminal de um SET é determinada a partir de uma média dos estados de carga da ilha. Quando vários SETs são conectados, o estado de carga da ilha de um SET é afetado pelos estados de carga dos SETs vizinhos. Deste modo, um SET isolado pode apresentar uma corrente terminal bastante diferente quando colocado em um circuito (YU; HWANG; AHN, 1999).

A solução proposta por Yu *et al.* consiste em fazer a capacitância da interconexão ser grande o suficiente para que os transistores sejam isolados uns dos outros. Para encontrar o valor limite a partir do qual os dispositivos podem ser tratados de forma isolada, os autores utilizaram o inversor mono-elétron, ilustrado na Figura 29. Os valores utilizados no circuito do inversor estão descritos na Tabela 1.

Tabela 1: Valores dos parâmetros do circuito inversor.

Parâmetro	Valor
$C_n, n = 1, 2, 3, 4$	1,6 aF
$R_n, n = 1, 2, 3, 4$	100 M $\Omega$
$V_B$	30 mV
$C_g$	3,2 aF
$C_L$	0,5 $\mu$ F

Considerando a Figura 29, Yu *et al.* analisaram a corrente terminal do SET formado pelas junções 3 e 4, uma vez que ele é influenciado pelo primeiro transistor (formado pelas junções 1 e 2). Através da variação de  $C_L$ , os autores concluíram que, se a condição  $C_L > 6,25C_j$ , em que  $C_L$  é a capacitância da interconexão e  $C_j$  é a capacitância da junção-túnel, for satisfeita, os SETs podem ser tratados de forma isolada e, portanto, a simulação em SPICE é válida. Para chegar a tal conclusão, a resposta do circuito inversor conforme  $C_L$  era variado, em SPICE, foi comparada com a resposta obtida através do KOSEC, um simulador mono-elétron baseado no método de MC.

Sendo assim, os autores propuseram o primeiro modelo macro para um transistor mono-elétron. Conforme ilustrado na Figura 30, em um modelo macro, a característica I-V do SET é obtida através de um circuito equivalente. Este circuito é composto por dispositivos convencionais, em uma topologia capaz de emular as características terminais do SET.

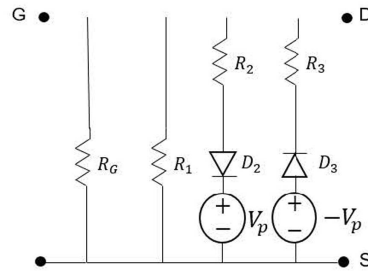


Figura 30: Circuito equivalente para o SET proposto por Yu *et al.* Os terminais marcados com G, S e D correspondem aos terminais de *gate*, fonte e dreno, respectivamente - modificada de (YU; HWANG; AHN, 1999).

Na Figura 30,  $R_G$  deve ser suficientemente grande para inibir o fluxo de corrente entre os terminais G e S. No modelo de Yu *et al.*,  $R_G = 100 \text{ M}\Omega$ . A característica simétrica da curva  $I_{DS}-V_{DS}$  é incorporada ao modelo através dos componentes  $R_2, D_2, V_p$  e  $R_3, D_3, -V_p$ . Os diodos e as fontes de tensão permitem fluxo de corrente em ambos os sentidos, dependendo do valor de  $V_{DS}$ . As oscilações de Coulomb, presentes na curva  $I_{DS}-V_{GS}$ , são implementadas através dos resistores  $R_1, R_2$  e  $R_3$ . Estes componentes são expressos pelas seguintes equações:

$$R_1(V_G) = CR1 + CR2 \cos(CF1.V_G), \quad (4.12)$$

$$R_2(V_G) = R_3(V_G) = \frac{CV_p}{CI2 - \frac{2CV_p}{R_1(V_G)}}, \quad (4.13)$$

em que,  $CR1$ ,  $CR2$ ,  $CF1$ ,  $CV_p$  e  $CI2$  são parâmetros usados para ajustar a curva I-V a diversos valores de  $V_G$ . A Tabela 2 detalha os valores utilizados no trabalho de Yu *et al.*

É válido mencionar que o parâmetro  $CF1$  pode ser calculado por (4.14) e que  $CV_p$  corresponde ao valor absoluto das fontes  $V_p$ , ou seja,  $CV_p$  é o valor limite a partir do qual há condução de corrente em um dos ramos entre os terminais D e S.

Tabela 2: Valores dos parâmetros de ajuste de curva utilizados por Yu *et al.*

Parâmetro	Valor
CR1	300 M $\Omega$
CR2	100 M $\Omega$
CF1	40 V <sup>-1</sup>
$CV_p$	0,02 V
CI2	0,2 nA

$$CF1 = \frac{2C_G}{e}, \quad (4.14)$$

em que,  $C_G$  é a capacitância de *gate* e  $e$  é a carga elementar.

A fim de validar a funcionalidade de seu modelo, Yu *et al.* utilizaram dois circuitos: o inversor mono-elétron e uma porta não-ou composta por oito SETs. Os autores simularam estes circuitos em SPICE, através do modelo proposto, e em KOSEC. A temperatura utilizada nas simulações foi de 30 K. Os resultados das simulações foram comparados, o que comprovou que os resultados obtidos pelo modelo macro possuem precisão razoável. Quanto ao tempo computacional, a simulação em SPICE foi consideravelmente mais rápida. O tempo total de simulação em SPICE foi quatro ordens de grandeza menor que o tempo total demandado pelo KOSEC.

Uma vez que os resultados obtidos por Yu *et al.* foram positivos, alguns pesquisadores buscaram formas de aumentar a precisão dos modelos macro. Em 2003, Wu e Lin (WU; LIN, 2003) propuseram modificações no modelo de Yu *et al.* A fim de reduzir a corrente de fuga entre os terminais G e S da Figura 30, Wu e Lin propuseram o modelo ilustrado na Figura 31.

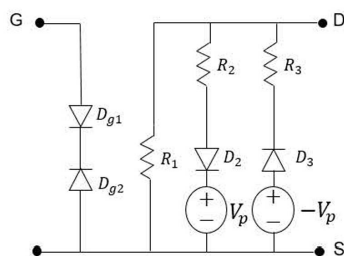


Figura 31: Circuito equivalente para o SET proposto por Wu e Lin. Os terminais marcados com G, S e D correspondem aos terminais de *gate*, fonte e dreno, respectivamente - modificada de (WU; LIN, 2003).

Na Figura 31 nota-se que a modificação proposta por Wu e Lin consiste em substituir o resistor  $R_G$  por dois diodos em conexão série com polaridades opostas,  $D_{g1}$  e  $D_{g2}$ . As equações que expressam os parâmetros também foram alteradas:

$$R1(V_G, V_{ds}) = CR1 + CR2[\cos(CF1\pi V_G) + 1]2^{CV_p - V_{ds}^x}, \quad (4.15)$$

$$R2(V_G, V_{ds}) = R3(V_G, V_{ds}) = \frac{CV_p}{CI2 - \frac{2CV_p}{R1(V_G, V_{ds})}}, \quad (4.16)$$

o parâmetro  $x$  é uma função da temperatura. Os parâmetros CR1 e CR2 são dados pelas equações (4.17) e (4.18), respectivamente. CF1 é dado por (4.14),  $CV_p$  e CI2 possuem os mesmos valores usados por Yu, como consta na Tabela 2.

$$CR1 = 4R_T, \quad (4.17)$$

$$CR2 = 1,33R_T, \quad (4.18)$$

em que  $R_T$  é a resistência de tunelamento da junção.

Para validar a funcionalidade de seu modelo, Wu e Lin levantaram as curvas  $I_{DS}-V_{DS}$  e  $I_{DS}-V_{GS}$  e utilizaram como *benchmark* os resultados obtidos pelo SIMON. A comparação mostrou que o modelo de Wu e Lin apresenta boa concordância com as curvas do SIMON. Os autores também simularam um inversor mono-elétron e compararam seu resultado com o obtido através do modelo de Yu *et al.* Apesar de o resultado de Wu e Lin estar mais próximo do *benchmark* que o de Yu *et al.*, ele ainda encontra-se longe de apresentar a precisão do SIMON.

Outra desvantagem apresentada pelo modelo macro de Yu, proposto em 1999, é o fato de ele ser aplicável apenas a análises estacionárias. Em 2005, o próprio Yu desenvolveu um novo modelo macro, que poderia ser utilizado para análises transientes em SPICE (YU; HWANG; AHN, 2005). Para análises transientes, os autores chegaram à conclusão de que quando a condição  $C_L > 64.C_j$  é satisfeita, o SET pode ser tratado de forma isolada. O circuito equivalente é retratado na Figura 32.

Os parâmetros do circuito equivalente são encontrados por meio das equações (4.19) e (4.20).

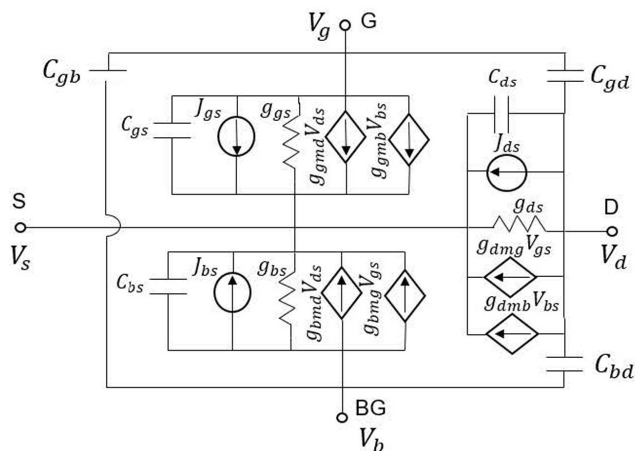


Figura 32: Circuito equivalente proposto por Yu *et al.* Os terminais marcados com G, BG, S e D correspondem aos terminais de *gate*, *backgate*, fonte e dreno, respectivamente - modificada de (YU; HWANG; AHN, 2005).

$$\begin{bmatrix} J_{ds} \\ J_{gs} \\ J_{bs} \end{bmatrix} = \begin{bmatrix} g_{ds} & g_{dmg} & g_{dmb} \\ g_{gmd} & g_{gs} & g_{gmb} \\ g_{bmd} & g_{bmg} & g_{bs} \end{bmatrix} \begin{bmatrix} V_{ds} \\ V_{gs} \\ V_{bs} \end{bmatrix} - \begin{bmatrix} I_d \\ I_g \\ I_b \end{bmatrix}, \quad (4.19)$$

$$\begin{bmatrix} g_{ds} & g_{dmg} & g_{dmb} \\ g_{gmd} & g_{gs} & g_{gmb} \\ g_{bmd} & g_{bmg} & g_{bs} \end{bmatrix} = \begin{bmatrix} \frac{\partial I_d}{\partial V_{ds}} & \frac{\partial I_d}{\partial V_{gs}} & \frac{\partial I_d}{\partial V_{bs}} \\ \frac{\partial I_g}{\partial V_{gs}} & \frac{\partial I_g}{\partial V_{gs}} & \frac{\partial I_g}{\partial V_{bs}} \\ \frac{\partial I_b}{\partial V_{gs}} & \frac{\partial I_b}{\partial V_{gs}} & \frac{\partial I_b}{\partial V_{bs}} \end{bmatrix}, \quad (4.20)$$

em que  $V_{ds}$  é a tensão entre os terminais de dreno e fonte,  $V_{gs}$  é a tensão entre os terminais de *gate* e fonte,  $V_{bs}$  é a tensão entre os terminais de *backgate* e fonte,  $J_{ds}$  é a fonte de corrente equivalente entre dreno e fonte,  $J_{gs}$  é a fonte de corrente equivalente entre *gate* e fonte,  $J_{bs}$  é a fonte de corrente equivalente entre *backgate* e fonte,  $I_d$  é a corrente fluindo para o dreno,  $I_g$  é a corrente fluindo para o *gate* e  $I_b$  é a corrente fluindo para o *backgate*.

A validação da funcionalidade do modelo foi feita através da implementação de um inversor mono-elétron, uma porta não-e e um flip flop JK. Os resultados obtidos (curvas I-V e tempo de simulação) são comparados com as aproximações de MC e da EM. O modelo macro é aproximadamente duas vezes mais rápido que os demais métodos. Quanto à precisão, a modelagem macro, conforme esperado, é a menos precisa.

Karimian *et al.* propuseram alterações no modelo de Wu *et al.* (KARIMIAN et al., 2009). Com este novo modelo os autores buscaram aumentar a precisão e possibilitar o cálculo do tempo de atraso entre entrada e saída. Para tanto, foi utilizado um circuito com capacitor chaveado.

Conforme ilustrado na Figura 33, o capacitor  $C$  e as chaves  $\phi_1$  e  $\phi_2$  formam o que os autores nomearam de bloco quantizador. A carga armazenada no capacitor executa a quantização. As chaves são comandadas por dois sinais de *clock*, retratados na Figura 34.

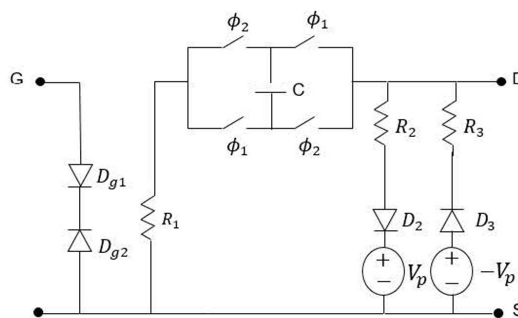


Figura 33: Circuito equivalente proposto por Karimian *et al.* Os terminais marcados com G, S e D correspondem aos terminais de *gate*, fonte e dreno, respectivamente - modificada de (KARIMIAN et al., 2009).

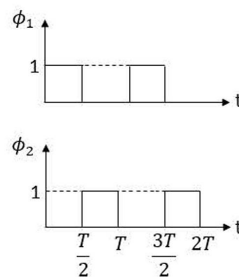


Figura 34: Sinais de *clock* utilizados no modelo de Karimian *et al.* Os sinais não possuem sobreposição, cada sinal controla duas chaves de forma simultânea.  $T$  é o período dos sinais - modificada de (KARIMIAN et al., 2009).

Caso os sinais de entrada e saída não possuam variações durante o período  $T$ , o bloco quantizador pode ser aproximado por uma resistência equivalente no valor de  $T/4C$  (KARIMIAN et al., 2009). Os componentes do circuito da Figura 33 são calculados através das equações (4.22) e (4.23). Os parâmetros  $CR_2$ ,  $CF_1$ ,  $CV_p$  e  $CI_2$  possuem os mesmos valores utilizados por Wu *et al.*, discutidos anteriormente.

O parâmetro  $CR1$  é relacionado à capacitância  $C$  do capacitor chaveado através de (4.21).

$$CR1 = \frac{T}{4C}, \quad (4.21)$$

em que  $T$  é o período dos sinais  $\phi_1$  e  $\phi_2$ .

$$R_1(V_g, V_{ds}) = CR2[\cos(CF1\pi V_g) + Y]2^{CV_p - V_{ds}x}, \quad (4.22)$$

$$R_2(V_g, V_{ds}) = R_3(V_g, V_{ds}) = \frac{CV_p}{CI2 - \frac{2CV_p}{(R_1(V_g, V_{ds}) + CR1)}}, \quad (4.23)$$

em que  $Y$  é definido pelos autores como um parâmetro positivo e  $x$  é uma função da temperatura (da mesma forma que para o modelo de Wu *et al*).

Para fins de validação, Karimian utilizou como *benchmark* curvas  $I_{DS}-V_{DS}$  obtidas por meio do SIMON. As curvas obtidas com o modelo proposto e com o modelo de Wu foram comparadas com o *benchmark*, para uma temperatura de 30 K. Os resultados comprovam que o modelo da Karimian aproxima-se mais dos resultados do SIMON que o modelo de Wu.

Uma das mais recentes propostas de modelo macro para o SET foi realizada por Jain *et al.* (JAIN *et al.*, 2015). Baseados no modelo original de Yu (YU; HWANG; AHN, 1999), os autores propuseram alterações no circuito equivalente e nas equações que definem os parâmetros. A justificativa para tais alterações consiste em melhorar a característica I-V na região de bloqueio de Coulomb.

Deste modo, a nova proposta substitui o resistor  $R_1$ , ver Figura 30, por uma fonte de corrente controlada por tensão. O circuito equivalente é ilustrado na Figura 35.

Os valores dos componentes do circuito equivalente são dados pelas equações (4.24)-(4.26).

$$R_1(V_{gs}) = \frac{\frac{CV_p}{CI2 - \frac{2CV_p}{CR1 - CR2 \cos(CF1\pi V_{gs})}}}{K_1}, \quad (4.24)$$

$$R_2(V_{gs}) = \frac{\frac{CV_p}{CI2 - \frac{2CV_p}{CR1 + CR2 \cos(CF1\pi V_{gs})}}}{K_2}, \quad (4.25)$$

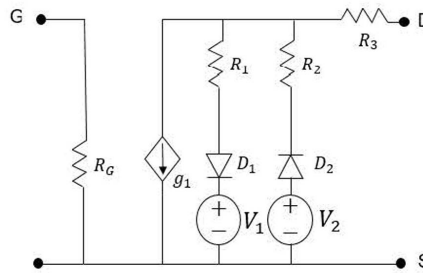


Figura 35: Circuito equivalente proposto por Jain *et al.* Os terminais marcados com G, S e D correspondem aos terminais de *gate*, fonte e dreno, respectivamente.  $g_1$  é a fonte de corrente controlada por tensão e  $R_3$  foi incluído no modelo para facilitar a medição da corrente no dreno - modificada de (JAIN *et al.*, 2015).

$$g_1(V_{gs}) = \frac{CR3(\pi \cdot V_{gs})}{K_3}, \quad (4.26)$$

em que  $K_1$ ,  $K_2$  e  $K_3$  são fatores de escala,  $CR2 = 220 \text{ M}\Omega$ ,  $CR3 = 0,5 \text{ n}\Omega^{-1}$ , e  $R_G = 100 \text{ G}\Omega$ . Os parâmetros  $CF1$ ,  $CI2$ ,  $CR1$  e  $CV_p$  possuem os mesmos valores usados por Yu, expressos na Tabela 2.

Os fatores  $K_1$  e  $K_2$  permitem a variação do intervalo de cálculo da corrente de dreno, sem alterar a característica da curva. Outra vantagem destes fatores é a possibilidade de optar entre um transistor simétrico ( $K_1 = K_2$ ) ou assimétrico ( $K_1 \neq K_2$ ). O fator  $K_3$  controla o deslocamento vertical da curva  $I_{DS}-V_{DS}$  para diferentes valores de  $V_{GS}$ . Para justificar o aumento no valor de  $R_G$  (de  $100 \text{ M}\Omega$  utilizado por Yu para  $100 \text{ G}\Omega$ ), os autores simularam a característica estática de um inversor monoeletrôn para diversos valores de  $R_G$ . Os resultados demonstram que para  $50 \text{ G}\Omega$ , a característica se aproxima do ideal.

A curva  $I_{DS}-V_{DS}$  do modelo de Jain foi comparada com as curvas dos modelos de Yu, de Wu e Lin, do modelo MIB e do SIMON, usada como *benchmark*. Durante o bloqueio de Coulomb, as curvas do modelo MIB e do modelo de Jain são as mais próximas do *benchmark*. Um bom diferencial do trabalho de Jain *et al.* consiste na simulação de um circuito integrador construído a partir do modelo proposto. Deste modo, os autores comprovam que seu modelo pode ser utilizado em circuitos analógicos e em análises transientes.

Os modelos macro apresentam, em geral, precisão inferior à apresentada por modelos probabilísticos do SET. Contudo, o considerável ganho em eficiência computacional faz com que os pesquisadores busquem alternativas que forneçam maior

precisão às simulações. Conforme detalhado nesta seção, há diversos trabalhos neste sentido. Pode-se afirmar que a proposta de Yu *et al.* tem guiado estes esforços desde 1999.

### 4.3 CONSIDERAÇÕES FINAIS

Neste capítulo, buscou-se cobrir os principais aspectos relativos à modelagem caixa branca do SET. Foram explanados os métodos de modelagem e de simulação da Equação Mestra e de Monte Carlo. Um amplo panorama englobando diversos modelos já propostos para o transistor mono-elétron foi traçado. As técnicas de modelagem analítica e macro foram explicitadas e os principais trabalhos relativos à modelagem do SET foram revistos.

## 5 MÉTODOS APLICADOS

No Capítulo 3, foram discutidas as etapas que constituem a modelagem caixa preta: coleta de dados, determinação de estrutura, estimação de parâmetros e validação. O objetivo deste capítulo é apresentar os métodos que foram utilizados neste trabalho para a determinação de estrutura e estimação de parâmetros.

A etapa de estimação de parâmetros, em um problema de identificação, é fundamentalmente um problema de otimização (AGUIRRE, 2015). Como será explicitado neste capítulo, os métodos aplicados nesta etapa buscam a minimização de uma função de custo. Deste modo, a primeira parte deste capítulo apresenta alguns conceitos de otimização determinística.

Na segunda parte deste capítulo, buscou-se detalhar os métodos que auxiliaram na construção dos melhores modelos obtidos para o SET neste trabalho. É, no entanto, válido ressaltar que existem diversas técnicas que ajudam a determinar a estrutura de um modelo e a estimar seus parâmetros. A escolha de qual técnica aplicar pode se basear na literatura disponível, no conhecimento que o projetista tem sobre o sistema, na aplicação desejada para o modelo, dentre outros fatores (ALVES, 2009).

### 5.1 CONCEITOS INICIAIS

Nesta seção será apresentada a base conceitual necessária para a compreensão do método de Levenberg-Marquardt, empregado na estimação de parâmetros dos modelos obtidos neste trabalho.

Dados o ponto  $a \in \mathbb{R}^n$  e o número  $r \in \mathbb{R}_+$ , define-se a bola aberta de centro em  $a$  e raio  $r$  como o conjunto  $B(a; r)$  dos pontos  $x \in \mathbb{R}^n$  cuja distância ao ponto  $a$  é menor do que  $r$  (LIMA, 2013).

$$B(a; r) = \{x \in \mathbb{R}^n; \|x - a\| < r\}, \quad (5.1)$$

considere que  $a \in X \subset \mathbb{R}^n$ , pode-se dizer que o ponto  $a$  é interior ao conjunto  $X$  quando, para algum  $r > 0$ , tem-se  $B(a; r) \subset X$ . Ou seja, todos os pontos suficientemente próximos do ponto  $a$  também pertencem ao conjunto  $X$  (LIMA, 2013).

Deste modo, pode-se definir os pontos de mínimo global,  $x_g^*$ , e de mínimo local,  $x_l^*$ , para uma função  $f(\cdot)$  em  $\mathbb{R}^n$ . Seja  $D \subset \mathbb{R}^n$  e uma função  $f : D \rightarrow \mathbb{R}$ , o ponto  $x_g^* = [x_{1g} \ x_{2g} \ \dots \ x_{ng}]$  é ponto de mínimo global, se  $f(x_g^*) < f(x) \ \forall x \in D$ . Se, entretanto, existir apenas uma bola aberta,  $B$ , de centro em  $a$ , com  $r > 0$ , tal que  $f(x_l^*) < f(x) \ \forall x \in B$ , então o ponto  $x_l^* = [x_{1l} \ x_{2l} \ \dots \ x_{nl}]$  é ponto de mínimo local (LIMA, 2013).

O teorema de máximos e mínimos auxilia na busca por estes pontos. Considere a definição do gradiente de uma função  $f(x)$  em  $\mathbb{R}^n$ :

$$\nabla f = \frac{\partial f(x)}{\partial x}. \quad (5.2)$$

**Teorema 1 (Máximos e Mínimos).** *Seja  $x$  um ponto interior de  $D$ , suponha que  $\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_1}$ ,  $\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_2}$ , ...,  $\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_n}$  existem. Se  $x$  é ponto de mínimo ou máximo de  $f$ , então  $\nabla f$  é o vetor nulo.*

Uma vez que as funções de interesse neste trabalho têm domínio com natureza vetorial,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , define-se a Jacobiana de  $f$  como:

$$J_f(x) = \nabla f = \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \quad \dots \quad \frac{\partial f(x)}{\partial x_n} \right]^T, \quad (5.3)$$

em que  $x_1, x_2, \dots, x_n$  são componentes do vetor  $x$ .

A matriz Hessiana da função  $f$  é dada por (5.4).

$$H_f(x) = \nabla^2 f = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n x_1} & \frac{\partial^2 f(x)}{\partial x_n x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}, \quad (5.4)$$

Os métodos de minimização de uma função de custo,  $f(\cdot)$ , podem ser classificados com base na informação necessária sobre  $f(\cdot)$  para que realizem a busca por  $x_g^*$  (ISERMANN; MÜNCHHOF, 2011):

- Métodos de ordem zero: dependem apenas da avaliação de  $f(\cdot)$ ;
- Métodos de primeira ordem: avaliam a função,  $f(\cdot)$ , e sua Jacobiana,  $\nabla f$ ;
- Métodos de segunda ordem: avaliam a função,  $f(\cdot)$ , sua Jacobiana,  $\nabla f$ , e sua Hessiana,  $\nabla^2 f$ .

Outro conceito muito utilizado na avaliação de funções de custo é o teorema de Taylor. Expandir uma função através do teorema de Taylor tem por objetivo reduzir o custo computacional da implementação (OROSKI, 2015).

**Teorema 2 (Teorema de Taylor).** *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é continuamente diferenciável e  $p \in \mathbb{R}^n$ , então:*

$$f(x + p) = f(x) + \nabla f(x + \alpha p)^\top p, \quad (5.5)$$

em que  $\alpha \in (0, 1)$ . Como  $f$  é continuamente diferenciável, pode-se aplicar o operador  $\nabla$  mais uma vez:

$$\nabla f(x + p) = \nabla f(x) + \int_0^1 \nabla^2 f(x + \alpha p)^\top p d\alpha, \quad (5.6)$$

assim,

$$f(x + p) = f(x) + \nabla f(x)^\top p + \frac{1}{2} p^\top \nabla^2 f(x + \alpha p)^\top p. \quad (5.7)$$

A seguir serão apresentados os métodos de busca em linha e região de confiança.

## 5.2 MECANISMOS DE BUSCA

Os métodos de otimização, em geral, seguem o modelo de convergência iterativa (OROSKI, 2015):

$$x_{k+1} = x_k + \alpha_k \Delta x_k, \quad (5.8)$$

sendo  $\Delta x_k$  o passo que é ponderado pelo escalar  $\alpha_k$ . Uma vez que estes parâmetros tenham sido escolhidos adequadamente, espera-se que:

$$f(k+1) < f(k), \quad (5.9)$$

em que se considerou um problema de minimização. Pode-se expressar este método pelo algoritmo 1.

---

**Algoritmo 1: Métodos de direção descendente.**

---

**Entrada:** Ponto inicial  $x \in \text{dom } f$

1 **início**

2 **Enquanto**  $\nabla f(x_k) \neq 0$  **faça**

1. Determinar a direção de busca,  $\Delta x$ ;

2. *Line Search*: escolher o passo,  $\alpha > 0$ ;

3. Atualizar:  $x_{k+1} = x_k + \alpha \Delta x$ ;

3 **fim**

---

No Algoritmo 1 o termo *Line Search*, ou Busca em Linha, refere-se ao segundo passo seguido por algoritmos de minimização de funções. Em geral, estes algoritmos primeiramente definem uma direção de busca e, posteriormente, passam a buscar em uma reta formada pela direção previamente determinada (BOYD; VANDENBERGHE, 2009). A próxima seção apresenta os principais métodos de busca em linha.

### 5.2.1 Busca em Linha

Uma vez que já se tenha definida uma direção de busca, a busca em linha procura pelo mínimo da função na reta definida por esta direção. Os dois tipos mais comuns de busca em linha são a busca exata em linha e o “*backtracking*”.

### 5.2.1.1 Busca Exata em Linha

A busca exata em linha tem por objetivo encontrar um determinado  $\alpha \in \mathbb{R}$  que minimize a função  $f$  ao longo do raio  $(x + \alpha\Delta x)$ :

$$\alpha = \operatorname{argmin} f(x + \alpha\Delta x), \quad (5.10)$$

em que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\alpha \geq 0$ .

Resolver (5.10) através da busca exata em linha é, em geral, computacionalmente caro e inexato (BOYD; VANDENBERGHE, 2009).

### 5.2.1.2 Backtracking

Este método de busca em linha usa dois parâmetros,  $\alpha$  e  $\beta$ . A ideia geral do *backtracking* está expressa no Algoritmo 2.

---

#### Algoritmo 2: *Backtracking*

---

**Entrada:** Ponto inicial  $x \in \operatorname{dom} f$ ,  $\alpha \in (0, 0,5)$  e  $\beta \in (0, 1)$

$u = 1$ ;

**1 início**

**2** | **Enquanto**  $f(x + \Delta x) \geq f(x) + \alpha u \nabla f(x)^\top \Delta x \neq 0$  **faça**

**3** |  $u = \beta u$ ;

**4 fim**

---

No algoritmo de *backtracking*, o valor do passo,  $u$ , é inicializado com valor unitário. Este passo é reduzido por um fator  $\beta$ , até que o critério de parada seja atingido (BOYD; VANDENBERGHE, 2009).

## 5.2.2 Região de Confiança

Os métodos de busca por região de confiança são alternativas aos métodos de busca em linha. A estratégia da região de confiança é aplicada no algoritmo de Levenberg-Marquardt.

Os métodos de região de confiança definem regiões em volta da iteração atual. Nesta iteração, é gerado um modelo,  $m_k(p)$ , que represente adequadamente

a função objetivo,  $f$ . Então, escolhe-se um passo que minimize o modelo na região de confiança (OROSKI, 2015).

$$\min_{p \in \mathbb{R}^n} m_k(p) = f(x_k) + \nabla f(x_k)^\top p + \frac{1}{2} p^\top M_k p, \quad (5.11)$$

em que  $M_k$  é uma matriz simétrica.

Normalmente, a região de confiança é uma bola definida por  $\|p\| \leq \Delta$ , sendo  $\Delta > 0$  o raio da região de confiança (NOCEDAL; WRIGHT, 1999). Um índice que avalia a adequação do modelo,  $m_k(p)$ , à função objetivo,  $f$ , é expresso por (5.12).

$$\rho = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)}. \quad (5.12)$$

As buscas em linha e por região de confiança diferem na forma como escolhem a direção de busca e o passo em cada iteração (NOCEDAL; WRIGHT, 1999). Enquanto a busca em linha começa por definir uma distância apropriada em uma direção fixa, a busca por região de confiança estabelece uma distância máxima, o raio da região de confiança, para depois estabelecer uma distância e um passo apropriados (OROSKI, 2015).

O Algoritmo 3 descreve o mecanismo de busca por região de confiança.

### 5.3 MÉTODO DE NEWTON

O método de Newton pode ser definido como um algoritmo de otimização de direção descendente e de segunda ordem (BOYD; VANDENBERGHE, 2009). Este método usa a expansão da função objetivo em séries de Taylor, (5.13). O algoritmo 4 descreve o método de Newton.

$$f(x_0 + \Delta x) \approx f(x_0) + \nabla f(x_0) \Delta x + \frac{1}{2} \Delta x^\top \nabla^2 f(x_0) \Delta x. \quad (5.13)$$

O gradiente de (5.13) é dado por:

$$\frac{\partial f(x_0 + \Delta x)}{\partial \Delta x} = \nabla f(x_0) + \nabla^2 f(x_0) \Delta x. \quad (5.14)$$

---

**Algoritmo 3: Busca por Região de Confiança**


---

**Entrada:**  $\Delta > 0$ ,  $\Delta_0 \in (0, \Delta)$  e  $\eta \in (0, \frac{1}{4})$

- 1 **início**
- 2   **Para**  $k = 0, 1, 2, \dots$  **faça**
- 3     Obter  $p_k$  que resolva aproximadamente (5.11);
- 4     Avaliar (5.12);
- 5     **Se**  $\rho < \frac{1}{4}$  **então**
- 6        $\Delta_{k+1} = \frac{1}{4} \|p_k\|$ ;
- 7     **Senão**
- 8       **Se**  $\rho > \frac{3}{4}$  e  $\|p_k\| = \Delta_k$  **então**
- 9          $\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$ ;
- 10      **Senão**
- 11        $\Delta_{k+1} = \Delta_k$ ;
- 12      **Se**  $\rho_k > \eta$  **então**
- 13        $x_{k+1} = x_k + p_k$ ;
- 14      **Senão**
- 15        $x_{k+1} = x_k$ ;
- 16 **fim**

---

Considerando que o ponto  $x_g^*$  minimiza a função  $f(x)$ , tem-se que  $\nabla f(x_g^*) = 0$  e,

$$0 = \nabla f(x_0) + \nabla^2 f(x_0) \Delta x. \quad (5.15)$$

O passo do método de Newton,  $\Delta x$ , é, então:

$$\Delta x = -[\nabla^2 f(\bar{x})]^{-1} \nabla f(\bar{x}). \quad (5.16)$$

Em notação recursiva, tem-se:

$$x_{k+1} = x_k - [\nabla^2 f(\bar{x})]^{-1} \nabla f(\bar{x}). \quad (5.17)$$

O Algoritmo 4 descreve o método de Newton.

---

**Algoritmo 4: Método de Newton**


---

**Entrada:**  $x_0 \in \mathbb{R}^n$ ;

**1 início**
**2** Calcule  $\Delta x_0$ , solução de (5.16);

**3**  $x_1 = x_0 + \Delta x_0$ ;

**4**  $k = 1$ ;

**5 Enquanto**  $\nabla f(x_k) \neq 0$  **faça**
**6** Calcule  $\Delta x_k$ , solução de (5.16);

**7** Determine  $x_{k+1}$ :

**8**  $x_{k+1} = x_k + \Delta x_k$ ;

**9**  $k = k + 1$ ;

**10 fim**


---

Os métodos de segunda ordem são conhecidos pela rápida convergência (OROSKI, 2015). Entretanto, o uso da matriz Hessiana,  $\nabla^2 f(x)$  e a necessidade de sua inversão são duas desvantagens destes métodos. Nem sempre a matriz Hessiana será conhecida e, mesmo em casos em que sua obtenção seja possível, a inversão desta matriz pode tornar o problema numérico mal condicionado (ISERMANN; MÜNCHHOF, 2011). No método de Newton, conforme explicitado pelo Algoritmo 4, é necessário calcular a matriz Hessiana a cada iteração. Isto faz com que o custo computacional deste método seja elevado (BOYD; VANDENBERGHE, 2009).

## 5.4 MÉTODO DE GAUSS-NEWTON

Da mesma forma que o método de Newton, o método de Gauss-Newton é um método de segunda ordem. A diferença consiste no fato de que o método de Gauss-Newton busca simplificar o processo de inversão da matriz Hessiana, a fim de reduzir complicações decorrentes de mal condicionamento numérico (OROSKI, 2015).

Considere a seguinte função de custo:

$$f(x) = \sum_{i=1}^N \xi_i(x)^2, \quad (5.18)$$

a aplicação do operador  $\nabla$  nesta função, resulta em:

$$\nabla f(x) = J_\xi(x)^\top \xi(x), \quad (5.19)$$

em que  $J_\xi(x) = \left[ \frac{\partial \xi_i}{\partial x_j} \right]$  para  $J_\xi(x) \in \mathbb{R}^{m \times n}$ ,  $i = 1, 2, \dots, n$  e  $j = 1, 2, \dots, m$ .

Aplicando novamente operador  $\nabla$ , agora em (5.19), tem-se:

$$\nabla^2 f(x) = \sum_{i=1}^n \nabla \xi_i(x) \nabla \xi_i(x)^\top + \sum_{i=1}^n \xi_i(x) \nabla^2 \xi_i(x). \quad (5.20)$$

Geralmente, a primeira parcela da soma em (5.20) é mais importante do que a segunda parcela (NOCEDAL; WRIGHT, 1999). Deste modo, considera-se a seguinte simplificação:

$$\nabla^2 f(x) \approx J_\xi(x)^\top J_\xi(x). \quad (5.21)$$

Apesar da simplificação expressa em (5.21), o método de Gauss-Newton fornece desempenho muito similar ao método de Newton (NOCEDAL; WRIGHT, 1999). O Algoritmo 5 descreve os passos do método de Gauss-Newton.

---

**Algoritmo 5: Método de Gauss-Newton**

---

**Entrada:**  $x_0 \in \mathbb{R}^n$ ;

1 **início**

2     Calcule  $\Delta x_0$ , solução de:

3      $\nabla^2 f(x_0) \Delta x_0 = -\nabla f(x_0)$ ;

4     em que  $\nabla^2 f(x_0) = J_\xi(x_0)^\top J_\xi(x_0)$ ;

5      $x_1 = x_0 + \Delta x_0$ ;

6      $k = 1$ ;

7     **Enquanto**  $\nabla f(x_k) \neq 0$  **faça**

8         Calcule  $\Delta x_k$ , solução de:

9          $\nabla^2 f(x_k) \Delta x_k = -\nabla f(x_k)$ ;

10        Determine  $x_{k+1}$ :

11         $x_{k+1} = x_k + \Delta x_k$ ;

12         $k = k + 1$ ;

13 **fim**

---

## 5.5 LEVENBERG-MARQUARDT

O algoritmo de Levenberg-Marquardt, LM, é um método de segunda ordem. Desta forma, este método possui a habilidade de convergir rapidamente a partir de um ponto inicial,  $x_0$ , mesmo que este ponto esteja fora da região de convergência de

outros métodos (MARQUARDT, 1963). O algoritmo de LM é um dos mais eficientes e mais utilizados métodos de otimização, especialmente na área de Identificação de Sistemas (ISERMANN; MÜNCHHOF, 2011).

O algoritmo de LM usa a estratégia de busca por região de confiança. Esta estratégia evita o principal problema do método de Gauss-Newton, que ocorre quando a Jacobiana,  $J_\xi(x)$ , passa a ter posto não pleno (NOCEDAL; WRIGHT, 1999). Pode-se, então, entender que o método de Levenberg-Marquardt faz a seguinte modificação no método de Gauss-Newton:

$$\begin{cases} x_{k+1} &= x_k + \Delta x_k \\ \Delta x_k &= -[J_\xi(x)^\top J_\xi(x) + \lambda I]^{-1} [J_\xi(x)^\top J_\xi(x)], \end{cases} \quad (5.22)$$

em que  $\lambda \in \mathbb{R}$  e  $I$  é a matriz identidade.

A ideia do método de LM é perturbar a matriz Hessiana, aproximada por (5.21), considerando  $\nabla^2 f(x) + \lambda I$ , para  $\lambda > 0$ . A expansão em séries de Taylor é aplicada no algoritmo de LM. O Algoritmo 6 descreve o mecanismo de busca do método de LM.

---

**Algoritmo 6: Levenberg-Marquardt**

---

**Entrada:**  $x_0 \in \mathbb{R}^n$ ;

1 **início**

2     **Calcule**  $\Delta_{x_0}$ , **solução de:**

3      $[\nabla^2 f(x_0) + \lambda I] \Delta x_0 = -\nabla f(x_0)$ ;

4     **em que**  $\nabla^2 f(x_0) = J_\xi(x_0)^\top J_\xi(x_0)$ ;

5      $x_1 = x_0 + \Delta_{x_0}$ ;

6      $k = 1$ ;

7     **Enquanto**  $\nabla f(x_k) \neq 0$  **faça**

8         **Calcule**  $\Delta_{x_k}$ , **solução de:**

9          $[\nabla^2 f(x_k) + \lambda I] \Delta x_k = -\nabla f(x_k)$ ;

10         **Determine**  $x_{k-1}$ :

11          $x_{k-1} = x_k + \Delta_{x_k}$ ;

12          $k = k + 1$ ;

13 **fim**

---

As próximas seções discutem métodos frequentemente aplicados na identificação de sistemas.

## 5.6 MÉTODO DOS MÍNIMOS QUADRADOS

O Método dos Mínimos Quadrados, MMQ, é um dos mais conhecidos e aplicados nas diversas áreas de tecnologia (AGUIRRE, 2015). A ideia básica deste método teve origem nos estudos de Gauss sobre Astronomia (DAVIS, 1963). Nestes estudos, Gauss refere-se ao fato de que observações e medições são apenas aproximações da realidade. Desta forma, quaisquer cálculos que se baseiem nestas medições são, por sua vez, aproximações. Um dos aspectos importantes, levantado por Gauss, diz respeito ao número necessário de observações para que a aproximação feita seja satisfatória. Uma vez que devem ser considerados os erros de medição, a quantidade requerida de observações costuma ser maior do que o número estritamente necessário à determinação de grandezas desconhecidas. Essas observações redundantes possibilitam a redução do efeito do erro.

A estimação de parâmetros na modelagem caixa preta pode ser vista como uma solução de um sistema de equações, em que os parâmetros são as incógnitas. Na presença de ruído que corrompa as medições, o número de observações,  $N$ , deve ser maior do que o número de parâmetros a se determinar,  $n$ . Isto faz com que haja mais equações do que incógnitas, ou seja, o sistema de equações será sobredeterminado (OROSKI, 2015).

Considere uma função  $f(\mathbf{x}, \boldsymbol{\theta}) : \mathbb{R}^n \longrightarrow \mathbb{R}$ , em que  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_N]$  e  $\boldsymbol{\theta} \in \mathbb{R}^n$  é o vetor de parâmetros que caracteriza a função  $f(\cdot)$ . Pode-se representar  $f(\cdot)$  da seguinte forma (AGUIRRE, 2015):

$$y = f(\mathbf{x}, \boldsymbol{\theta}). \quad (5.23)$$

A representação em (5.23) define um conjunto de  $N$  equações em que  $N$  observações de  $y$  são escritas em função de  $\mathbf{x}$  e  $\boldsymbol{\theta}$ :

$$\begin{aligned} y_1 &= f(x_1, \boldsymbol{\theta}), \\ y_2 &= f(x_2, \boldsymbol{\theta}), \\ &\vdots \\ y_N &= f(x_N, \boldsymbol{\theta}), \end{aligned} \quad (5.24)$$

sendo  $y_i$  a  $i$ -ésima observação de  $y$  e  $\mathbf{x} = [x_i^1 \ x_i^2 \ \dots \ x_i^n]^T$  são as  $i$ -ésimas observações dos  $n$  elementos do vetor  $\mathbf{x}$ . Cada equação de (5.24) pode ser chamada de restrição. Desta forma, (5.24) explicita  $N$  restrições de (5.23).

Para que se consiga determinar o vetor  $\theta \in \mathbb{R}^n$  a partir das restrições expressas em (5.24), são necessárias, no mínimo,  $n$  equações. As restrições (5.24) podem ser escritas na forma de (5.25) (AGUIRRE, 2015).

$$y = x^T \theta. \quad (5.25)$$

O conjunto de equações em (5.24) pode ser escrito na forma matricial:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^n \\ x_2^1 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \vdots \\ x_n^1 & x_n^2 & \cdots & x_n^n \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix},$$

$$y = X\theta, \quad (5.26)$$

em que  $X \in \mathbb{R}^{n \times n}$  e  $x_i^j$  é a  $i$ -ésima observação da variável  $x^j$ . As variáveis independentes  $x^1, x^2, \dots, x^n$  são chamadas de regressores e  $\theta$  é o vetor de parâmetros a determinar.

No caso em que  $N = n$ , o sistema descrito por (5.26) pode ser resolvido através de (5.27)

$$\theta = X^{-1}y, \quad (5.27)$$

em que  $X$  é uma matriz quadrada, não singular e, portanto, passível de inversão.

No caso em que  $N > n$ , o sistema de equações é sobredeterminado, ou seja,  $X \in \mathbb{R}^{N \times n}$ ,  $y \in \mathbb{R}^{N \times 1}$  e  $\theta \in \mathbb{R}^{n \times 1}$ . Uma vez que a matriz  $X$  agora não é quadrada, a solução do sistema por meio de (5.27) não é possível de forma direta. Entretanto, pode-se escrever a equação normal de (5.26), pré-multiplicando a equação pela matriz  $X^T$ :

$$X^T y = X^T X \theta. \quad (5.28)$$

A multiplicação da transposta de uma matriz pela matriz original resulta em uma matriz quadrada, deste modo, pode-se escrever (AGUIRRE, 2015):

$$\theta = [X^T X]^{-1} X^T y, \quad (5.29)$$

em que  $X^T X$  é uma matriz não singular. A matriz  $[X^T X]^{-1} X^T$  é conhecida como matriz pseudo-inversa (STRANG, 2009).

Cabe ressaltar que, enquanto (5.27) é a única solução de um sistema de equações com  $n$  restrições, (5.29) é apenas uma das soluções do sistema sobre-determinado.

Uma vez que o vetor de parâmetros tenha sido estimado, pode-se definir o resíduo,  $\xi$ , como o erro cometido quando se tenta explicar uma medida real,  $y$ , através de uma medida estimada,  $\hat{y}$ :

$$\begin{aligned}\hat{y} &= X\hat{\theta}, \\ y &= \hat{y} + \xi,\end{aligned}\tag{5.30}$$

em que  $\hat{\theta} \in \mathbb{R}^n$  é o vetor de parâmetros estimado e  $\xi \in \mathbb{R}^N$  é o vetor de resíduos, que pode ser escrito como  $\xi = y - X\hat{\theta}$ .

O MMQ busca minimizar uma função de custo,  $J_{MQ}$ . Como o nome do método sugere, esta função de custo nada mais é do que o somatório do quadrado dos resíduos, dado por (5.31).

$$J_{MQ}(\theta, Z^N) = \sum_{i=1}^N \xi(i)^2 = \xi^T \xi = \|\xi\|^2,\tag{5.31}$$

em que  $Z^N$  é o conjunto de dados utilizado. A função  $J_{MQ}$  pode ser descrita como um índice que mede a qualidade do ajuste de  $X\hat{\theta}$  aos dados reais  $y$  (LJUNG, 1999).

Substituindo-se  $\xi$  em (5.31), obtém-se (os argumentos de  $J_{MQ}$  foram omitidos, por simplicidade):

$$\begin{aligned}J_{MQ} &= (y - X\hat{\theta})^T (y - X\hat{\theta}) \\ J_{MQ} &= y^T y - y^T X\hat{\theta} - \hat{\theta}^T X^T y + \hat{\theta}^T X^T X\hat{\theta}.\end{aligned}\tag{5.32}$$

Uma vez que se busca minimizar  $J_{MQ}$  com respeito a  $\hat{\theta}$ , deve-se resolver  $\frac{\partial J_{MQ}}{\partial \hat{\theta}} = 0$ . Conforme visto na Seção 5.1,  $\frac{\partial J_{MQ}}{\partial \hat{\theta}}$  é chamado de gradiente da função de custo.

Para derivar parcialmente (5.32) é necessário utilizar algumas propriedades relativas à diferenciação de matrizes. Para relembrar as propriedades aqui aplicadas, considere uma matriz  $A \in \mathbb{R}^{n \times n}$  e os vetores  $x$  e  $y \in \mathbb{R}^n$ . Desta forma, as seguintes relações são verdadeiras:

$$\begin{aligned}
\frac{\partial(\mathbf{x}^\top \mathbf{y})}{\partial \mathbf{y}} &= \mathbf{x}; \\
\frac{\partial \mathbf{y}^\top \mathbf{x}}{\partial \mathbf{y}} &= \mathbf{x}; \\
\frac{\partial \mathbf{x}^\top A \mathbf{x}}{\partial \mathbf{x}} &= (A + A^\top) \mathbf{x}.
\end{aligned} \tag{5.33}$$

Usando as propriedades em (5.33), a derivada parcial de (5.32) com relação ao vetor  $\hat{\boldsymbol{\theta}}$  é dada por:

$$\begin{aligned}
\frac{\partial J_{MQ}}{\partial \hat{\boldsymbol{\theta}}} &= -(\mathbf{y}^\top X)^\top - X^\top \mathbf{y} + (X^\top X + X^\top X) \hat{\boldsymbol{\theta}} \\
&= -X^\top \mathbf{y} - X^\top \mathbf{y} + 2X^\top X \hat{\boldsymbol{\theta}}.
\end{aligned} \tag{5.34}$$

Igualando-se (5.34) a zero, chega-se à mesma solução obtida através do método da pseudoinversa, expressa em (5.29) e repetida abaixo.

$$\hat{\boldsymbol{\theta}} = [X^\top X]^{-1} X^\top \mathbf{y}.$$

Por fim, para garantir que  $\hat{\boldsymbol{\theta}}$  seja mínimo, a matriz Hessiana deve satisfazer (5.35).

$$\frac{\partial^2 J_{MQ}}{\partial \hat{\boldsymbol{\theta}}^2} = 2X^\top X > 0. \tag{5.35}$$

A matriz  $X^\top X$  é definida positiva<sup>1</sup> por construção, ou seja, (5.35) é verdadeira.

Neste ponto, a notação utilizada será modificada para que se torne mais próxima àquela encontrada na literatura de Identificação de Sistemas. A matriz de regressão,  $X$ , pode incluir valores relativos à entrada do sistema a ser modelado,  $u$ , e à sua saída,  $y$ . Diz-se, então, que este modelo é do tipo ARX. Neste caso, em que os regressores referem-se tanto à entrada quanto à saída, a matriz de regressão é chamada de  $\Psi$ .

<sup>1</sup>Uma matriz  $A \in \mathbb{R}^{n \times n}$  é definida positiva, se  $\mathbf{z}^\top A \mathbf{z} > 0$  para todo vetor não-nulo  $\mathbf{z} \in \mathbb{R}^n$ . A principal característica de uma matriz definida positiva é o fato de que todos os seus autovalores são positivos.

$$\Psi = \begin{bmatrix} y(n_y) & \cdots & y(1) & u(n_y + 1) & u(n_y) & \cdots & u(n_y - n_u + 1) \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ y(k-1) & \cdots & y(k-n_y) & u(k) & u(k-1) & \cdots & u(k-n_u) \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ y(N-1) & \cdots & y(N-n_y) & u(N) & u(N-1) & \cdots & u(N-n_u) \end{bmatrix}, \quad (5.36)$$

em que  $n_y \in \mathbb{N}^*$  e  $n_u \in \mathbb{N}$  são as regressões máximas de entrada e de saída, respectivamente.  $N$  é o número de observações, ou número de amostras dos sinais. Reescrevendo a Equação (5.29) com a notação descrita acima, tem-se:

$$\hat{\theta} = [\Psi^T \Psi]^{-1} \Psi^T \mathbf{y}. \quad (5.37)$$

A solução dada por (5.37) minimiza a função de custo  $J_{MQ}$ , expressa em (5.31), e sua principal vantagem reside em ser o mínimo global de  $J_{MQ}$ , para modelos que sejam lineares nos parâmetros (LJUNG, 1999).

Existem algumas propriedades do estimador de Mínimos Quadrados, MQ, que fazem dele um dos mais utilizados na área de identificação. Uma das mais vantajosas diz respeito à ausência de polarização do estimador, desde que as hipóteses de Gauss-Markov sejam atendidas (AGUIRRE, 2015).

A definição de polarização de um estimador e a apresentação das hipóteses de Gauss-Markov são melhor compreendidas a partir de alguns conceitos. Primeiramente, deve-se definir a esperança matemática de uma Variável Aleatória<sup>2</sup>, VA,  $E\{\cdot\}$ . Deste modo, considere uma VA,  $\alpha$ , discreta e com uma função massa de probabilidade  $p(\alpha_i)$ . A esperança matemática desta variável é dada por:

$$E\{\alpha\} = \sum_{i=1}^{\infty} \alpha_i p(\alpha_i). \quad (5.38)$$

A covariância mede a interdependência entre duas VAs discretas,  $\alpha$  e  $\beta$ :

$$Cov\{\alpha, \beta\} = E\{(\alpha - E\{\alpha\})(\beta - E\{\beta\})\}. \quad (5.39)$$

<sup>2</sup>Uma VA,  $X$ , pode ser definida como o mapeamento dos resultados de um experimento,  $S$ , em valores reais,  $X : S \rightarrow \mathbb{R}$ . O resultado do lançamento de um dado é um exemplo clássico de VA.

Finalmente, a variância de uma VA mede o quão dispersa uma dada VA é em relação ao seu valor esperado. Deste modo, a variância é um caso particular da covariância, quando  $\alpha = \beta$ :

$$\text{Var}\{\alpha\} = E\{(\alpha - E\{\alpha\})^2\}. \quad (5.40)$$

As hipóteses de Gauss-Markov são elencadas abaixo:

- A esperança matemática para o vetor de resíduos,  $\xi$ , indexado por  $i \in \mathbb{N}$ , deve ser nula para qualquer  $i$ :

$$E\{\xi_i\} = 0, \quad \forall i; \quad (5.41)$$

- A variância do vetor de resíduos,  $\xi$ , indexado por  $i \in \mathbb{N}$ , deve ser constante para qualquer  $i$ :

$$\text{Var}\{\xi_i\} = \sigma^2, \quad \forall i; \quad (5.42)$$

- A covariância entre os termos  $\xi_i$  e  $\xi_j$ , do vetor de resíduos, deve ser nula para  $i \neq j$ :

$$\text{Cov}\{\xi_i, \xi_j\} = 0, \quad \forall i \neq j, \quad (5.43)$$

em que  $i$  e  $j \in \mathbb{N}$ .

A polarização do estimador, também chamada de *bias*, é definida como:

$$b = E\{\hat{\theta}\} - \theta, \quad (5.44)$$

em que  $b$  é a polarização,  $\hat{\theta}$  é o vetor de parâmetros estimado e  $\theta$  é o vetor de parâmetros real.

A polarização do estimador indica a existência de um erro sistemático, que não pode ser removido através do aumento do número de observações,  $N$  (LEMMA; RAMASAMY; SHUHAIMI, 2010). Assim, nos estimadores livres de polarização, a seguinte equação é verdadeira:

$$\lim_{N \rightarrow +\infty} E\{\hat{\theta}\} = \theta, \quad (5.45)$$

ou seja, o valor estimado aproxima-se do valor real conforme o número de observações tende ao infinito.

O método explicitado nesta seção é conhecido como MMQ clássico. Existem outras versões deste método, tais como o MMQ ponderado e o MMQ estendido, elaboradas para solucionar problemas que não são passíveis de solução através do MMQ clássico (por exemplo, sistemas cujas medições são contaminadas por ruído não branco) (AGUIRRE, 2015).

## 5.7 MÉTODO DOS MÍNIMOS QUADRADOS ORTOGONAL

O MMQ, apresentado na seção passada, é uma das ferramentas mais básicas em identificação. Sua grande utilidade deriva de sua simplicidade e de sua aplicabilidade em problemas de diversas naturezas. Entretanto, quando a matriz de regressão,  $\Psi$ , possui um número elevado de colunas, o que é comum na estimação de modelos não lineares, o uso da matriz pseudoinversa, (5.37), pode resultar em estimativas deficientes para os parâmetros desejados (AGUIRRE, 2015).

O método dos **M**ínimos **Q**uadrados **O**rtogonais, MQO, propõe a decomposição da matriz  $\Psi$  conforme (5.46).

$$\Psi = WA, \quad (5.46)$$

em que  $W$  é uma matriz com colunas ortogonais<sup>3</sup>, obtida através da ortogonalização de  $\Psi$ , e  $A$  é uma matriz triangular superior.

Existem diferentes algoritmos criados para ortogonalizar o espaço de colunas de uma matriz. A seguir, será explicitado o método de Gram-Schmidt.

<sup>3</sup>Dois vetores são ortogonais, se  $\langle x, y \rangle = 0$ , em que  $\langle \cdot \rangle$  é o produto interno entre os vetores  $x$  e  $y$  (LIMA, 2008).

O algoritmo de Gram-Schmidt calcula as matrizes  $A$  e  $W$  através dos seguintes passos:

$$\begin{cases} w_1 = \psi_1 \\ a_{j,i} = \frac{\langle w_i, \psi_j \rangle}{\langle w_i, w_i \rangle} \quad 1 \leq j < i \\ w_i = \psi_i - \sum_{j=1}^{i-1} a_{j,i} w_j, \end{cases} \quad (5.47)$$

em que  $w_i$  e  $\psi_i$  são vetores que representam as colunas das matrizes  $W$  e  $\Psi$ , respectivamente. A matriz  $A$  é dada por (5.48).

$$A = \begin{bmatrix} 1 & a_{1,2} & a_{1,3} & \cdots & a_{1,n_\theta} \\ 0 & 1 & a_{2,3} & \cdots & a_{2,n_\theta} \\ \vdots & \ddots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & 1 & a_{n_\theta-1,n_\theta} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (5.48)$$

em que  $A \in \mathbb{R}^{n_\theta}$  e  $n_\theta$  é a dimensão do vetor de parâmetros.

Considerando que  $\mathbf{y} = \Psi \hat{\boldsymbol{\theta}}$ , a equação normal é dada por (5.49).

$$\Psi^T \mathbf{y} = \Psi^T \Psi \hat{\boldsymbol{\theta}}. \quad (5.49)$$

A substituição de (5.46) em (5.49) resulta em:

$$A^T W^T \mathbf{y} = A^T W^T W A \hat{\boldsymbol{\theta}}. \quad (5.50)$$

Uma vez que  $W$  é uma matriz cujas colunas são ortogonais entre si,  $W^T W = D$ , sendo  $D$  uma matriz diagonal, definida positiva (STRANG, 2009). Assim:

$$\begin{aligned} A^T W^T \mathbf{y} &= A^T D A \hat{\boldsymbol{\theta}} \\ W^T \mathbf{y} &= D A \hat{\boldsymbol{\theta}} \\ D^{-1} W^T \mathbf{y} &= A \hat{\boldsymbol{\theta}}, \end{aligned} \quad (5.51)$$

definindo-se um vetor de parâmetros auxiliar  $\hat{\mathbf{g}} = A \hat{\boldsymbol{\theta}}$ , tem-se:

$$\hat{\mathbf{g}} = D^{-1} W^T \mathbf{y}. \quad (5.52)$$

O vetor de parâmetros original,  $\hat{\boldsymbol{\theta}}$ , é calculado por:

$$\hat{\boldsymbol{\theta}} = A^{-1} \hat{\mathbf{g}}. \quad (5.53)$$

Também é possível efetuar o cálculo de  $\hat{g}$  com o uso de (5.54).

$$\hat{g}_i = \frac{\langle w_i, y \rangle}{\langle w_i, w_i \rangle}, \quad (5.54)$$

em que  $i = 1, 2, \dots, n_\theta$  e  $n_\theta$  é a dimensão do vetor de parâmetros original.

Na próxima seção, a taxa de redução de erro, uma técnica de ranqueamento de regressores, será explicitada.

## 5.8 TAXA DE REDUÇÃO DE ERRO

Os métodos apresentados nas seções anteriores, MMQ e MQO, assumem que as variáveis independentes do problema de identificação já foram determinadas. Ou seja, para aplicar estes métodos, os regressores que compõem o modelo devem ser conhecidos. Isto equivale a dizer que a estrutura do modelo já foi determinada, e o projetista pode prosseguir para a estimação de parâmetros.

A aplicação do MQO em conjunto com a taxa de redução de erro, *Error Reduction Ratio* - ERR, é uma ferramenta muito aplicada na determinação de estrutura de modelos NARMAX (BILLINGS, 2013). Para definir a ERR, considere o seguinte modelo NARMAX:

$$\begin{aligned} y(k) &= \psi(k)^T \hat{\theta} + \xi(k) \\ &= \sum_{i=1}^{n_\theta} \hat{\theta}_i \psi_i(k) + \xi(k), \end{aligned} \quad (5.55)$$

em que  $\hat{\theta} \in \mathbb{R}^{n_\theta}$  é o vetor de parâmetros estimado,  $\xi(k)$  é o resíduo para a  $k$ -ésima aproximação de  $y$ ,  $i \in \mathbb{N}$  e  $\psi_i$  é a  $i$ -ésima coluna da matriz de regressão.

Pode-se construir o seguinte modelo auxiliar, a partir da aplicação do MQO, que ortogonaliza as colunas da matriz de regressão,  $\Psi$ :

$$y(k) = \sum_{i=1}^{n_\theta} \hat{g}_i w_i(k) + \xi(k), \quad (5.56)$$

em que  $\hat{g} \in \mathbb{R}^{n_\theta}$  é o vetor de parâmetros estimado através do MQO e  $w_i$  é a  $i$ -ésima coluna da matriz de regressores  $W$ , obtida através da ortogonalização de  $\Psi$ . Já que as colunas de  $W$  são ortogonais, tem-se que  $\langle w_i, w_j \rangle = 0$  para  $i \neq j$ .

A soma dos valores quadráticos de  $y(k)$  pode ser calculada por meio do produto interno  $\langle \mathbf{y}, \mathbf{y} \rangle$  ou  $\mathbf{y}^T \mathbf{y}$ :

$$\begin{aligned} \frac{1}{N} \sum_{k=1}^N y(k)^2 &= \frac{1}{N} \langle \mathbf{y}, \mathbf{y} \rangle \\ \langle \mathbf{y}, \mathbf{y} \rangle &= \sum_{i=1}^{n_\theta} \hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle + \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle, \end{aligned} \quad (5.57)$$

em que se considerou a ortogonalidade entre os regressores,  $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = 0$ , e entre os regressores e os resíduos,  $\langle \mathbf{w}_i, \boldsymbol{\xi} \rangle = 0$  (AGUIRRE, 2015).

A Equação (5.57) diz que a soma dos valores quadráticos de  $y(k)$  pode ser explicada através da soma dos valores quadráticos de cada regressor, respectivamente ponderados pelos quadrados de seus parâmetros, desde que se use uma base ortogonal. A parcela  $\langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle$  corresponde à parcela não explicada pelos regressores (BILLINGS, 2013). Deste modo, pode-se usar (5.57) para quantificar a importância de cada regressor, de forma individual.

Quando se inclui o  $i$ -ésimo regressor ao modelo, a parcela de  $\langle \mathbf{y}, \mathbf{y} \rangle$  que este regressor explica é  $\hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle$ . Deste modo, pode-se definir a taxa de redução de erro devido à inclusão do  $i$ -ésimo regressor (AGUIRRE, 2015):

$$[ERR]_i = \frac{\hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}. \quad (5.58)$$

O valor de  $ERR_i$  em (5.58) encontra-se dentro do intervalo  $[0, 1]$ . A ERR individual pode ser interpretada como a porcentagem da saída  $\mathbf{y}$  que é explicada pelo regressor  $\mathbf{w}_i$  (AGUIRRE, 2015).

Sendo assim, pode-se definir um espaço de regressores candidatos e aplicar MQO associado à ERR para determinar quais regressores são mais relevantes para o modelo.

## 5.9 MÉTODOS DE PREDIÇÃO DE ERRO

Métodos de predição de erro, *Prediction Error Methods* - PEM, determinam o vetor de parâmetros através da minimização da seguinte função de custo (ASTRÖM,

1980):

$$J_{PEM}(\hat{\theta}, Z^N) = \sum_{k=1}^N \xi(k|k-1, \hat{\theta})^2 \quad (5.59)$$

$$\xi(k|k-1, \hat{\theta}) = y(k) - \hat{y}(k|k-1),$$

em que  $Z^N = \{u(1), y(1), \dots, u(N), y(N)\}$  é o conjunto de dados,  $\hat{y}(k|k-1)$  é a predição de um passo à frente da saída e  $\xi(k|k-1, \hat{\theta})$  é o erro de um passo à frente.  $\hat{y}(k|k-1)$  indica que o valor da saída no instante  $k$  foi determinado a partir de informação disponível até o instante  $k-1$ .

Considere o seguinte modelo geral:

$$y(k) = \psi_{yu}^\top(k-1) \hat{\theta} + \xi(k), \quad (5.60)$$

$$\psi_{yu}(k-1) = [y(k-1) \dots y(k-n_y) \ u(k-1) \dots u(k-n_u)],$$

em que  $\psi_{yu}(k-1)$  é um vetor com  $n_\theta$  variáveis regressoras de entrada,  $u$ , e de saída,  $y$ , tomadas até o instante  $(k-1)$ . Se o vetor  $\psi_{yu}(k-1)$  for composto pelos regressores indicados em (5.60), o modelo é do tipo ARX, e sua predição de um passo à frente pode ser escrita como (5.61).

$$\hat{y}(k|k-1) = \psi_{yu}^\top(k-1) \hat{\theta}. \quad (5.61)$$

Para as classes AR e ARX, o estimador que minimiza (5.59) é o estimador de mínimos quadrados. Deste modo, pode-se considerar que o MMQ é um estimador que pertence à classe PEM.

Quando a estimação de parâmetros precisa ser realizada por algoritmos recursivos, o erro de predição é chamado de inovação e não é igual ao resíduo (AGUIRRE, 2015). Considerando a Equação (5.62), a parcela  $G[y(k) - \hat{y}(k|k-1)]$  representa a correção feita pelo algoritmo recursivo na estimativa anterior  $\hat{y}(k-1)$ . Esta correção é conhecida como inovação.

$$\hat{y}(k) = \hat{y}(k-1) + G[y(k) - \hat{y}(k|k-1)], \quad (5.62)$$

em que  $G \in \mathbb{R}$  é o ganho utilizado para corrigir a estimativa  $\hat{y}(k-1)$ .

## 5.10 ERRO MÉDIO QUADRÁTICO

Para mensurar o grau de adequação de um modelo aos dados empíricos amostrados a partir do sistema que se pretende representar, pode-se recorrer a algumas métricas. Neste trabalho, o *Mean Squared Error*, MSE, foi escolhido como métrica, devido à ausência de polarização e à variância mínima (ALSMADI, 2003).

Para ter mínima variância, os valores estimados devem se concentrar em valores próximos ao real. Um método utilizado para testar estimadores que combina ambos, polarização tendendo a zero conforme  $N$  cresce e variância mínima, é o MSE, dado como (ALSMADI; WILKES, 2002):

$$MSE(\hat{\beta}) = E\{(\hat{\beta} - \beta)^2\}, \quad (5.63)$$

em que  $\beta$  é uma VA e  $\hat{\beta}$  é o valor estimado desta VA.

O MSE também pode ser expresso por:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y - \hat{y})^2, \quad (5.64)$$

em que  $y, \hat{y} \in \mathbb{R}$ .

## 5.11 CONSIDERAÇÕES FINAIS

Neste capítulo foi estabelecida a base teórica para a compreensão do método de LM, aplicado nos processos de otimização deste trabalho. Foram embasados os mecanismos de busca exata em linha, *backtracking* e região de confiança. Os principais conceitos envolvendo os métodos de Newton, Gauss-Newton e Levenberg-Marquardt foram apresentados, bem como os algoritmos básicos destes métodos. Ainda, foi visto que o método dos mínimos quadrados possui ampla aplicabilidade na estimação de parâmetros. Sua versão ortogonal pode ser aplicada tanto na estimação de parâmetros, quanto na determinação de estrutura, associado à taxa de redução de erro. Os métodos de predição de erro também foram conceituados, uma vez que são frequentemente aplicados na literatura e apresentam relevância especial para modelos cuja finalidade é predição de comportamento. Por fim, conceituou-se o MSE, que

será aplicado como índice de medição da adequação dos modelos aos dados coletados do sistema.

## 6 MODELAGEM CAIXA PRETA DO SET

Neste capítulo, os passos seguidos para a obtenção de um modelo caixa preta do transistor mono-elétron serão apresentados. Foram realizadas várias simulações, a fim de obter modelos com estrutura simples e que representassem adequadamente o SET.

O primeiro problema a ser tratado diz respeito à escolha da representação matemática adequada. Conforme visto na Seção 3.6, há várias opções de classes de modelos que podem ser utilizadas. Desta forma, as primeiras simulações a serem descritas têm caráter investigativo, no sentido de determinar qual classe seria mais adequada à representação matemática do transistor. Desta forma, a configuração destas simulações é simples: sem inclusão de ruído, à temperatura  $T = 0$  K e sem variação da tensão de polarização,  $V_{DS}$ . Uma vez evidenciada a classe de modelos que obteve o melhor desempenho, segundo as métricas consideradas, procedeu-se à construção de modelos que considerassem a inclusão de ruído, a variação de  $V_{DS}$  e em temperaturas de operação diferentes.

### 6.1 ESCOLHA DA REPRESENTAÇÃO MATEMÁTICA

As próximas seções dedicam-se a explicar os passos seguidos na realização das simulações cujos resultados orientaram a escolha da representação matemática ou classe de modelos. Estes passos serão elencados de acordo com a sequência de etapas da modelagem caixa preta, apresentada no Capítulo 3.

#### 6.1.1 Coleta de Dados

Conforme explicitado na Seção 2.2, o SET é classificado como um sistema não linear. Deste modo, a fim de que as dinâmicas do transistor sejam excitadas

adequadamente, é necessário escolher um sinal com um rico perfil de amplitudes. Considerando a discussão apresentada na Seção 3.2, buscou-se um sinal que efetivamente excitasse as duas principais dinâmicas do transistor: o bloqueio de Coulomb e as oscilações de Coulomb.

Para que o transistor apresente regiões de oscilação em sua saída, é preciso que  $V_{DS}$  seja mantido em valor constante e que haja uma variação monotônica em sua tensão de *gate*, ver Seção 2.1.5. Deste modo, um sinal PRMLS puro não seria adequado para a identificação do SET, uma vez que não possui variação monotônica. A solução encontrada consiste em inserir, de forma aleatória, rampas no sinal PRMLS original. O sinal resultante, mostrado na Figura 36, será chamado de PRMLS híbrido.

A Figura 37 apresenta um esquemático das simulações efetuadas para coleta de dados neste trabalho. Na primeira simulação, o sinal da Figura 36 foi aplicado como tensão de *gate*,  $V_{GS}$ . Os parâmetros do transistor utilizados estão descritos na Tabela 3.

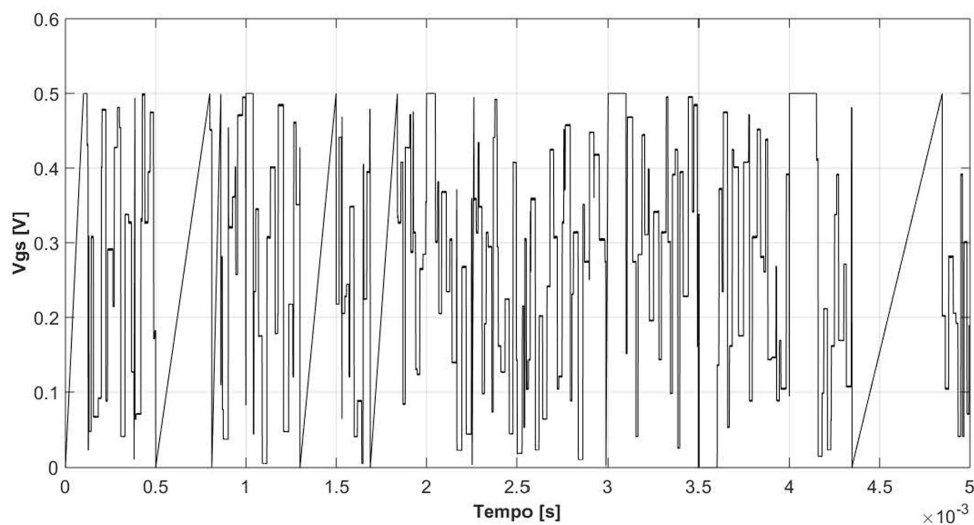


Figura 36: Sinal de entrada aplicado ao transistor mono-elétron na etapa de estimação.

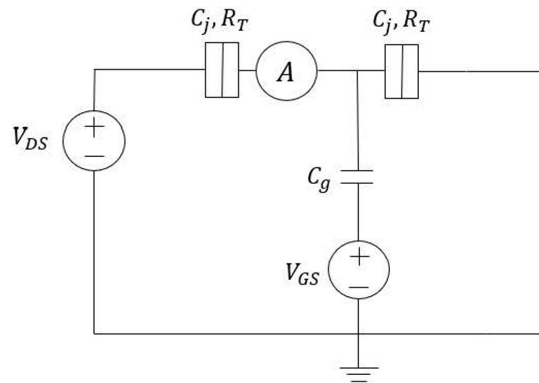


Figura 37: Esquemático da simulação de coleta de dados para a modelagem caixa preta do SET.  $V_{GS}$  é o sinal mostrado na Figura 36. O transistor é simétrico, ou seja, as junções-túnel possuem os mesmos parâmetros,  $C_j = C_g = 1$  aF e  $R_T = 100$  k $\Omega$ .  $V_{DS} = 20$  mV é a tensão de polarização do transistor. A saída é a corrente entre dreno e fonte, medida pelo amperímetro.

Tabela 3: Parâmetros do transistor mono-elétron utilizados na coleta de dados.

Parâmetro	Valor
$C_g, C_{j1}, C_{j2}$	1 aF
$R_{T1}, R_{T2}$	100 k $\Omega$
$V_{DS}$	20 mV
Temperatura	0 K

O circuito da Figura 37 foi simulado no SIMON. O período de amostragem,  $T_s$ , é de 100 ns. O número de amostras, nesta etapa, é igual a 50 mil. Desta forma, as sequências de entrada e de saída são representadas por vetores de dimensão 50 mil. O tempo total de simulação foi ajustado para acomodar 50 mil amostras com  $T_s = 100$  ns, ou seja, é igual a 5 ms. Tanto os parâmetros da Tabela 3 quanto o valor do período de amostragem foram escolhidos com base no circuito modelo, proposto por Wasshuber, para a demonstração das características de oscilação e de bloqueio de Coulomb (WASSHUBER, 2001).

A resposta do transistor ao sinal mostrado na Figura 36 é ilustrada na Figura 38.

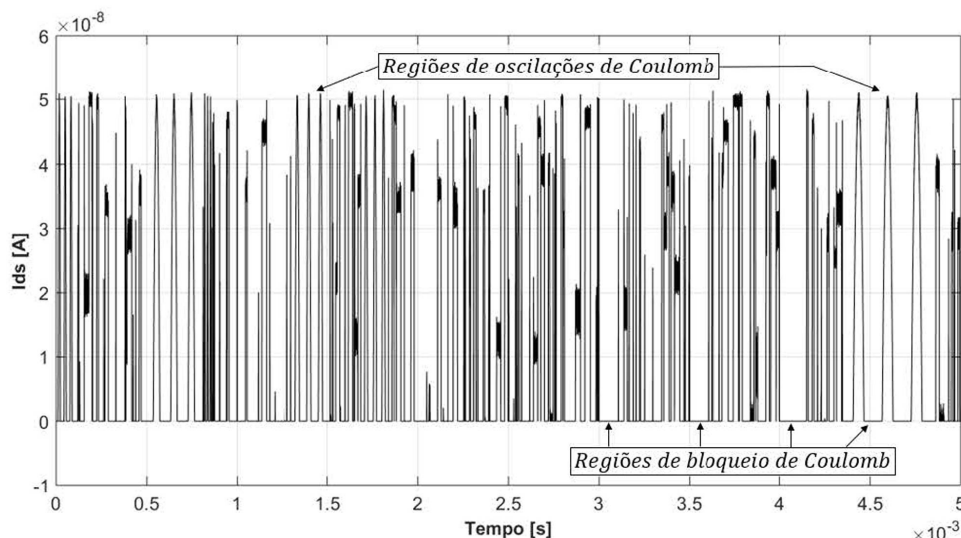


Figura 38: Corrente entre dreno e fonte,  $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 3 quando a tensão  $V_{GS}$  mostrada na Figura 36 é aplicada à sua entrada.

Na Figura 38, algumas regiões de bloqueio e de oscilações de Coulomb foram marcadas com o propósito de ressaltar que o sinal coletado na saída do transistor, de fato, apresenta as principais dinâmicas do SET. É válido pontuar que a saída apresenta oscilações em todos os intervalos em que há variação monotônica na entrada (rampas). Deste modo, pode-se concluir que os sinais coletados são adequados para a identificação do transistor mono-elétron nas condições delimitadas pela simulação. Sendo assim, pode-se prosseguir para a etapa de determinação de estrutura do modelo.

### 6.1.2 Determinação de Estrutura

A determinação de estrutura, conforme apresentada na Seção 3.3, consiste, por exemplo, em determinar o número de polos de uma função de transferência. Entretanto, antes que se parta para esta abordagem, é necessário saber se uma função de transferência é adequada para modelar o sistema desejado. Sendo assim, primeiramente, deve-se escolher uma representação matemática para, posteriormente, determinar a estrutura desta representação.

Uma vez que o transistor mono-elétron é um sistema não linear, as representações matemáticas consideradas para sua modelagem são não lineares, ver Seção 3.6. As

seguintes representações foram testadas: NARX, Hammerstein, Wiener e Hammerstein-Wiener. As métricas utilizadas para a comparação das representações foram o MSE e o tempo médio demandado para a simulação da etapa de validação do modelo.

Os dados coletados na etapa anterior foram utilizados para estimar modelos pertencentes às classes NARX, Hammerstein, Wiener e Hammerstein-Wiener. O melhor modelo obtido em cada uma dessas classes foi selecionado de acordo com o menor MSE. Por fim, os melhores modelos foram comparados em relação ao MSE e ao tempo de simulação. O modelo com os menores MSE e tempo de simulação foi selecionado para representar o SET. O *benchmark* utilizado para o cálculo do MSE foi a curva de saída do SET, obtida através do SIMON. A Figura 39 ilustra o esquemático de escolha da classe de modelos.

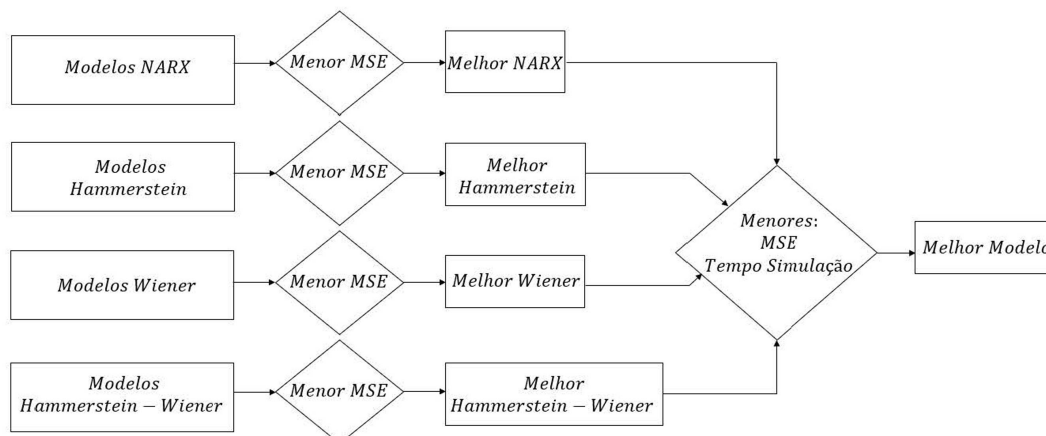


Figura 39: Esquemático detalhando os passos e métricas que guiaram a escolha da representação matemática para a modelagem caixa preta do SET.

Conforme já discutido nos capítulos anteriores, a determinação de estrutura de um modelo é a etapa que faz com que a Identificação de Sistemas seja muitas vezes chamada de “arte” (AGUIRRE, 2015). Isto ocorre pela falta de uma técnica consolidada que auxilie o projetista a determinar a estrutura de modelos matemáticos (BILLINGS, 2013).

Para o modelo NARX, foram aplicados os métodos MQO, descrito na Seção 5.7, associado à ERR, descrita na seção 5.8, para o ranqueamento de regressores. O espaço de regressores candidatos foi construído a partir das restrições expressas na Tabela 4.

Tabela 4: Restrições aplicadas à construção do espaço de regressores candidatos para o modelo NARX.  $n_u$  e  $n_y$  representam as regressões máximas de entrada e de saída, respectivamente.

Regressão Máxima	Ordem de <i>kernel</i> máxima
$n_u = n_y = 4$	2

É válido ressaltar que foram testadas ordens de regressão até sete e ordens de *kernel* até três. As restrições expressas na Tabela 4 ilustram o caso para o qual se obteve o melhor modelo. As características das matrizes de regressão estão expostas na Tabela 5.

Tabela 5: Matrizes de Regressão para o espaço de regressores candidatos.

Matriz	Ordem de “kernel”	Regressores	Dimensão (linhas $\times$ colunas)
$M_u$	1	$u(k - i_u)$ $i_u = 0, 1, \dots, n_u$	$49996 \times 5$
$M_y$	1	$y(k - i_y)$ $i_y = 1, 2, \dots, n_y$	$49996 \times 4$
$M_{uu}$	2	$u(k - i_u)u(k - j_u)$ $i_u, j_u = 0, 1, \dots, n_u$	$49996 \times 15$
$M_{yy}$	2	$y(k - i_y)y(k - j_y)$ $i_y, j_y = 1, 2, \dots, n_y$	$49996 \times 10$
$M_{uy}$	2	$u(k - i_u)y(k - i_y)$ $i_u = 0, 1, \dots, n_u$ $i_y = 1, 2, \dots, n_y$	$49996 \times 20$

A construção das matrizes de regressão é explicitada nas Equações (6.1)-(6.5). A notação foi simplificada de tal forma que  $u(k - 1) = u_{k-1}$ . Cabe ressaltar que os elementos das matrizes representam colunas, ou seja, o regressor  $u_{k-1}$ , por exemplo, não é um escalar, mas sim um vetor tal que  $u_{k-1} \in \mathbb{R}^{49996}$ . O número 49996 vem do fato de que, uma vez que a regressão máxima é igual a quatro, as quatro primeiras amostras são utilizadas como condições iniciais.

$$M_u = \begin{bmatrix} \mathbf{u}_k & \mathbf{u}_{k-1} & \mathbf{u}_{k-2} & \mathbf{u}_{k-3} & \mathbf{u}_{k-4} \end{bmatrix}. \quad (6.1)$$

$$M_y = \begin{bmatrix} \mathbf{y}_{k-1} & \mathbf{y}_{k-2} & \mathbf{y}_{k-3} & \mathbf{y}_{k-4} \end{bmatrix}. \quad (6.2)$$

$$M_{uu} = \begin{bmatrix} \mathbf{u}_k \mathbf{u}_k & \cdots & \mathbf{u}_k \mathbf{u}_{k-4} & \mathbf{u}_{k-1} \mathbf{u}_{k-1} & \cdots & \mathbf{u}_{k-1} \mathbf{u}_{k-4} \\ \mathbf{u}_{k-2} \mathbf{u}_{k-2} & \cdots & \mathbf{u}_{k-2} \mathbf{u}_{k-4} & \mathbf{u}_{k-3} \mathbf{u}_{k-3} & \mathbf{u}_{k-3} \mathbf{u}_{k-4} & \mathbf{u}_{k-4} \mathbf{u}_{k-4} \end{bmatrix}. \quad (6.3)$$

$$M_{yy} = \begin{bmatrix} \mathbf{y}_{k-1} \mathbf{y}_{k-1} & \mathbf{y}_{k-1} \mathbf{y}_{k-2} & \cdots & \mathbf{y}_{k-1} \mathbf{y}_{k-4} & \mathbf{y}_{k-2} \mathbf{y}_{k-2} \\ \cdots & \mathbf{y}_{k-2} \mathbf{y}_{k-4} & \mathbf{y}_{k-3} \mathbf{y}_{k-3} & \mathbf{y}_{k-3} \mathbf{y}_{k-4} & \mathbf{y}_{k-4} \mathbf{y}_{k-4} \end{bmatrix}. \quad (6.4)$$

$$M_{uy} = \begin{bmatrix} \mathbf{u}_k \mathbf{y}_{k-1} & \cdots & \mathbf{u}_{k-1} \mathbf{y}_{k-1} & \cdots & \mathbf{u}_{k-1} \mathbf{y}_{k-4} & \mathbf{u}_{k-2} \mathbf{y}_{k-1} & \cdots \\ \mathbf{u}_{k-2} \mathbf{y}_{k-4} & \cdots & \mathbf{u}_{k-3} \mathbf{y}_{k-1} & \cdots & \mathbf{u}_{k-3} \mathbf{y}_{k-4} & \cdots & \mathbf{u}_{k-4} \mathbf{y}_{k-4} \end{bmatrix}. \quad (6.5)$$

A matriz de regressão,  $\Psi_{NARX}$ , é construída pela concatenação das matrizes constantes na Tabela 5. A dimensão da matriz  $\Psi_{NARX}$ , expressa em (6.6), é  $49996 \times 44$ .

$$\Psi_{NARX} = \begin{bmatrix} M_u & M_y & M_{uu} & M_{yy} & M_{uy} \end{bmatrix}. \quad (6.6)$$

Considerando a ordem máxima de regressão igual a quatro, o número de amostras  $N = 50$  mil e que  $\Psi_{NARX} \in \mathbb{R}^{49996 \times 44}$ , a equação matricial que se deseja resolver é dada por (6.7).

$$\begin{bmatrix} y(5) \\ \vdots \\ y(k) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} u(5) & \cdots & u(5)u(3) & \cdots & u(5)y(1) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ u(k) & \cdots & u(k)u(k-2) & \cdots & u(k)y(k-4) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ u(N) & \cdots & u(N)u(N-2) & \cdots & u(N)y(N-4) \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{44} \end{bmatrix}, \quad (6.7)$$

em que  $\theta \in \mathbb{R}^{44}$ .

Conforme já mencionado, o MQO foi utilizado em conjunto com a taxa de redução de erro, a fim de ranquear os melhores regressores e reduzir a dimensão do problema expresso em (6.7). Os cinco melhores regressores foram escolhidos para compor o modelo. Deste modo, após a aplicação do MQO, a matriz de regressão possui dimensão  $49996 \times 5$  e  $\theta \in \mathbb{R}^5$ .

A determinação de estrutura para os modelos de Hammerstein, de Wiener e de Hammerstein-Wiener envolve, além da escolha dos regressores, a escolha dos *breakpoints* das não linearidades. Neste trabalho, serão utilizadas não linearidades do tipo PWL, ver Seção 3.6.3.4. Para determinar qual tipo de não linearidade seria utilizada foi considerada a simplicidade da estrutura. Testes realizados através do MatLab, versão R2017a, comprovaram a adequação deste tipo de não linearidade na modelagem do transistor mono-elétron, os resultados destes testes serão apresentados no próximo capítulo.

Uma vez escolhido o tipo de não linearidade, para que a estrutura dos modelos orientados a blocos esteja completa, resta determinar o número de segmentos lineares na função PWL e os parâmetros da função de transferência que representa o bloco dinâmico linear. Ou seja, quantos polos, zeros e atrasos terá essa função de transferência. Existem infinitas combinações possíveis para estas variáveis e nenhuma metodologia bem estabelecida que auxilie o projetista nesta escolha (AGUIRRE, 2015). Deste modo, em Identificação de Sistemas, é comum que se chegue empiricamente a uma aproximação adequada (COELHO, 2016).

A Tabela 6 elenca quais são as características a serem determinadas para cada modelo de blocos concatenados. Foram realizados diversos testes com diferentes combinações destas características. A métrica utilizada na escolha da melhor estrutura foi o MSE.

Tabela 6: Características a serem determinadas na estrutura de modelos de blocos concatenados.  $L_{in}$  e  $L_{out}$  representam o número de segmentos lineares nas não linearidades de entrada e de saída, respectivamente.  $n_a$ ,  $n_b$  e  $n_k$  são a quantidade de polos, zeros e atrasos puros, respectivamente.

Bloco	Hammerstein	Wiener	Hammerstein-Wiener
Não Linearidade de Entrada	$L_{in}$	-	$L_{in}$
Bloco Dinâmico Linear	$n_a, n_b$ e $n_k$	$n_a, n_b$ e $n_k$	$n_a, n_b$ e $n_k$
Não Linearidade de Saída	-	$L_{out}$	$L_{out}$

A diferença na metodologia aplicada para obtenção de estrutura do modelo NARX polinomial e dos modelos de blocos concatenados se deve a dois motivos principais. Existem, atualmente, diversos trabalhos que debatem a determinação de estrutura de modelos NARX polinomiais (BILLINGS, 2013; ALVES, 2009). Já, para modelos orientados a blocos, a determinação da função não linear dificulta a obtenção da estrutura (GIRI; BAI, 2010). Deste modo, a determinação de estrutura de modelos de Hammerstein-Wiener, por exemplo, é um tema pouco frequente na literatura (BAI, 2002). Um segundo ponto relevante é o fato de que a ferramenta de simulação disponível no MatLab R2017a para obtenção de modelos caixa preta não apresenta a opção de estrutura de modelos NARX puramente polinomiais, o que impossibilita a realização de testes. A próxima seção detalha os métodos aplicados na estimação de parâmetros dos modelos.

### 6.1.3 Estimação de Parâmetros

A estimação de parâmetros dos modelos cujas estruturas foram escolhidas na etapa anterior foi tratada como um problema de otimização. Deste modo, foram escolhidas funções de custo e métodos de minimização para definir os parâmetros.

Para o modelo NARX, a função de custo é  $J_{MQ}$ , dada pela Equação (5.31). O MMQ clássico foi aplicado para realizar a minimização da função de custo.

Para os modelos de Hammerstein, Wiener e de Hammerstein-Wiener, o número de parâmetros a serem determinados é maior, uma vez que é preciso definir os *breakpoints* das funções PWL. Ou seja, é preciso estabelecer em que ponto um segmento linear termina e outro começa. Especialmente nos modelos de Hammerstein-

Wiener, que possuem não linearidades estáticas na entrada e na saída, a estimação dos parâmetros não é trivial (BAI, 2002).

Para que o ajuste do modelo aos dados simulados fosse aprimorado, nos modelos orientados a blocos, a estimação de parâmetros foi feita de forma recursiva. O critério de parada utilizado foi o número de iterações,  $n_{it} = 40$ . A função de custo para estes modelos é  $J_{PEM}$ , dada pela Equação (5.59). O algoritmo de Levenberg-Marquardt, explicitado na Seção 5.5, foi aplicado para realizar a busca do mínimo de  $J_{PEM}$ . É válido ressaltar que a otimização foi feita com a restrição de que o módulo dos polos da função de transferência respeitasse a inequação  $|\text{polo}| < 1$ . Esta restrição foi imposta visando a estabilidade dos modelos obtidos.

Como ponto inicial do algoritmo de Levenberg-Marquardt foi utilizada a melhor solução linear, ou seja, a solução encontrada pelo MMQ é utilizada como ponto inicial de otimização. Esta configuração para a escolha do ponto inicial é inerente ao MatLab.

Uma vez que a estrutura e os parâmetros dos modelos foram obtidos, o próximo passo consiste em realizar a validação destes modelos.

#### 6.1.4 Validação dos Modelos

O método de validação escolhido neste trabalho é conhecido como validação dinâmica e busca checar se o modelo é capaz de seguir o comportamento do sistema, em diferentes condições de interesse (AGUIRRE; BILLINGS, 1995).

Para o transistor mono-elétron, é importante que o modelo reproduza regiões em que o dispositivo encontra-se em bloqueio e em condução (característica oscilatória da corrente entre dreno e fonte).

O primeiro passo para se obter a validação do modelo é determinar uma métrica a partir da qual a adequação do comportamento do modelo ao comportamento do sistema será mensurada. Neste trabalho, a métrica utilizada é o MSE.

Para validar os modelos, foi utilizado o sinal apresentado na Figura 40. Uma característica importante deste sinal é sua correlação com o sinal utilizado na etapa de estimação, apresentado na Figura 36. Já que o sinal da etapa de estimação representa o conjunto de dados com os quais o modelo foi treinado, para que a validação seja efetuada corretamente, a correlação entre os sinais utilizados na estimação e na

validação deve ser baixa. De forma simplificada, pode-se dizer que o coeficiente de correlação quantifica a relação estatística de dependência entre duas variáveis.

O coeficiente de correlação entre os dois sinais foi calculado através da função  $corrcoef(V_1, V_2)$  do MatLab R2017a, que retorna o coeficiente de correlação entre as variáveis  $V_1$  e  $V_2$ . O valor encontrado foi de  $corrcoef(V_{GS,est}, V_{GS,val}) = 0,0144$ . Quanto mais próximo de zero for o coeficiente de correlação, mais descorrelacionados são os sinais.

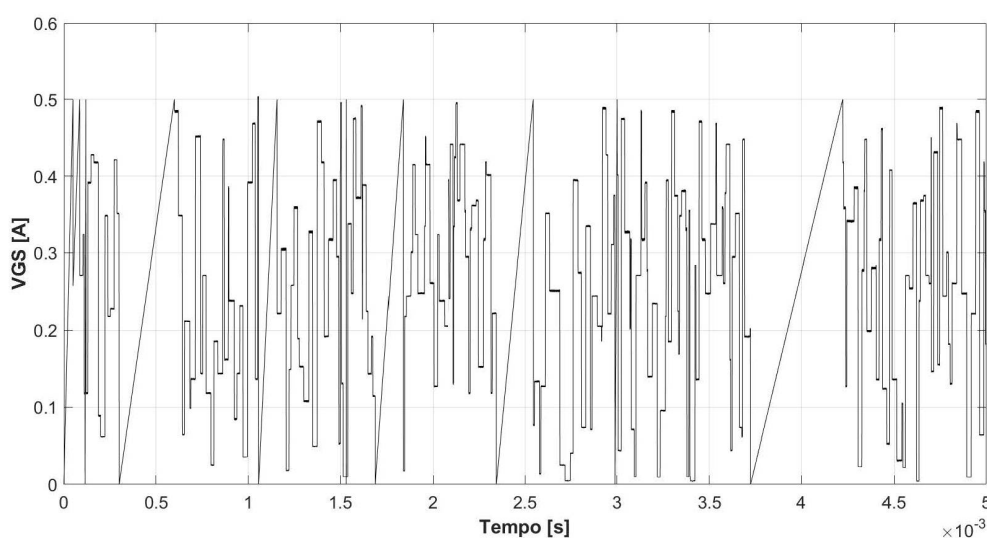


Figura 40: Sinal de entrada aplicado ao transistor mono-elétron na etapa de validação.

A resposta do SET ao sinal da Figura 40 é mostrada na Figura 41. É válido ressaltar que a simulação de coleta do conjunto de dados de validação possui a mesma configuração que a de estimação. Os parâmetros do transistor, período de amostragem e número de amostras são os mesmos que foram apresentados na seção 6.1.1.

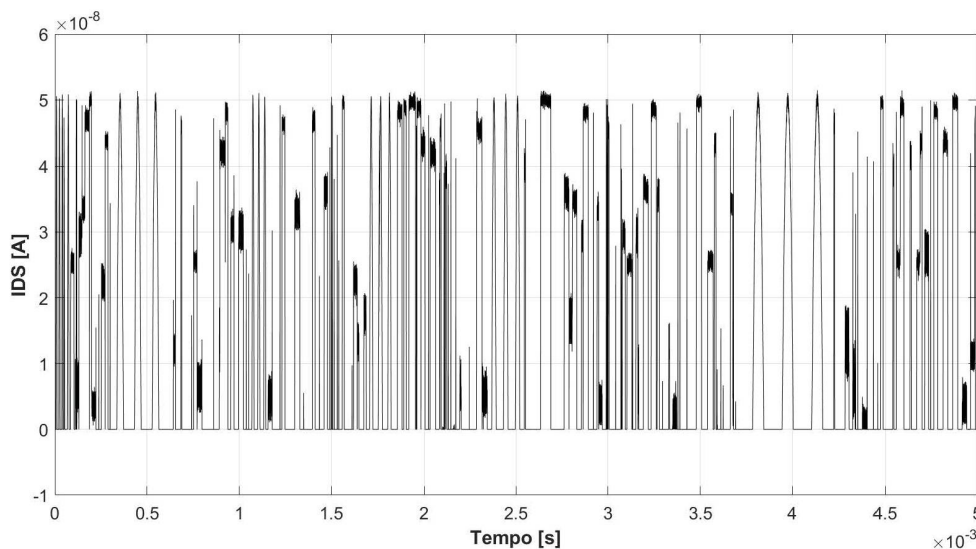


Figura 41: Corrente entre dreno e fonte,  $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 3 quando a tensão  $V_{GS}$  mostrada na Figura 40 é aplicada à sua entrada.

Após a realização da validação dos modelos NARX polinomial, de Hammersstein, de Wiener e de Hammerstein-Wiener, foi feita a comparação entre essas representações, a fim de determinar qual delas melhor se adapta à modelagem do SET. Na validação, foi calculado o MSE de cada modelo com relação à saída do sistema (a resposta do transistor, coletada por meio de simulação no SIMON), mostrada na Figura 41. Além do MSE, foi computado o tempo médio de simulação para cada modelo. A simulação de validação foi executada 50 vezes para cada modelo e o tempo médio calculado. Utilizou-se a função *tic toc* do Matlab R2017a a fim de computar o tempo demandado para a finalização da simulação. As simulações foram realizadas em um “notebook” Dell i7 com Windows 10. Estes resultados serão explicitados no próximo capítulo, ver Tabela 10. A representação matemática com a qual se construiu o modelo com menores MSE e tempo médio de simulação foi escolhida para representar o SET.

## 6.2 MODELOS: DIFERENTES CONDIÇÕES DE OPERAÇÃO

A seção anterior explicitou como foram feitos os testes que nortearam a escolha da representação matemática do transistor mono-elétron. Uma vez feita esta escolha, foram construídos modelos para condições diferentes de operação. Estas condições são representadas pelo aumento da temperatura e pela variação da tensão

de polarização do SET,  $V_{DS}$ . Nesta seção, são apresentadas as configurações das simulações feitas durante a obtenção destes modelos. Cabe pontuar que os métodos utilizados na estimação de parâmetros e validação destes modelos são idênticos aos descritos na Seção 6.1. Os sinais de entrada utilizados nas simulações são os mesmos das Figuras 36 e 40. A resposta do transistor mono-elétron varia conforme as condições de operação. As Figuras 42-46 ilustram a resposta do SET, simulada através do SIMON, para cada configuração de simulação.

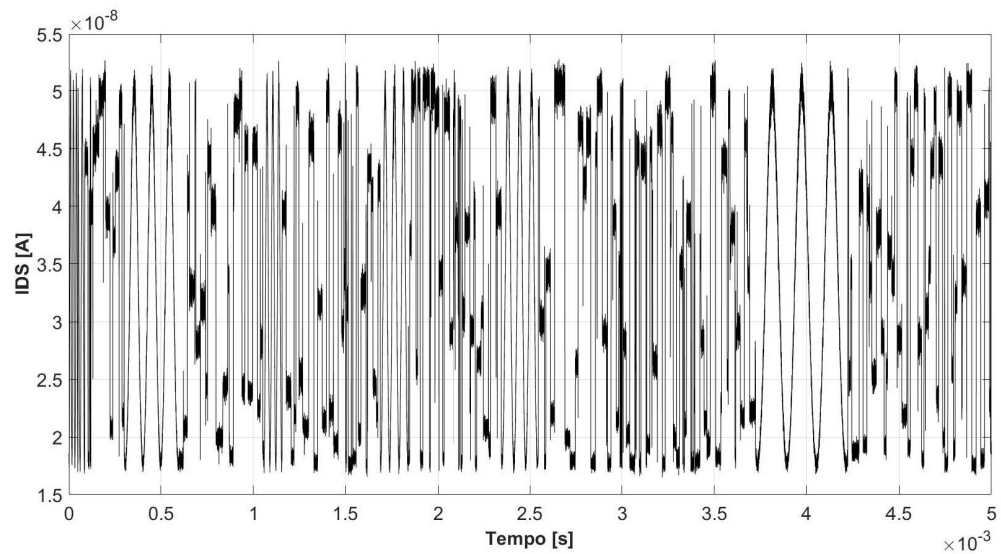


Figura 42: Corrente entre dreno e fonte,  $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão  $V_{GS}$  mostrada na Figura 40 é aplicada à sua entrada a  $T = 77$  K.

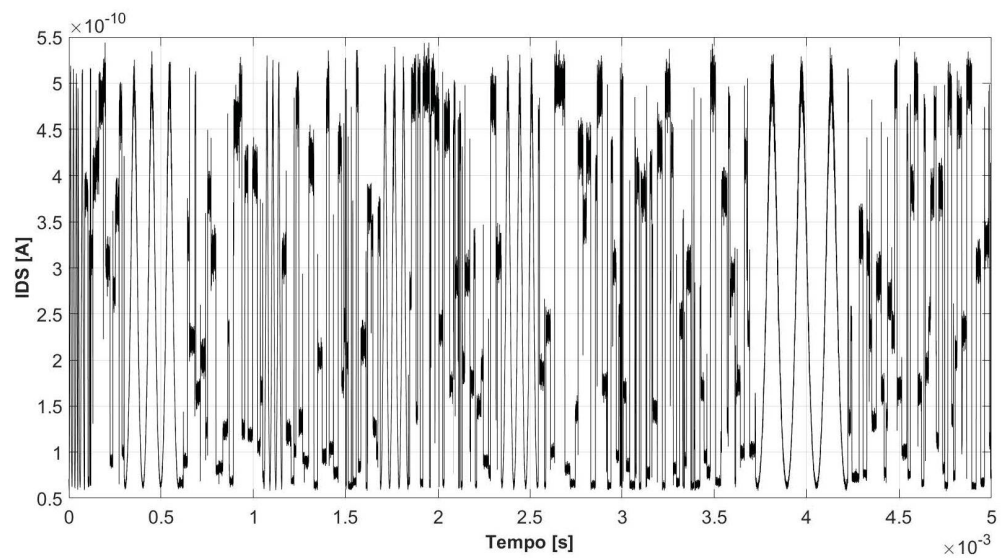


Figura 43: Corrente entre dreno e fonte,  $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão  $V_{GS}$  mostrada na Figura 40 é aplicada à sua entrada com  $T = 150$  K e  $V_{DS} = 20$  mV.

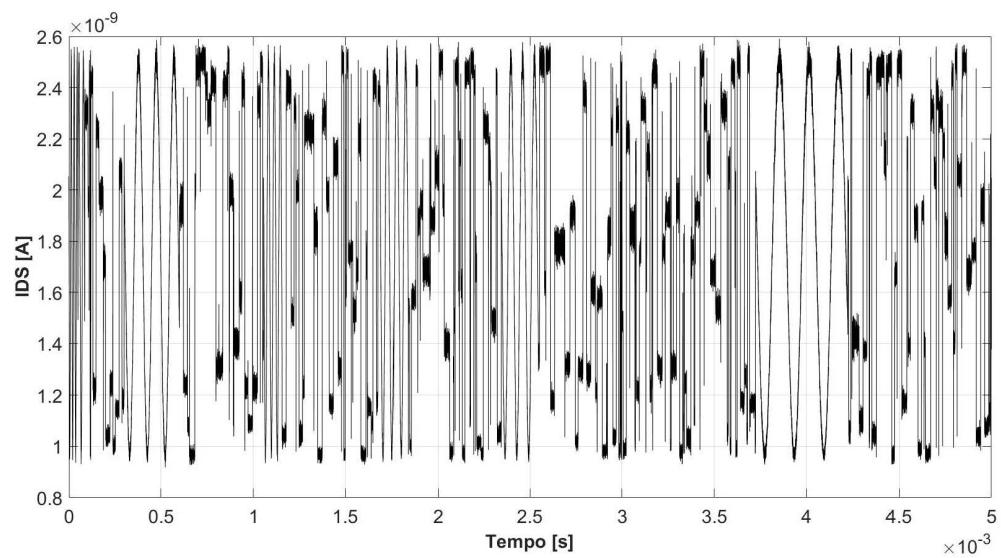


Figura 44: Corrente entre dreno e fonte,  $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão  $V_{GS}$  mostrada na Figura 40 é aplicada à sua entrada com  $T = 150$  K e  $V_{DS} = 100$  mV.

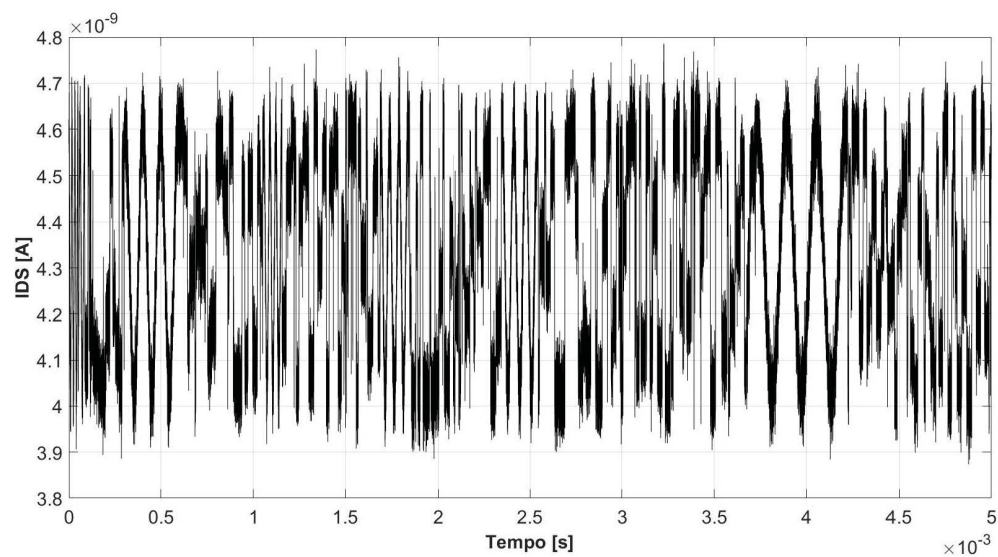


Figura 45: Corrente entre dreno e fonte,  $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão  $V_{GS}$  mostrada na Figura 40 é aplicada à sua entrada com  $T = 150$  K e  $V_{DS} = 180$  mV.

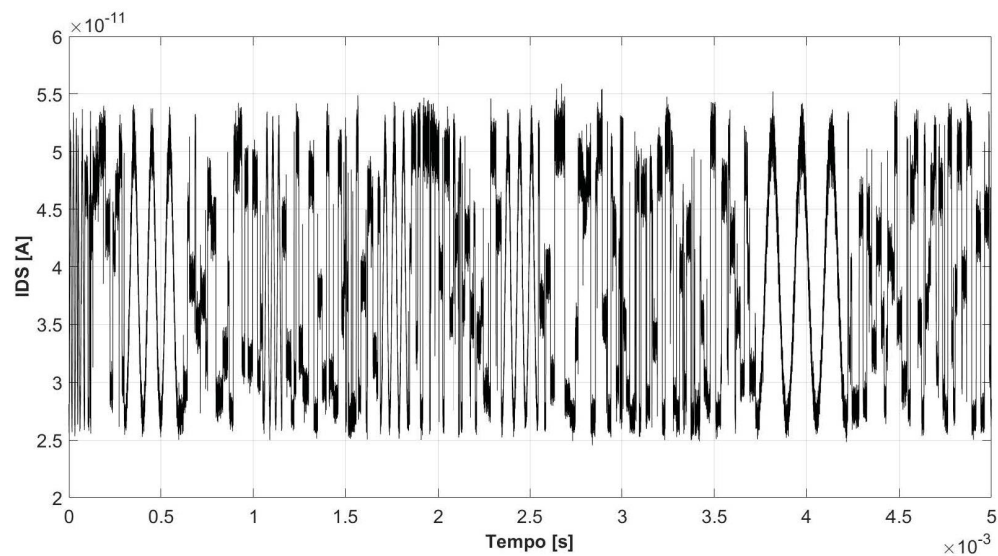


Figura 46: Corrente entre dreno e fonte,  $I_{DS}$ , para o SET com parâmetros detalhados na Tabela 7 quando a tensão  $V_{GS}$  mostrada na Figura 40 é aplicada à sua entrada com  $T = 298$  K e  $V_{DS} = 20$  mV.

As Figuras 47 e 48 ilustram a variação na resposta do transistor, conforme a temperatura e a tensão de polarização são alteradas. A Tabela 7 apresenta as

condições para as quais foram obtidos novos modelos. Cabe ressaltar que o SIMON inclui ruído<sup>1</sup> na simulação conforme a temperatura de operação aumenta (WASSHUBER, 2001).

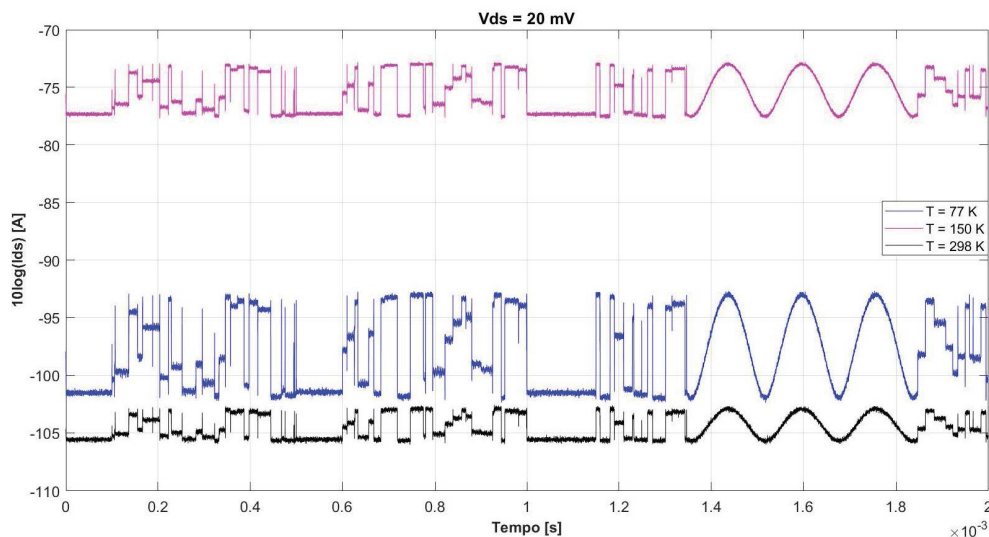


Figura 47: Corrente entre dreno e fonte,  $I_{DS}$ , para os diferentes valores de temperatura.

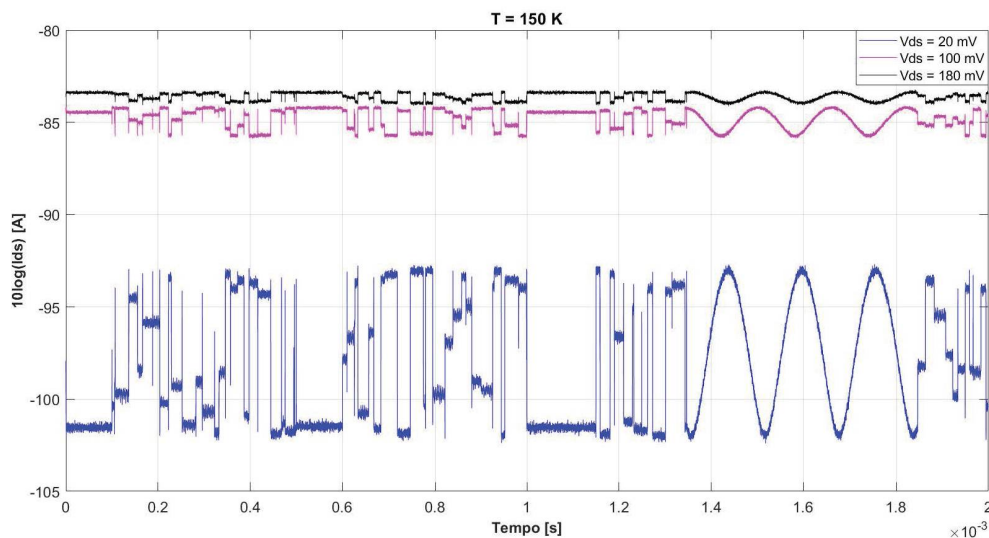


Figura 48: Corrente entre dreno e fonte,  $I_{DS}$ , para os diferentes valores de tensão de polarização.

<sup>1</sup>Neste trabalho, faz-se distinção entre os termos ruído, resíduo e inovação. O ruído refere-se à contaminação da saída do sistema. Ou seja, é um sinal que corrompe a medida de  $y(k)$ .

Tabela 7: Configurações utilizadas para a coleta de dados na construção de modelos em diferentes temperaturas de operação e com tensões de polarização variadas.

Parâmetro \ Temperatura	77 K	150 K	298 K
$C_g$	1 aF	1 aF	1 aF
$C_{j1}, C_{j2}$	1 aF	0,1 aF	0,1 aF
$R_{T1}, R_{T2}$	100 k $\Omega$	10 M $\Omega$	100 M $\Omega$
$V_{DS}$	20 mV	20, 100 e 180 mV	20 mV

Nota-se, na Tabela 7, que os parâmetros das junções-túnel foram alterados conforme a temperatura de simulação é aumentada. Esta alteração é fundamental para que o SET apresente bloqueio e oscilações de Coulomb em sua saída. As condições que devem ser satisfeitas para que a operação do transistor ocorra conforme o esperado foram explicitadas na Seção 2.1.4.

### 6.3 CONSIDERAÇÕES FINAIS

Este capítulo apresentou a metodologia específica aplicada na obtenção dos modelos. Primeiramente, foram estabelecidas condições de simulação ideais e, a partir dos dados coletados, foram testadas algumas representações matemáticas não lineares. Os modelos obtidos pertencentes às classes NARX polinomial, Hammerstein, Wiener e Hammerstein-Wiener foram comparados e ranqueados de acordo com os menores MSE e tempo médio de simulação. A segunda etapa abrange a obtenção de modelos em condições não ideais, por exemplo, a temperaturas diferentes de 0 K.

## 7 APRESENTAÇÃO E ANÁLISE DE RESULTADOS

Neste capítulo, os principais resultados serão expostos e analisados. A apresentação destes resultados terá a mesma sequência disposta no Capítulo 6: primeiramente serão mostrados os resultados obtidos nas simulações que auxiliaram na escolha da representação matemática e, posteriormente, serão expostos os resultados relativos aos modelos obtidos para diferentes condições de operação. A análise destes resultados é feita na última seção deste capítulo.

### 7.1 ESCOLHA DA REPRESENTAÇÃO MATEMÁTICA

O primeiro resultado a ser apresentado diz respeito às estruturas dos modelos. Estes modelos foram construídos para o transistor mono-elétron sob as condições de simulação apresentadas na Tabela 3.

Para o modelo NARX polinomial, a combinação do MQO e da taxa de redução de erro apontou como mais representativos, dentro do espaço de regressores candidatos escolhido, os regressores expressos em (7.1). É válido ressaltar que o modelo com o espaço completo de regressores candidatos possui uma ERR de aproximadamente 98 %. Já o modelo reduzido, com apenas 5 regressores, possui ERR de 96 %.

$$\hat{y}(k) = \hat{\theta}_1 u(k) + \hat{\theta}_2 u(k-1) + \hat{\theta}_3 u(k-2)u(k-4) + \hat{\theta}_4 u(k-3)u(k-3) + \hat{\theta}_5 \hat{y}(k-1)\hat{y}(k-1), \quad (7.1)$$

em que  $u(k) \in \mathbb{R}$  é o sinal de entrada aplicado na etapa de validação,  $\hat{y} \in \mathbb{R}$  é a saída estimada por meio dos regressores expressos na equação e  $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4$  e  $\hat{\theta}_5$  são os parâmetros a serem determinados.

Considerando a escolha do espaço de regressores candidatos, explicitada na Seção 6.1.2, e a construção da matriz de regressores  $\Psi_{NARX}$ , expressa em (6.6), sabe-se que o espaço de regressores é composto por 44 regressores (número de colunas da matriz  $\Psi_{NARX}$ ). Os cinco regressores mais representativos, expressos em (7.1), foram escolhidos para compor o modelo NARX.

O modelo NARX polinomial completo, após a aplicação do MMQ para determinação dos parâmetros, é dado por (7.2). A Figura 50 apresenta a dinâmica do modelo NARX sobreposta com a dinâmica do sistema.

$$\hat{y}(k) = -1,18 \times 10^{-8}u(k) + 1,32 \times 10^{-7}u(k-1) - 3,61 \times 10^{-7}u(k-2)u(k-4) - 2,59 \times 10^{-7}u(k-3)u(k-3) + 4,01 \times 10^{-7}\hat{y}(k-1)\hat{y}(k-1). \quad (7.2)$$

A seguir, serão apresentados os resultados para os modelos orientados a blocos. Para estes modelos, a Tabela 8 apresenta os detalhes da estrutura. A Equação (7.3) apresenta a forma da função de transferência utilizada como bloco dinâmico linear destes modelos.

Tabela 8: Estrutura dos modelos orientados a blocos para o transistor simétrico com  $C_g = C_j = 1$  aF,  $R_T = 100$  k $\Omega$ ,  $V_{DS} = 20$  mV e  $T = 0$  K.  $L_{in}$  e  $L_{out}$  representam o número de segmentos lineares nas não linearidades de entrada e de saída, respectivamente.  $n_a$ ,  $n_b$  e  $n_k$  são a quantidade de zeros, polos e atrasos puros, respectivamente.

Bloco	Hammerstein	Wiener	Hammerstein-Wiener
Não Linearidade de Entrada	$L_{in} = 13$	-	$L_{in} = 11$
Bloco Dinâmico Linear	$n_a = 2$	$n_a = 2$	$n_a = 2$
	$n_b = 4$	$n_b = 4$	$n_b = 4$
	$n_k = 1$	$n_k = 1$	$n_k = 1$
Não Linearidade de Saída	-	$L_{out} = 30$	$L_{out} = 14$

$$H(z) = \frac{I_{DS}(z)}{V_{GS}(z)} = \frac{z^{-1}(1 + \hat{\theta}_1 z^{-1} + \hat{\theta}_2 z^{-2})}{1 + \hat{\theta}_3 z^{-1} + \hat{\theta}_4 z^{-2} + \hat{\theta}_5 z^{-3} + \hat{\theta}_6 z^{-4}}, \quad (7.3)$$

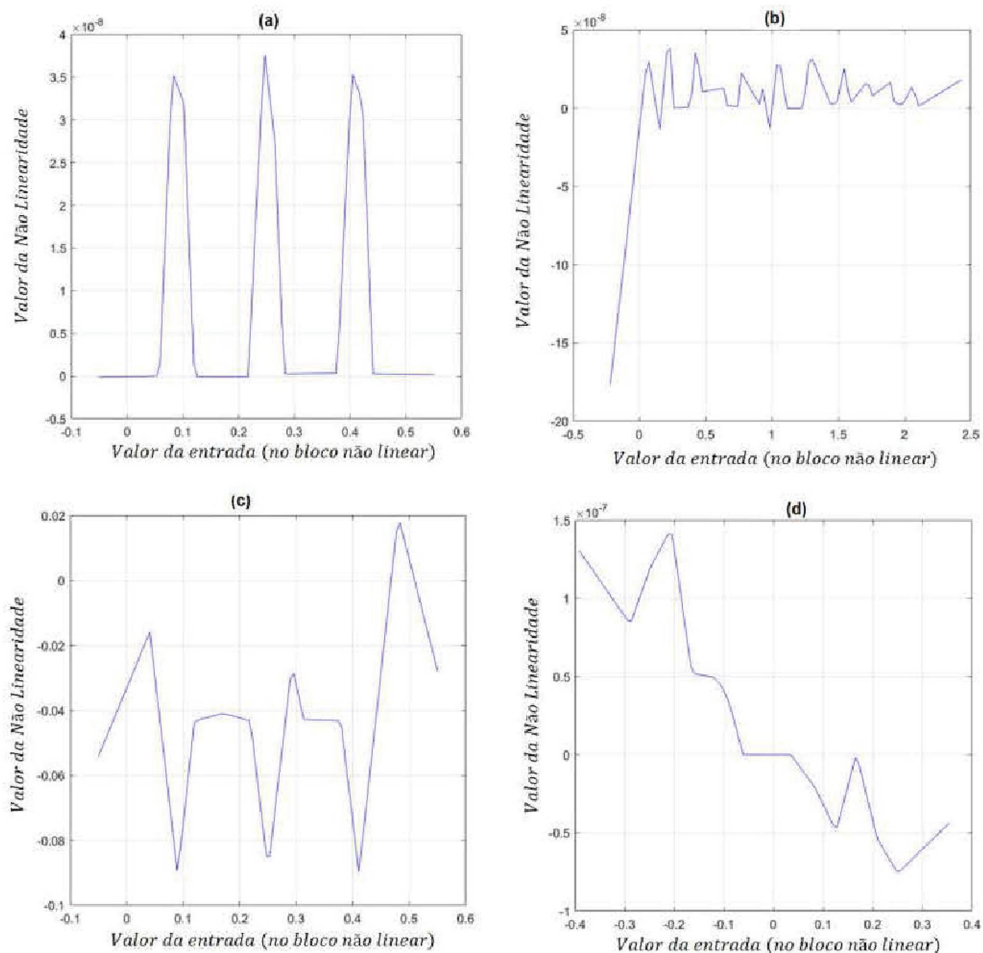
em que  $\hat{\theta}_1$ ,  $\hat{\theta}_2$ ,  $\hat{\theta}_3$ ,  $\hat{\theta}_4$ ,  $\hat{\theta}_5$  e  $\hat{\theta}_6$  são os parâmetros da função de transferência,  $H(z)$ , a serem determinados.

Deve-se ressaltar que, embora a estrutura do bloco dinâmico linear seja a mesma para as três classes de modelos orientados a blocos, os coeficientes da função de transferência variam, conforme pode ser visto na Tabela 9.

Tabela 9: Parâmetros dos blocos dinâmicos lineares para os modelos de Hammerstein, de Wiener e de Hammerstein-Wiener.

Parâmetro	Hammerstein	Wiener	Hammerstein-Wiener
$\hat{\theta}_1$	-0,0507	-0,0548	0,000538
$\hat{\theta}_2$	-0,9485	-0,9419	-0,9991
$\hat{\theta}_3$	-0,3795	0,3556	0,3406
$\hat{\theta}_4$	-0,9006	-0,8130	-0,9536
$\hat{\theta}_5$	0,3117	-0,4177	0,3407
$\hat{\theta}_6$	-0,0311	-0,1244	-0,04618

A Figura 49 apresenta os gráficos das não linearidades estáticas de cada um dos modelos de blocos concatenados.



**Figura 49: Não linearidades estáticas dos modelos de blocos concatenados. a) PWL de entrada para o modelo de Hammerstein. b) PWL de saída para o modelo de Wiener. c) e (d) PWLs de entrada e de saída, para o modelo de Hammerstein-Wiener, respectivamente.**

A Figura 50 apresenta as respostas obtidas com os modelos NARX, de Hammerstein, de Wiener e de Hammerstein-Wiener sobrepostas à resposta do sistema. Para facilitar a visualização, nesta figura são plotadas 20 mil das 50 mil amostras.

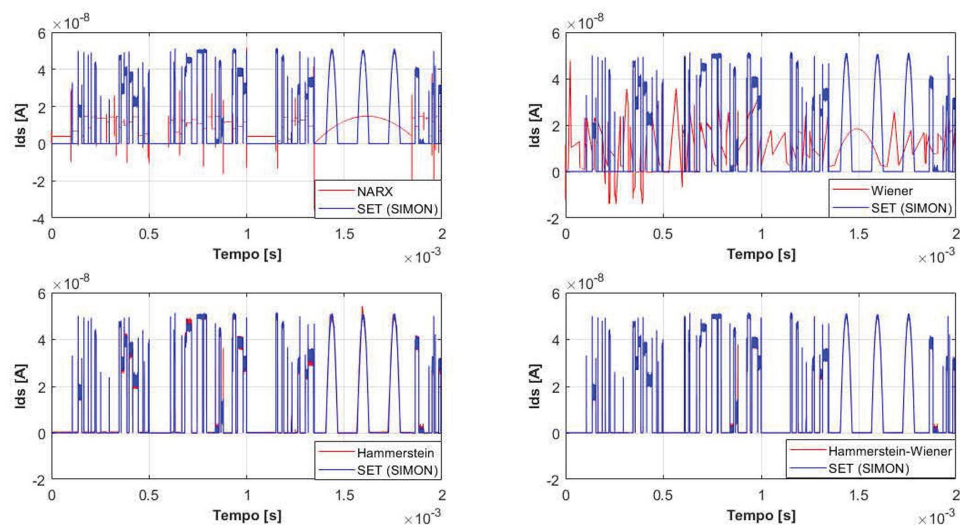


Figura 50: Resposta dos modelos NARX, de Hammerstein, de Wiener e de Hammerstein-Wiener com estruturas explicitadas na Tabela 8 e nas equações (7.2) e (7.3) e parâmetros expostos na Tabela 9. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = C_g = 1$  aF,  $R_T = 100$  k $\Omega$  e  $V_{DS} = 20$  mV. A simulação foi realizada a  $T = 0$  K.

Conforme já exposto, as métricas utilizadas para orientar a escolha da classe de modelos que melhor representaria o SET, dentro do escopo pretendido, foram o MSE e o tempo médio de simulação. A Tabela 10 apresenta os resultados obtidos. Para fins de comparação, nesta tabela foi incluído o tempo que o SIMON levou para executar a simulação do SET na etapa de validação. O gráfico da Figura 51 mostra, de forma resumida, o desempenho de cada uma das classes orientadas a blocos com relação às métricas consideradas. O modelo NARX polinomial não foi incluído neste gráfico por questões de escala. Como pode ser visto na Tabela 10, os valor do MSE para este modelo é consideravelmente maior do que para os demais modelos.

Tabela 10: Desempenho dos modelos com relação ao MSE e ao tempo médio de simulação.

Classe	MSE [ $fA^2$ ]	Tempo Médio de simulação [s]
NARX polinomial	5507	0,0541
Hammerstein	0,0086	0,1019
Wiener	0,3643	0,1020
Hammerstein-Wiener	0,0071	0,1017
SIMON	<i>Benchmark</i>	137 (uma única execução)

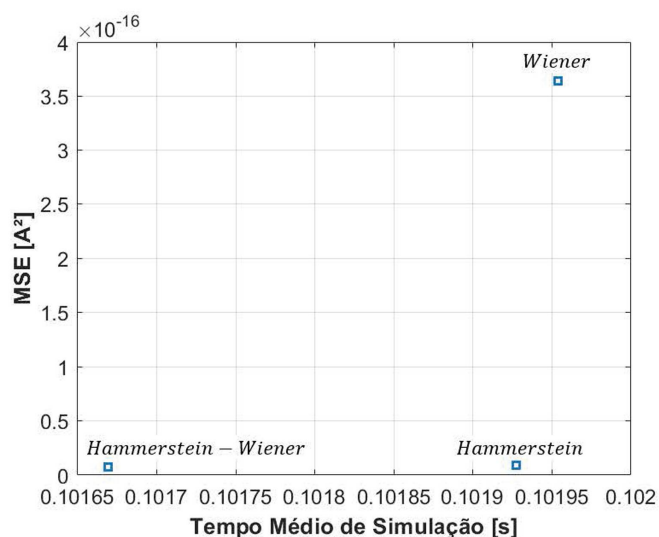


Figura 51: Gráfico tempo médio de simulação  $\times$  MSE para os modelos orientados a blocos.

Com base nos resultados apresentados acima, escolheu-se a classe de modelos de Hammerstein-Wiener para representar o transistor mono-elétron sob as condições simuladas. Na próxima seção, serão apresentados os detalhes dos modelos de Hammerstein-Wiener para outras condições de operação.

## 7.2 MODELOS DE HAMMERSTEIN-WIENER PARA O SET

Nesta seção, os modelos obtidos para o SET para as temperaturas de 77 K, 150 K e 298 K serão detalhados. A determinação de estrutura, estimação de

parâmetros e validação destes modelos seguiu os passos já descritos no Capítulo 6. Os detalhes das estruturas destes modelos encontram-se na Tabela 11 e nas equações (7.4)-(7.6), que apresentam a estrutura dos blocos dinâmicos lineares.

Tabela 11: Estrutura dos modelos de Hammerstein-Wiener para transistor com parâmetros expostos na Tabela 7.  $L_{in}$  e  $L_{out}$  representam o número de segmentos lineares nas não linearidades de entrada e de saída, respectivamente.  $n_a$ ,  $n_b$  e  $n_k$  são a quantidade de polos, zeros e atrasos puros.

T [K]	77	150	150	150	298
$V_{DS}$ [mV]	20	20	100	180	20
$L_{in}$	10	20	19	33	19
$n_a$	1	1	1	1	2
$n_b$	3	4	4	4	4
$n_k$	1	0	0	0	0
$L_{out}$	10	30	15	34	15

A Equação (7.4) mostra a estrutura da função de transferência,  $H_{77K}(z)$ , que representa o bloco dinâmico linear, para  $T = 77$  K,  $C_j = C_g = 1$  aF,  $R_T = 100$  k $\Omega$  e  $V_{DS} = 20$  mV.

$$H_{77K}(z) = \frac{I_{DS}(z)}{V_{GS}(z)} = \frac{z^{-1}(1 + \hat{\theta}_1 z^{-1})}{1 + \hat{\theta}_2 z^{-1} + \hat{\theta}_3 z^{-2} + \hat{\theta}_4 z^{-3}}. \quad (7.4)$$

Para os modelos em que  $T = 150$  K,  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 10$  M $\Omega$  e  $V_{DS} = 20$  mV, 100 mV e 180 mV, o bloco dinâmico linear representado por  $H_{150K}(z)$  é:

$$H_{150K}(z) = \frac{I_{DS}(z)}{V_{GS}(z)} = \frac{1 - \hat{\theta}_1 z^{-1}}{1 + \hat{\theta}_2 z^{-1} + \hat{\theta}_3 z^{-2} + \hat{\theta}_4 z^{-3} + \hat{\theta}_5 z^{-4}}. \quad (7.5)$$

Finalmente, para o modelo em que  $T = 298$  K,  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 100$  M $\Omega$  e  $V_{DS} = 20$  mV, a função de transferência,  $H_{298K}(z)$ , pode ser expressa por:

$$H_{298K}(z) = \frac{I_{DS}(z)}{V_{GS}(z)} = \frac{1 - \hat{\theta}_1 z^{-1} + \hat{\theta}_2 z^{-2}}{1 + \hat{\theta}_3 z^{-1} + \hat{\theta}_4 z^{-2} + \hat{\theta}_5 z^{-3} + \hat{\theta}_6 z^{-4}}. \quad (7.6)$$

Nas equações (7.4)-(7.6),  $\hat{\theta}_1$ ,  $\hat{\theta}_2$ ,  $\hat{\theta}_3$ ,  $\hat{\theta}_4$ ,  $\hat{\theta}_5$  e  $\hat{\theta}_6$  são os parâmetros das funções de transferência a serem determinados. Estes parâmetros têm valores diferentes para cada um dos modelos, conforme exposto na Tabela 12.

Tabela 12: Parâmetros dos blocos dinâmicos lineares para os modelos de Hammerstein, de Wiener e de Hammerstein-Wiener em diferentes condições de simulação.

T [K]	77	150	150	150	298
$V_{DS}$ [mV]	20	20	100	180	20
$\hat{\theta}_1$	-0,9822	-0,9830	-0,9800	0,9772	-0,6126
$\hat{\theta}_2$	-1,2595	-0,9815	-0,7689	0,9803	0,9308
$\hat{\theta}_3$	0,2976	-0,0251	-0,2313	0,02224	$-2,1714 \times 10^{-10}$
$\hat{\theta}_4$	-0,0248	0,0461	0,2610	0,0086	$1,1721 \times 10^{-10}$
$\hat{\theta}_5$	-	-0,0225	-0,2339	-0,0068	$4,0577 \times 10^{-11}$
$\hat{\theta}_6$	-	-	-	-	$4,8147 \times 10^{-11}$

As Figuras 52-54 ilustram as não linearidades estáticas utilizadas em cada modelo. As figuras 55-59 mostram a resposta do modelo sobreposta à resposta do sistema. A Tabela 13 apresenta o MSE de cada modelo com relação à saída do transistor mono-elétron coletada através de simulação no SIMON.

Tabela 13: MSE obtido pelos modelos de Hammerstein-Wiener em diferentes condições de simulação.

T [K]	77	150	150	150	298
$V_{DS}$ [mV]	20	20	100	180	20
MSE [ $A^2$ ]	$3,47 \times 10^{-18}$	$6,79 \times 10^{-23}$	$1,81 \times 10^{-21}$	$2,14 \times 10^{-21}$	$1,82 \times 10^{-24}$

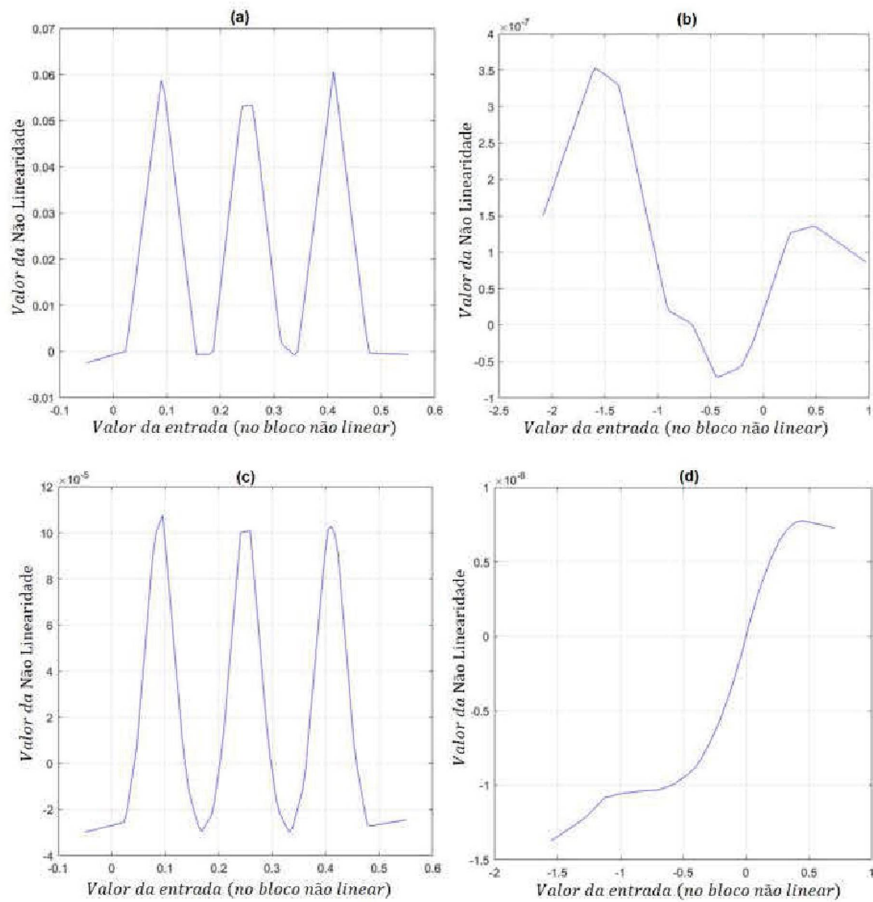


Figura 52: Não linearidades estáticas para os modelos de Hammerstein-Wiener em diferentes condições de operação. (a) PWL de entrada e (b) de saída para  $T = 77$  K e  $V_{DS} = 20$  mV. (c) PWL de entrada e (d) de saída para  $T = 150$  K e  $V_{DS} = 20$  mV.

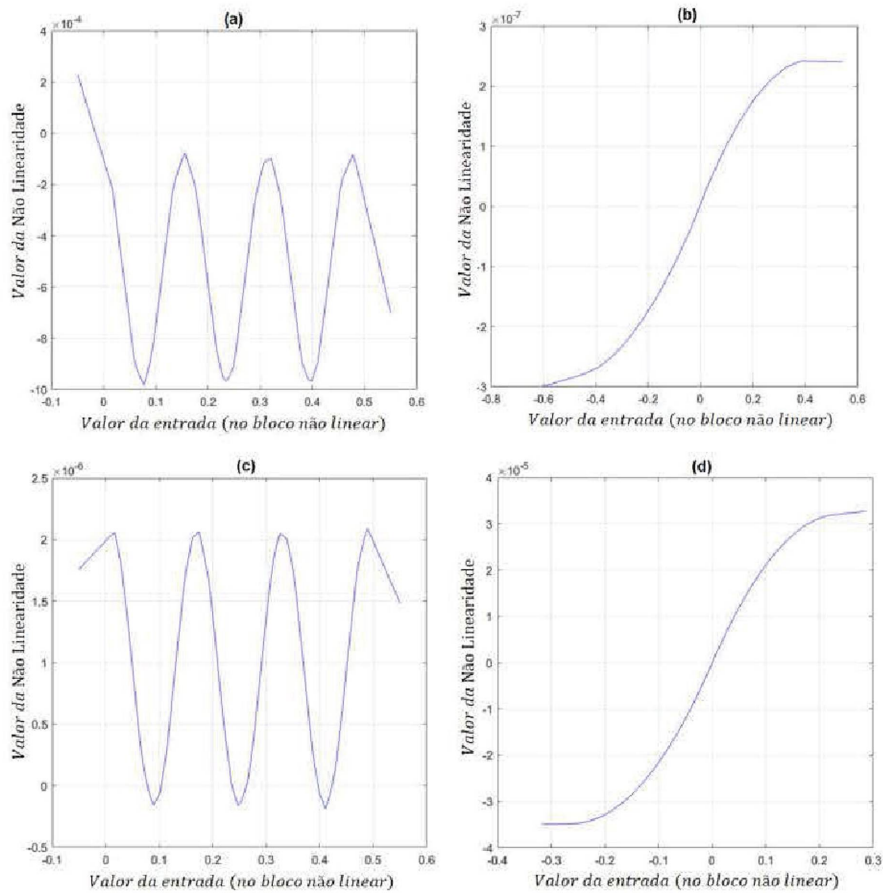


Figura 53: Não linearidades estáticas para os modelos de Hammerstein-Wiener em diferentes condições de operação. (a) PWL de entrada e (b) de saída para  $T = 150\text{ K}$  e  $V_{DS} = 100\text{ mV}$ . (c) PWL de entrada e (d) de saída para  $T = 150\text{ K}$  e  $V_{DS} = 180\text{ mV}$ .

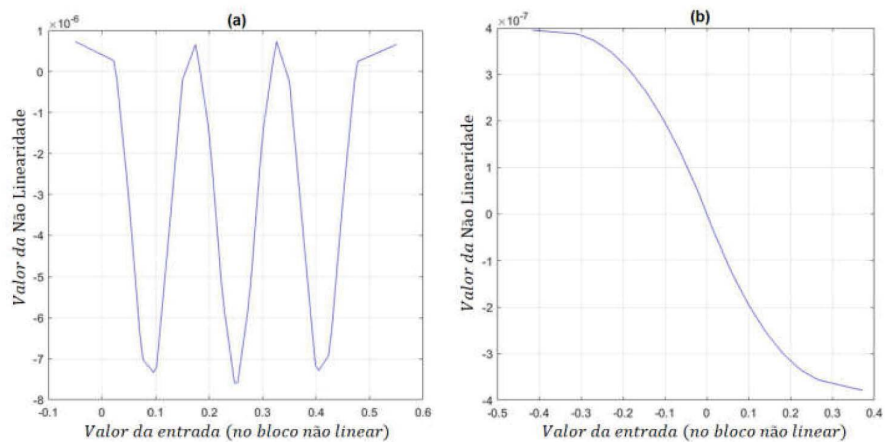


Figura 54: Não linearidades estáticas para os modelos de Hammerstein-Wiener em diferentes condições de operação. (a) PWL de entrada e (b) de saída para  $T = 298$  K e  $V_{DS} = 20$  mV

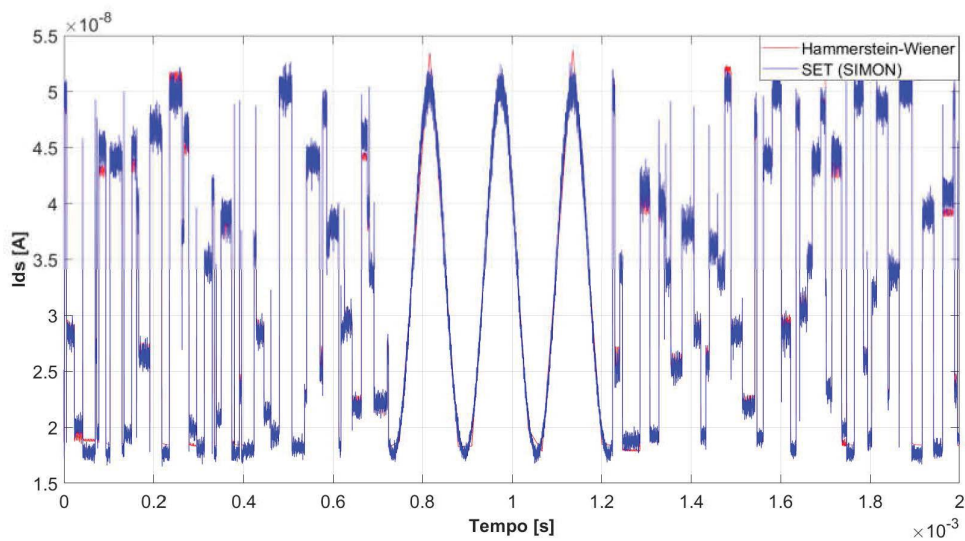


Figura 55: Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.4) e parâmetros expostos na Tabela 12. O transistor monoelétrón é simétrico e possui os seguintes parâmetros:  $C_j = C_g = 1$  aF,  $R_T = 100$  k $\Omega$  e  $V_{DS} = 20$  mV. A simulação foi realizada a  $T = 77$  K. O MSE é de  $3,47 \times 10^{-18}$  A<sup>2</sup>.

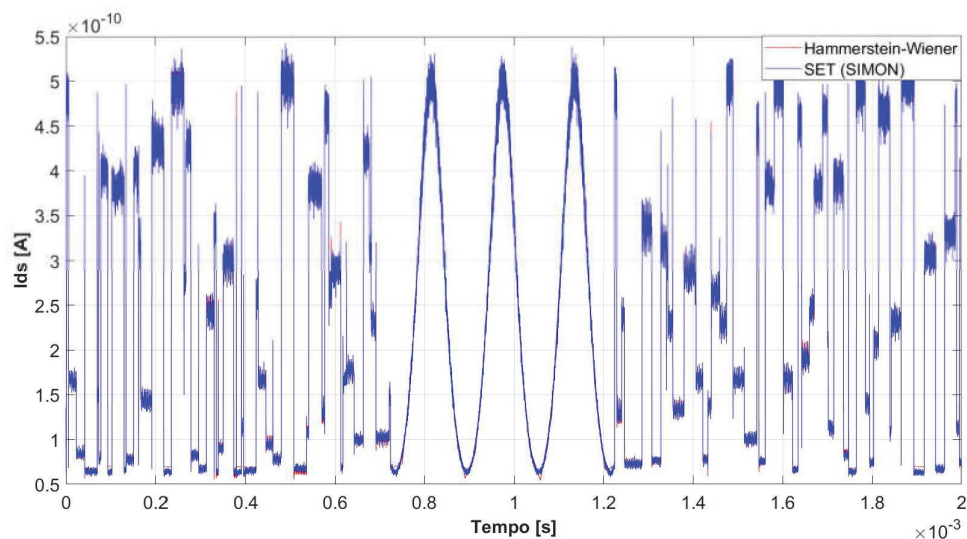


Figura 56: Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 10$  M $\Omega$  e  $V_{DS} = 20$  mV. A simulação foi realizada a  $T = 150$  K. O MSE é de  $6,79 \times 10^{-23}$  A<sup>2</sup>.

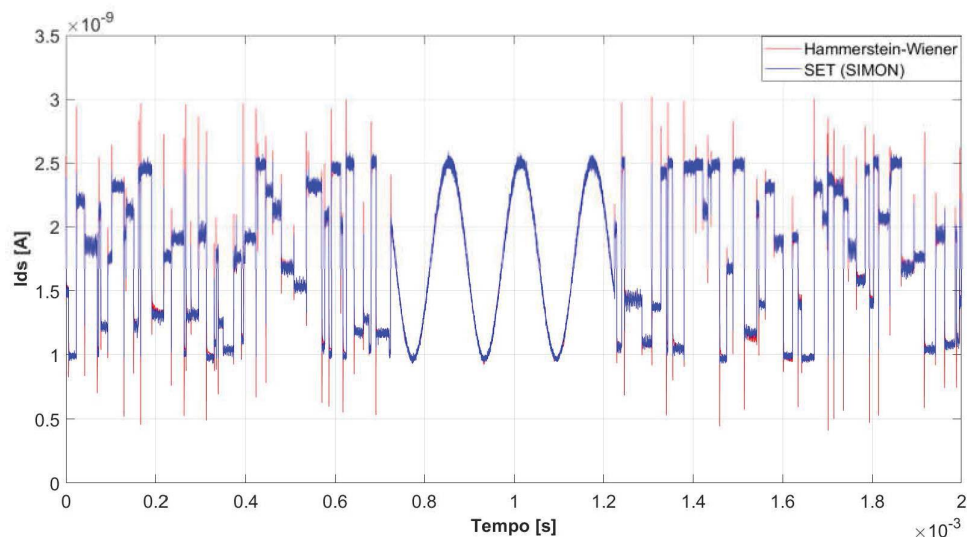


Figura 57: Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 10$  M $\Omega$  e  $V_{DS} = 100$  mV. A simulação foi realizada a  $T = 150$  K. O MSE é de  $1,81 \times 10^{-21}$  A<sup>2</sup>.

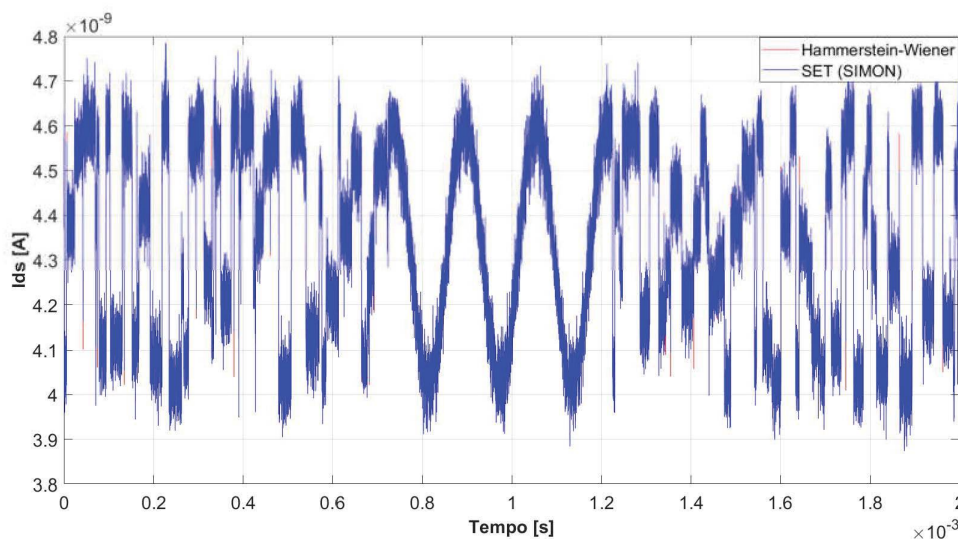


Figura 58: Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 10$  M $\Omega$  e  $V_{DS} = 180$  mV. A simulação foi realizada a  $T = 150$  K. O MSE é de  $2,14 \times 10^{-21}$  A<sup>2</sup>.

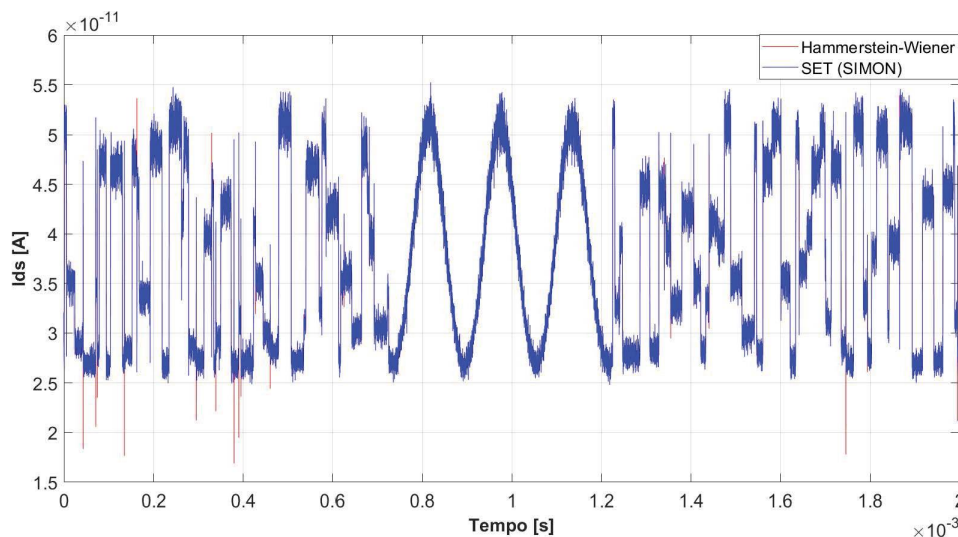


Figura 59: Resposta do modelo de Hammerstein-Wiener com estrutura explicitada na Tabela 11 e na Equação (7.5) e parâmetros expostos na Tabela 12. O transistor mono-elétron é simétrico e possui os seguintes parâmetros:  $C_j = 0,1$  aF,  $C_g = 1$  aF,  $R_T = 100$  M $\Omega$  e  $V_{DS} = 20$  mV. A simulação foi realizada a  $T = 298$  K. O MSE é de  $1,82 \times 10^{-24}$  A<sup>2</sup>.

### 7.3 ANÁLISE DE RESULTADOS

Nesta seção, a análise relativa aos resultados apresentados é realizada. Buscou-se, na argumentação aqui apresentada, seguir a ordem na qual os resultados foram expostos.

Desta forma, a discussão de resultados começa pela avaliação do modelo NARX polinomial. A escolha da representação polinomial, dentro da classe NARX, deve-se, principalmente, ao escopo de aplicação do modelo pretendido. Uma vez que o objeto de modelagem é um dispositivo eletrônico, o modelo matemático deve possuir estrutura simples e de fácil implementação em plataformas de simulação de circuitos. Como pode ser verificado em (7.2), a equação que representa o modelo NARX satisfaz o requisito de simplicidade. Destaca-se, aqui, que a redução do número de regressores, obtida por meio da ERR, teve pouco impacto na representatividade do modelo. Entretanto, a partir da Figura 50, fica claro que este modelo não é adequado para representar o SET.

Conforme já discutido no decorrer deste trabalho, um modelo para o transistor mono-elétron deve ser capaz de reproduzir as características referentes ao bloqueio e às oscilações de Coulomb. Pode-se verificar, analisando a Figura 50, que o modelo NARX polinomial obtido não é capaz de seguir a dinâmica do sistema nas regiões de bloqueio e de oscilação. A baixa representatividade deste modelo para o SET não é evidenciada apenas de maneira visual, mas também pelo alto valor do MSE, que é de  $5,507 \times 10^{-12} A^2$ .

É importante ressaltar, no entanto, que este resultado refere-se apenas a modelos NARX polinomiais obtidos através da metodologia testada neste trabalho. Ou seja, modelos construídos com o uso do MQO e da ERR com as seguintes restrições: máxima regressão de entrada e de saída igual a quatro e ordem de *kernel* máxima igual a dois. Existem diversas metodologias que podem ser aplicadas na obtenção de estrutura de modelos NARX. Não há, entretanto, um procedimento que seja mais recomendado do que outro ou que obtenha desempenho satisfatório para quaisquer representações matemáticas. Neste trabalho, a facilidade de determinar a estrutura compõe o conjunto de critérios considerado na escolha da representação matemática do modelo.

Por fim, ainda sobre o modelo NARX polinomial, a Tabela 10 mostra que o tempo médio de simulação para este modelo é mais baixo do que para os modelos

orientados a blocos. Este resultado era esperado, uma vez que a estrutura do modelo NARX polinomial obtido é bastante simples.

O desempenho dos modelos de blocos concatenados será analisado em conjunto. Primeiramente, serão analisados os resultados obtidos para simulações com  $T = 0$  K e  $V_{DS} = 20$  mV. A configuração das estruturas destes modelos é apresentada na Tabela 8. Percebe-se, por meio dos dados desta tabela, que a estrutura dos modelos de Hammerstein, de Wiener e de Hammerstein-Wiener é semelhante. O bloco dinâmico linear possui a mesma forma para as três classes de modelos. Ressalta-se, entretanto, que os parâmetros deste bloco são diferentes para cada modelo, como pode ser visto na Tabela 9. Estes parâmetros determinam os valores dos polos e zeros do modelo e, portanto, estabelecem características importantes como a estabilidade e a fase do modelo. Sendo assim, não se pode dizer que os modelos de blocos concatenados apresentados na Seção 7.1 possuem a mesma função de transferência.

Quanto ao processo de obtenção da estrutura dos modelos orientados a blocos, pode-se concluir que existe bastante subjetividade envolvida. A falta de uma metodologia bem estabelecida que auxilie na obtenção da estrutura de modelos orientados a blocos deve-se, em grande parte, à presença das não linearidades estáticas. Esta subjetividade, entretanto, é parcialmente compensada pela existência de ferramentas que auxiliam na obtenção empírica da estrutura. Estas ferramentas buscam facilitar a execução de testes que ajudem o projetista a chegar a uma estrutura cuja adequação à dinâmica do sistema seja satisfatória. Como exemplo, pode-se citar a *toolbox* Ident do MatLab.

Ainda sobre a subjetividade na obtenção da estrutura dos modelos, pontua-se que esta não é uma característica inerente apenas aos modelos de blocos concatenados. Na metodologia aplicada à obtenção da estrutura do modelo NARX, esta subjetividade se revela na definição das restrições com as quais o espaço de regressores candidatos é construído.

Destaca-se, ainda, que a estrutura dos modelos orientados a blocos deste trabalho é composta pela concatenação de segmentos de retas e por funções de transferência. Ou seja, são estruturas matemáticas simples, passíveis de serem implementadas em plataformas de simulação de circuitos.

A análise visual das respostas dos modelos de blocos concatenados, ilustradas na Figura 50, mostra que os modelos de Hammerstein e de Hammerstein-Wiener representam, com fidelidade satisfatória, as regiões em que o transistor se encontra

em bloqueio e em condução (característica oscilatória). Já o modelo de Wiener falha na representação dessas duas dinâmicas. O MSE, mais uma vez, corrobora o resultado evidenciado de forma visual. Na Tabela 10, nota-se que o MSE do modelo de Wiener é aproximadamente 423,6 % maior do que o MSE do modelo de Hammerstein e aproximadamente 513,1 % maior do que o MSE do modelo de Hammerstein-Wiener. Na comparação entre os modelos de Hammerstein e de Hammerstein-Wiener, percebe-se que o MSE para o modelo de Hammerstein é 12,1 % maior do que o MSE para o modelo de Hammerstein-Wiener. Este resultado evidencia a importância do uso de métricas que quantifiquem o desempenho dos modelos. A análise puramente visual poderia, neste caso, levar à conclusão errônea de que os modelos de Hammerstein e de Hammerstein-Wiener possuem desempenho equivalente.

Os tempos médios de simulação dos modelos orientados a blocos são, como se esperava, semelhantes, ver Tabela 10. Isto ocorre devido à similaridade da estrutura destes modelos. Os resultados de MSE e tempo médio de simulação, para os modelos de blocos concatenados, são resumidos na Figura 51. Esta figura ajuda a deixar mais clara a escolha pela representação matemática de Hammerstein-Wiener.

Desta forma, pode-se prosseguir à análise dos resultados obtidos com modelos de Hammerstein-Wiener para diferentes configurações de simulação. Conforme pode ser visto na revisão de modelos caixa branca do SET, apresentada na seção 4.2.1, existem dois parâmetros que afetam consideravelmente a operação do transistor: a temperatura,  $T$ , e a tensão de polarização,  $V_{DS}$ . Sendo assim, buscou-se desenvolver modelos para temperaturas e tensões diferentes, a fim de comprovar que a classe de modelos de Hammerstein-Wiener é capaz de representar o SET em diferentes condições de operação.

A primeira temperatura testada é de  $T = 77$  K. A estrutura do modelo construído para esta temperatura está expressa na Tabela 11 e em (7.4). Os parâmetros se encontram na Tabela 12. Para este modelo,  $V_{DS}$  foi mantido em 20 mV. Nota-se, na Tabela 12, que este modelo possui um bloco dinâmico linear com menos parâmetros, ou seja, a representação desta função de transferência é mais simples. A Figura 55 ilustra a resposta do modelo, cujo MSE é de  $3,47 \times 10^{-18} A^2$ . Nesta figura pode-se perceber a presença de ruído. Conforme já dito na Seção 6.2, o SIMON adiciona ruído conforme a temperatura aumenta. Nos Capítulos 2 e 4, foi discutida a influência da temperatura na operação do SET (devido à dependência dos parâmetros da junção-túnel com a temperatura) e na taxa de tunelamento. Apesar da presença de ruído, o

modelo apresentou boa concordância com a curva de saída do sistema, sendo capaz de representar as dinâmicas de bloqueio e oscilações de Coulomb.

Escolheu-se a temperatura intermediária de  $T = 150$  K para testar a representatividade da classe de Hammerstein-Wiener em valores diferentes de  $V_{DS}$ . As estruturas destes modelos estão explicitadas na Tabela 11 e em (7.5). Nota-se que o bloco dinâmico linear manteve a mesma estrutura nos modelos para diferentes valores de  $V_{DS}$ , os parâmetros deste bloco, entretanto, são diferentes para cada modelo e estão expressos na Tabela 12. Na Tabela 11 percebe-se que as não linearidades de entrada e saída possuem quantidades diferentes de segmentos lineares.

As figuras 56, 57 e 58 ilustram a resposta dos modelos de Hammerstein-Wiener a  $T = 150$  K com  $V_{DS}$  de 20, 100 e 180 mV, respectivamente. Aqui, deve-se destacar a presença de ruído na saída do transistor. Além disso, percebe-se a deterioração da relação sinal ruído, conforme o valor da tensão de polarização é aumentado. A amplitude das oscilações sofre com o aumento de  $V_{DS}$ , conforme discutido no Capítulo 4. Neste capítulo, pontuou-se que, para  $|V_{DS}| > \frac{e}{C_{ilha}}$ , o bloqueio de Coulomb não é mais observado. Para a temperatura de 150 K, a capacitância da ilha<sup>1</sup> é dada por  $C_{ilha} = 2C_j + C_g$ . Ou seja  $C_{ilha} = 2 \cdot 0,1 + 1 = 1,2$  aF. Desta forma, para  $V_{DS} > 133$  mV, o funcionamento do dispositivo é prejudicado. O valor de 180 mV para  $V_{DS}$  explica a resposta observada na Figura 58. O modelo, entretanto, obteve um MSE de  $2,14 \times 10^{-21} A^2$ , demonstrando-se capaz de representar a saída coletada no simulador.

Por fim, testou-se a representatividade da classe de Hammerstein-Wiener para um SET simulado a  $T = 298$  K. A estrutura deste modelo está explicitada na Tabela 11 e em (7.6). Os parâmetros se encontram na Tabela 12. A resposta deste modelo é ilustrada na Figura 59 na qual, novamente, destaca-se a presença de ruído. Para este modelo, obteve-se um MSE de  $1,82 \times 10^{-24} A^2$ .

Percebe-se que os modelos de Hammerstein-Wiener apresentaram resultados satisfatórios para todas as condições de simulação testadas. Pode-se dizer que esta classe de representações matemáticas, com não linearidades do tipo PWL, mostrou-se adequada para representar o transistor mono-elétron. Deve-se, porém, destacar que não há um único modelo que represente o transistor em qualquer condição de operação. Os modelos aqui apresentados, como todos os modelos já desenvolvidos para o transistor, possuem escopos de validade bem definidos. Estes escopos con-

<sup>1</sup>Já que o SET é simétrico, a capacitância da junção-túnel é multiplicada por 2.

sideram as condições em que as simulações de coleta de dados foram feitas e os parâmetros do transistor sendo simulado.

No próximo capítulo são apresentadas as considerações finais e os trabalhos futuros que podem ser elaborados a partir do desenvolvimento feito neste trabalho.

## 7.4 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os principais resultados deste trabalho. Os resultados da comparação entre as classes de modelos, com relação ao MSE e ao tempo médio de simulação, foram expostos e a escolha da classe de Hammerstein-Wiener justificada. Posteriormente, mostrou-se os resultados obtidos para diferentes configurações de simulação. Por fim, as principais discussões relativas aos resultados foram feitas.

## 8 CONCLUSÕES

Neste capítulo, serão apresentadas as considerações finais relativas à proposta da tese. Estas discussões consideram o estado da arte atual e os resultados apresentados neste trabalho. Os objetivos serão retomados e as publicações obtidas até o presente momento apresentadas.

### 8.1 Discussões

O objetivo principal desta tese de doutorado foi desenvolver um modelo caixa preta para o transistor mono-elétron. Tal objetivo se justifica devido ao grande potencial de aplicações do SET na indústria eletrônica. Para que este potencial seja efetivamente realizado, modelos precisos e eficientes são necessários.

Sendo assim, foi apresentado um amplo panorama dos modelos de transistor mono-elétron já propostos. Primeiramente, evidenciou-se a originalidade da ideia de utilizar conhecimentos relativos à Identificação de Sistemas na modelagem do SET. Esta abordagem é inédita na literatura.

No que concerne aos modelos já desenvolvidos, estes pertencem a dois grandes grupos: modelos analíticos ou modelos macro para SPICE. Existem diversos pontos, comuns às abordagens analítica e macro, que indicam aprimoramentos necessários, conforme discutido a seguir.

Primeiramente, ressalta-se a falta de utilização de métricas para quantificar a qualidade do modelo. Foram revisados em torno de 40 artigos, entre os mais atuais, relativos à elaboração de modelos para dispositivos mono-elétron, apenas um deles utiliza o MSE como métrica. Os demais trabalhos fazem uma comparação visual entre as curvas utilizadas como *benchmark* e as apresentadas em seus resultados. Esta metodologia apresenta pouca informação ao leitor e impede comparações mais precisas entre os modelos já desenvolvidos. Cabe ressaltar que destas comparações

poderiam resultar novas ideias de modelagem e configurações de parâmetros. Além disto, não existe muita clareza quanto aos parâmetros utilizados nas equações dos modelos. Este fato prejudica a reprodutibilidade das propostas.

Deste modo, a ideia de modelar o transistor por meio da Identificação de Sistemas pode fornecer ganhos, como a simplicidade e maior precisão dos modelos. Além disso, esta abordagem confere rigor à metodologia de modelagem, uma vez que se seguem passos previamente determinados nas etapas de testes e validação dos modelos.

A partir dos resultados obtidos, pode-se concluir que os modelos de Hammerstein-Wiener apresentam o melhor desempenho no problema de identificação do SET. A Tabela 10 e a Figura 51 deixam claro o melhor desempenho do modelo de Hammerstein-Wiener na representação do SET sob as condições estabelecidas no Capítulo 6. Dentre os modelos testados, o modelo de Hammerstein-Wiener possui o menor MSE,  $0,0071 fA^2$ , e o menor tempo médio de simulação. Além disso, a estrutura deste modelo é simples, podendo ser implementada em plataformas de simulação de circuitos. Julgando a simplicidade do modelo e sua capacidade de representar o sistema, além do baixo esforço computacional demandado, escolheu-se o modelo de Hammerstein-Wiener para representar o SET.

Uma vez feita esta escolha, procedeu-se à realização de simulações cujos resultados pudesse comprovar que modelos caixa preta são capazes de representar o comportamento do transistor mono-elétron em diferentes situações. Estes resultados foram exibidos no Capítulo 7 e analisados na Seção 7.3.

A análise gráfica das respostas obtidas, corroborada pelo MSE, permitiu concluir que a hipótese levantada neste trabalho foi confirmada. Ou seja, a modelagem caixa preta do transistor não é apenas possível, mas também é eficiente. Desta forma, considera-se que a principal contribuição deste trabalho não diz respeito ao modelo caixa preta em si, mas sim à aplicação de conhecimentos de Identificação de Sistemas em modelos de dispositivos mono-elétron.

## 8.2 Propostas para Trabalhos Futuros

O desenvolvimento deste trabalho resultou na obtenção de modelos caixa preta para o transistor mono-elétron em diferentes situações de operação. Como

continuação direta deste desenvolvimento, imagina-se a modelagem caixa cinza do transistor. Esta abordagem poderia fazer uso do vasto conhecimento disponível sobre o funcionamento do transistor. Também pode-se supor que a abordagem caixa cinza seja capaz de fornecer resultados promissores no que concerne à predição de comportamento.

A obtenção dos dados aplicados no processo de identificação foi feita por meio de simulação. A plataforma utilizada na simulação dos experimentos de coleta de dados é o SIMON. Embora esta metodologia seja amplamente utilizada na literatura, já que o SIMON é conhecido como a plataforma que fornece resultados mais precisos para estruturas nanoeletrônicas, deve-se considerar a coleta de dados experimentais para a realização da identificação do transistor. A falta de uma técnica de fabricação consolidada, para dispositivos mono-elétron, e o fato de que não há abundância de transistores mono-elétron para a realização destes experimentos dificulta a obtenção destes dados. Mesmo assim, esta é uma oportunidade de melhoria que deve ser considerada em novos trabalhos de modelagem caixas preta ou cinza do SET.

Outro desenvolvimento diretamente conectado a este trabalho concerne a aplicação dos modelos caixa preta em ferramentas de simulação de circuitos. Pode-se utilizar a plataforma de programação disponível em SPICE, por exemplo, para implementar as equações que compõem o modelo do transistor. Uma vez que a modelagem caixa preta mostrou-se eficaz em várias condições de operação, tem-se como perspectiva a simulação de circuitos nanoeletrônicos mais complexos e, até mesmo, híbridos.

Existem, ainda, algumas oportunidades, identificadas durante a revisão bibliográfica, que podem ser exploradas. Dentre estas, ressalta-se o trabalho de Lientschnig *et al.* ao desenvolver a relação de recursão expressa na equação (4.11). Esta relação configura um estimador recursivo não ótimo, utilizado no cálculo das probabilidades de estado. Pode-se aplicar o filtro de Kalman estendido a fim de minimizar a variância do estimador, aproximando-o do seu valor ótimo. Esta também é uma abordagem inédita com potencial para gerar ganho de precisão.

## REFERÊNCIAS

- ABUTALEB, M. M. A New Static Differential Design Style for Hybrid SET-CMOS Logic Circuits. *Journal of Computational Electronics*, v. 14, p. 329–340, 2015.
- AGUIRRE, L. A. A Nonlinear Correlation Function for Selecting the Delay Time in Dynamical Reconstructions. *Physics Letters A*, v. 203, p. 88–94, 1995.
- AGUIRRE, L. A. *Introdução à Identificação de Sistemas - Técnicas Lineares e Não Lineares Aplicadas a Sistemas Reais*. 4<sup>o</sup> ed. revista. ed. [S.I.]: Editora UFMG, 2015.
- AGUIRRE, L. A.; BILLINGS, S. A. Dynamical Effects of Overparametrization in Nonlinear Models. *Physica D*, v. 80, p. 26–40, 1995.
- AGUIRRE, L. A.; COELHO, M. C. S.; CORRÊA, M. V. On the Interpretation and Practice of Dynamical Differences between Hammerstein and Wiener Models. *IEEE Proc. Control Theory Appl.*, v. 152, p. 349–356, 2005.
- ALLEC, N.; KNOBEL, R. G.; SHANG, L. SEMSIM: Adaptive Multiscale Simulation for Single-Electron Devices. *IEEE Trans. Nanotechnol.*, v. 7-3, p. 351–354, 2008.
- ALSMADI, A. Cumulant-Based Approach to FIR System Identification. *Journal of Circuit Theory and Applications*, v. 31, p. 625–636, 2003.
- ALSMADI, A.; WILKES, D. M. Robust and Accurate ARX and ARMA Model Order Estimation of Non-Gaussian Processes. *IEEE Transactions on Signal Processing*, v. 50, p. 759763, 2002.
- ALVES, M. A. *Utilização de Auto-Consistência como Ferramenta Auxiliar na Seleção de Estrutura de Modelos NARX Polinomiais*. Tese (Doutorado) — Universidade Federal de Minas Gerais - UFMG, 2009.
- ANDERSON, H. L. Metropolis, Monte Carlo and the MANIAC. *Los Alamos Science*, v. 14, p. 96–108, 1986.
- ASTRÖM, K. J. Maximum Likelihood and Prediction Error Methods. *Automatica*, v. 16, p. 551–574, 1980.
- ASTRÖM, K. J.; HÄGGLUND, T. *Advanced PID Control*. [S.I.]: ISA, 2011.
- AVERIN, D. V.; LIKHAREV, K. K. Single-Electronics: Correlated Transfer of Single Electrons and Cooper Pairs in Small Tunnel Junctions. *Mesoscopic Phenomena in Solids*, p. 173–271, 1991.
- BAI, E. A Blind Approach to the Hammerstein-Wiener Model Identification. *Automatica*, v. 38, p. 967–979, 2002.

- BEENAKKER, C. W. J. Theory of Coulomb-Blockade Oscillations in the Conductance of a Quantum Dot. *Physical Review B*, v. 44, p. 1646–1656, 1991.
- BILLINGS, S. A. *Nonlinear System Identification: NARMAX methods in the time, frequency and spatio-temporal domains*. [S.I.]: Wiley, 2013.
- BILLINGS, S. A.; VOON, W. S. F. Correlation Based Model Validity Tests for Nonlinear Models. *International Journal of Control*, v. 44, p. 235–244, 1986.
- BOHR, M. T. Nanotechnology Goals and Challenges for Electronic Applications. *IEEE Transactions on Nanotechnology*, v. 1, p. 56–62, 2002.
- BOOM, A. J. W. van den; ENDEN, A. W. M. van den. The Determination of the Orders of Process and Noise Dynamics. *Automatica*, v. 10, p. 245–256, 1974.
- BORN, M. Physical Aspects of Quantum Mechanics. *Nature*, v. 119, p. 354–357, 1927.
- BOYD, S.; VANDENBERGHE, L. *Convex Optimization*. 1<sup>o</sup>. ed. [S.I.]: Cambridge Press, 2009.
- CHAOHONG, H. et al. Analog-to-digital Converter based on Single-Electron Tunneling Transistors. *Microelectronics Journal*, v. 37(2), p. 94–97, 2006.
- CHEN, R. H. MOSES: a General Monte Carlo Simulator for Single Electron Circuits. *The Electrochemical Society*, v. 96-2, p. 576, 1996.
- CHIBENI, S. S. Certezas e Incertezas sobre as Relações de Heisenberg. *Revista Brasileira de Ensino de Física*, v. 27, p. 181–192, 2004.
- COELHO, A. A. R. *Identificação de Sistemas Dinâmicos Lineares*. 2<sup>a</sup>. [S.I.]: UFSC, 2016.
- DAVIS, C. H. *A Translation of Gauss Theoria Motus*. [S.I.]: Cambridge, 1963.
- DORF, R.; BISHOP, R. *Sistemas de Controle Moderno*. 8<sup>a</sup>. [S.I.]: LTC, 2001.
- ECKHARDT, R. Stan Ulam, John Von Neumann and the Monte Carlo Method. *Los Alamos Science*, Special Issue 1987, p. 131–143, 1987.
- ELKAIN, G. H. System Identification-based Control of an Unmanned Autonomous Wind-Propelled Catamaran. *Elsevier - Control Engineering Practice*, v. 17(02), p. 158–169, 2009.
- FERRY, D. K. *Transport in Nanostructures*. [S.I.]: Cambridge University Press, 1991.
- FONSECA, L. R. C.; KOROTKOV, A. N.; LIKHAREV, K. A Numerical Study of the Dynamics and Statistics of Single Electron Systems. *Journal of Applied Physics*, v. 78-5, p. 3238–3251, 1995.
- GARNER, C. M. et al. Challenges for Dielectric Materials in Future Integrated Circuit Technology. *Microelectronics Reliability*, v. 47, p. 937–943, 2007.
- GEERLIGS, L. J. et al. Frequency-Locked Turnstile for Single Electrons. *Physical Review Letters*, v. 64-22, p. 2691–2694, 1990.

- GEERLIGS, L. J.; MOOJI, J. E. Charging Effects and Turnstile Clocking of Single Electrons in Small Tunnel Junctions. *Granular Nanoelectronics*, p. 393–411, 1991.
- GEROUSIS, C.; GREPIOTIS, A. Reconfigurable Gate Array Architecture for Logic Functions in Tunneling Transistors Technology. *Microelectronics Journal*, v. 44(8), p. 706–711, 2013.
- GIRI, F.; BAI, E. *Block Oriented Non-linear System Identification*. [S.l.]: Springer, 2010.
- GLAZMAN, L. I.; SHEKHTER, R. I. Coulomb Oscillations of the Conductance in a Laterally Confined Heterostructure. *Journal of Physics: Condensed Matter*, v. 1(33), p. 5811–5815, 1989.
- GOSER, K. F. et al. Aspects of Systems and Circuits for Nanoelectronics. *Proceedings of the IEEE*, v. 85(4), p. 558–573, 1997.
- GRABERT, H.; DEVORET, M. H. *Single Charge Tunneling - Coulomb Blockade Phenomena in Nanostructures*. [S.l.]: Plenum Press, 1992.
- GUIMARÃES, J. G. *Arquiteturas de Redes Neurais Nanoeletrônicas para Processadores em Escala Giga ou Tera*. Tese (Doutorado) — Universidade de Brasília - UnB, 2005.
- HADLEY, P.; LIENTSCHNIG, G.; LAI, M. Single-Electron Transistors. *Proceedings of the 29th International Symposium on Compound Semiconductors*, p. 125–132, 2002.
- HAIN, L. More from Moore. *Nature*, v. 520, p. 408(1), 2015.
- HANSON, G. W. *Fundamentals in Nanoelectronics*. [S.l.]: Prentice Hall, 2008.
- HASANEEN, E. A. M.; WAHAB, M. A. A.; AHMED, M. G. Exact Analytical Model of Single Electron Transistor for Practical IC Design. *Microelectronics Reliability*, v. 51, p. 733–745, 2011.
- HEINZEL, T. *Mesoscopic Electronics in Solid State Nanostructures*. 3ª. [S.l.]: Wiley-VCH, 2007.
- INOKAWA, H.; TAKAHASHI, Y. A Compact Analytical Model for Asymmetric Single-Electron Tunneling Transistors. *IEEE Transactions on Electronic Devices*, v. 50, p. 455–461, 2003.
- ISERMANN, R.; MÜNCHHOF, M. *Identification of Dynamic Systems: An Introduction with Applications*. [S.l.]: Springer-Verlag Berlin Heidelberg, 2011. (Advanced Textbooks in Control and Signal Processing).
- JAIN, A. et al. A new SPICE Macro Model of Single Electron Transistor for Efficient Simulation of Single-Electronic Circuits. *Analog Integrated Circuits and Signal Processing*, v. 82, p. 653–662, 2015.
- JAIN, A.; NAMERIAKPAM, B. S.; SARKAR, S. K. A New Compact Analytical Model of Single Electron Transistor for Hybrid SET-MOS Circuits. *Solid State Electronics*, v. 104, p. 90–95, 2015.

- JAIN, A.; SARKAR, S. K. Design and Reliability Analysis of a 4:1 MUX using Single Electron Tunneling Technology based Threshold Logic Gate. *Journal of Electron Devices*, v. 15, p. 1241–1248, 2004.
- JAIN, A. et al. Stability and Reliability Analysis of Hybrid CMOS-SET Circuits - a New Approach. *Journal of Computational and Theoretical Nanoscience*, v. 11(12), p. 2519–2524, 2014.
- JANA, A. et al. Design and Simulation of Hybrid CMOS-SET Circuits. *Microelectronics Reliability*, v. 53(4), p. 592–599, 2013.
- KARIMIAN, M. et al. A New SPICE Macro-Model for Simulation of Single Electron Circuits. *International Conference on Microelectronics*, p. 228–231, 2009.
- KIRIHARA, M.; NAKAZATO, K.; WAGNER, M. Hybrid Circuit Simulator Including a Model for Single Electron Tunneling Devices. *Japanese Journal of Applied Physics*, v. 38, p. 2028–2032, 1999.
- KULIK, I. O.; SHEKHTER, R. I. Kinetic Phenomena and Charge Discreteness Effects in Granulated Media. *Soviet Physics JETP*, v. 41, p. 308–316, 1975.
- LAVIEVILLE, R. et al. 350 K Operating Silicon Nanowire Single Electron/Hole Transistors Scaled down to 3.4 nm Diameter and 10 nm Gate Length. *EUROSOI-ULIS 2015*, 2015.
- LEMMA, D. T.; RAMASAMY, M.; SHUHAIMI, M. System Identification using Orthonormal Basis Filters. *Journal of Applied Sciences*, v. 10, p. 2516–2522, 2010.
- LIENTSCHNIG, G.; WEYMANN, I.; HADLEY, P. Simulating Hybrid Circuits of Single-Electron Transistors and Field-Effect Transistors. *Japanese Journal of Applied Physics*, v. 42, p. 6467–6472, 2003.
- LIKHAREV, K. K. Single-Electron Devices and Their Applications. *Proceedings of the IEEE*, v. 87, p. 606–632, 1999.
- LIMA, E. L. *Álgebra Linear*. 1°. ed. [S.I.]: Editora do IMPA, RJ, 2008.
- LIMA, E. L. *Espaços Métricos*. 2°. ed. [S.I.]: Editora do IMPA, RJ, 2013.
- LJUNG, L. *System Identification - Theory for the User*. [S.I.]: Prentice Hall - PTR, 1999.
- MAHAPATRA, S.; IONESCU, A. M.; BANERJEE, K. A Quasi-Analytical SET Model for Few Electron Circuit Simulation. *IEEE Electronic Device Letters*, v. 23, p. 366–368, 2002.
- MAHAPATRA, S. et al. Modelling and Analysis of Power Dissipation in Single Electron Logic. *IEDM*, p. 323–326, 2002.
- MAHAPATRA, S. et al. SETMOS: A Novel True Hybrid SET-MOS High Current Coulomb Blockade Oscillation Cell for Future Nanoscale Analog ICs. *IEDM Tech. Dig.*, p. 703–706, 2003.

- MAHAPATRA, S. et al. Analytical Modeling of Single-Electron Transistor for Hybrid CMOS-SET Analog IC Design. *IEEE Transactions on Electronic Devices*, v. 51(11), p. 1772–1782, 2004.
- MARQUARDT, D. W. An Algorithm for Least Squares Estimation of Nonlinear Parameters. *Journal of Society for Industrial and Applied Mathematics*, v. 11(2), p. 431–441, 1963.
- MERZBACHE, E. The Early History of Quantum Tunneling. *Physics Today*, v. 55, p. 44–49, 2002.
- METROPOLIS, N.; ULAM, S. The Monte Carlo Method. *Journal of the American Statistical Association*, v. 44, p. 335–341, 1949.
- MONTEIRO, G. V. S. *Controle Não Linear*. 2ª. [S.I.]: Escola Superior de Tecnologia, 2003.
- NEMETH, J.; KOLLAR, I. Identification of Volterra Kernels using Interpolation. *IEEE Transactions on Instrumentation and Measurements*, v. 51, p. 770–775, 2002.
- NOCEDAL, J.; WRIGHT, S. *Numerical Optimization*. 1ª. ed. [S.I.]: NY Springer, 1999.
- OHATA, A.; TORIUMI, A.; UCHIDA, K. Coulomb Blockade Effects in Edge Quantum Wire SOI MOSFETs. *Japanese Journal of Applied Physics*, v. 36, p. 1686–1689, 1997.
- ONO, Y. et al. Manipulation and Detection of Single Electrons for Future Information Processing. *Journal of Applied Physics*, v. 97(3), p. 1–9, 2005.
- OROSKI, E. *Identificação de Sistemas Não Lineares Utilizando Modelos NARX, Funções Ortogonais e Otimização Heurística*. Tese (Doutorado) — Universidade de Brasília - UnB, 2015.
- PAPOULIS, A.; PILLAI, S. U. *Probability, Random Variables and Stochastic Processes*. 4ª. [S.I.]: McGraw Hill, 2002.
- PARK, H. et al. Nanomechanical Oscillations in a Single-c60 Transistor. *Nature*, v. 407, p. 57–60, 2000.
- PASHKIN, Y. A.; NAKAMURA, Y.; TSAI, J. S. Room-temperature All Single-Electron Transistor made by Electron-Beam Lithography. *Applied Physics Letters*, v. 76, p. 2256–2258, 2000.
- PAUL, D. J. Nanoelectronics. *Encyclopedia of Physical Science and Technology*, v. 10, p. 285–301, 2002.
- PRUVOST, B.; MIZUTA, H.; ODA, S. Voltage-Limitation Free Analytical Single-Electron Transistor Model Incorporating the Effects of Spin-Degenerate Discrete Energy States. *Journal of Applied Physics*, v. 103, p. 1–10, 2008.
- RADWAN, M. S. et al. A Computationally Efficient Model of Single Electron Transistor. *Microelectronics Journal*, v. 46, p. 301–309, 2015.

- RAJA, T.; AGRAWAL, V. D.; BUSHNELL, M. L. A Tutorial on the Emerging Nanotechnology Devices. *Proceedings of the 17th International Conference on VLSI Design (VLSID04)*, p. 343–360, 2004.
- ROSA, A. *Identificação de Sistemas Não Lineares Utilizando Modelos de Volterra baseados em Funções Ortonormais de Kautz e Generalizadas*. Tese (Doutorado) — Uicamp, 2009.
- ROSA, A.; CAMPELLO, R.; AMARAL, W. Exact Search Directions for Optimizations of Linear and Nonlinear Models based on Generalized Orthonormal Functions. *IEEE Trans. on Automatic Control*, v. 54, p. 2757–2772, 2009.
- SHIN, S. J. et al. Si-based Ultrasmall Multiswitching Single-Electron Transistor Operating at Room-temperature. *Applied Physics Letters*, v. 97(10), p. 103101–103103, 2010.
- SHIN, S. J. et al. Room-temperature Charge Stability Modulated by Quantum Effects in a Nanoscale Silicon Island. *Nano Letters*, v. 11(4), p. 1591–1597, 2011.
- STRANG, G. *Álgebra Linear e suas Aplicações*. Tradução da 4<sup>o</sup>. [S.l.]: Cengage Learning, 2009.
- TAO, G.; TIAN, M. Discrete-Time Adaptive Control of Systems with Multisegment Piecewise-Linear Nonlinearities. *IEEE Trans. Automat. Control*, v. 43, p. 719–723, 1998.
- TOUMEY, C. Less is Moore. *Nature Nanotechnology*, v. 11, p. 2–3, 2016.
- UCHIDA, K. et al. Analytical Single-Electron Transistor (SET) Model for Design and Analysis of Realistic (SET) Circuits. *Japanes Journal of Applied Physics*, v. 39, p. 2321–2324, 2000.
- VOLTERRA, V. *Theory of Functionals and of Integral and Integro-Differential Equations*. [S.l.]: Dover Phoenix, 1959.
- VÖRÖS, J. Parameter Identification of Wiener Systems with Multisegment Piecewise-Linear Nonlinearities. *Systems and Control Letters*, v. 56, p. 99–105, 2007.
- WANG, X.; POROD, W. Analytical I-V Model for Single-Electron Transistors. *VLSI Design*, v. 13, p. 189–192, 2001.
- WASSHUBER, C. *Computational Single-Electronics*. 1<sup>a</sup>. [S.l.]: Springer-Verlag Wien, 2001.
- WASSHUBER, C.; KSOINA, H.; SELBERHERR, S. SIMON - A Simulator for Single-Electron Tunnel Devices and Circuits. *IEEE Trans. Comput.-Aided Des.*, v. 16, p. 937–944, 1997.
- WU, Y.; LIN, S. An Improved Single-Electron Transistor Model for SPICE Application. *Nanotechnology*, v. 3, p. 321–324, 2003.
- YU, Y. et al. Implementation of Single Electron Circuit Simulation by SPICE: KOSECSPICE. *Proceedings of Asia Pacific Workshop on Fundamental and Application of Advanced Semiconductor Device*, p. 85–90, 2000.

YU, Y. S.; HWANG, S. W.; AHN, D. Macromodeling of Single-Electron Transistors for Efficient Circuit Simulation. *IEEE Transactions on Electronic Devices*, v. 46, p. 1667–1671, 1999.

YU, Y. S.; HWANG, S. W.; AHN, D. Transient Modelling of Single-Electron Transistors for Efficient Circuit Simulation by SPICE. *IEE Proceedings - Circuits, Devices and Systems*, v. 152, p. 691–696, 2005.

ZARDALIDIS, G.; KARAFYLLIDIS, I. Design and Simulation of a Nanoelectronic Single Electron Control NOT Gate. *Microelectronics Journal*, v. 37(2), p. 94–97, 2006.

ZARDALIDIS, G.; KARAFYLLIDIS, I. G. SECS: a New Single-Electron-Circuit Simulator. *IEEE Trans. Circuits Syst.*, v. 55-9, p. 2774–2784, 2008.

## ANEXO A – Equação Mestra

### A.1 Método de Hadley e Lientschnig (WASSHUBER, 2001)

Considere um sistema com  $N$  condutores, a carga em um condutor  $i$  pode ser descrita através da matriz de capacitâncias do circuito,  $C$ , conforme (A.1).

$$q_i = \sum_{j=1}^N C_{ij} v_j, \quad i = 1, 2, \dots, N. \quad (\text{A.1})$$

Os elementos  $C_{ii}$ , pertencentes à diagonal de  $C$ , correspondem à capacitância total do condutor  $i$ . Os elementos fora da diagonal,  $C_{ij}$ , representam a capacitância entre os condutores  $i$  e  $j$ . A tensão nodal, portanto, pode ser escrita conforme (A.2).

$$v_i = \sum_{j=1}^N C_{ij}^{-1} q_j, \quad i = 1, 2, \dots, N. \quad (\text{A.2})$$

A energia eletrostática do circuito, então, é escrita como se segue:

$$U = \frac{1}{2} \mathbf{q}^T \mathbf{v} = \frac{1}{2} \sum_{i=1}^N q_i v_i = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N C_{ij}^{-1} q_i q_j. \quad (\text{A.3})$$

Na Equação (A.3),  $\mathbf{q} \in \mathbb{R}$  é o vetor de cargas nodais e  $\mathbf{v} \in \mathbb{R}$  é o vetor de tensões nodais. Desta forma, as cargas e tensões nodais podem ser dadas pelas derivadas parciais da energia eletrostática:

$$q_i = \frac{\partial U}{\partial v_i} \quad v_i = \frac{\partial U}{\partial q_i}. \quad (\text{A.4})$$

Sendo assim, a variação total da energia eletrostática pode ser escrita em função de suas derivadas parciais. Escolhendo as cargas nodais como variáveis independentes, que descrevem o estado do circuito, a variação total em  $U$  é:

$$dU(\mathbf{q}) = \sum_{i=1}^N \frac{\partial U}{\partial q_i} dq_i, \quad (\text{A.5})$$

ou

$$dU(\mathbf{q}) = \sum_{i=1}^N v_i dq_i. \quad (\text{A.6})$$

Assim, a variação  $dU$  causada pela adição de uma carga infinitesimal,  $dq_j$ , ao nó  $j$  é dada por:

$$dU = v_j dq_j = \sum_{i=1}^N C_{ji}^{-1} q_i dq_j, \quad (\text{A.7})$$

em que foi utilizada (A.2).

Para determinar a variação causada pela adição de uma carga elementar,  $e$ , ao nó  $j$ , basta integrar (A.7) de 0 até  $e$ . O resultado está expresso em (A.8), em que se considera a energia de carregamento como uma parcela da variação total da energia eletrostática.

$$\Delta U = ev_j + C_{jj}^{-1} \frac{e^2}{2}. \quad (\text{A.8})$$

Considere que um elétron tunele de um nó inicial,  $i$ , para um nó final,  $f$ . As tensões iniciais são  $v_i$ , para o nó  $i$ , e  $v_f$ , para o nó  $f$ . A variação na energia eletrostática é, então, calculada em duas etapas: retirar um elétron do nó  $i$  e adicionar um elétron ao nó  $f$ . Remover um elétron do nó  $i$  é equivalente a adicionar uma carga  $e$  ao nó  $i$ . Assim, a variação da energia eletrostática na primeira etapa será:

$$\Delta U_1 = ev_i + C_{ii}^{-1} \frac{e^2}{2}. \quad (\text{A.9})$$

Após a primeira etapa, o estado do circuito foi modificado. Isto ocorre, pois a remoção da carga  $-e$  do nó  $i$  altera as tensões nodais do circuito. Desta forma, considere que  $v'_f$  e  $v'_i$  são as tensões nodais dos nós envolvidos no tunelamento após a primeira etapa.

$$v'_i = v_i + eC_{ii}^{-1} \quad v'_f = v_f + eC_{if}^{-1}. \quad (\text{A.10})$$

A variação de energia causada pela adição de um elétron ao nó  $f$ , segunda etapa, será:

$$\Delta U_2 = -ev'_f + C_{ff}^{-1} \frac{e^2}{2}. \quad (\text{A.11})$$

Substituindo (A.10) para  $v'_f$  em (A.11), tem-se:

$$\Delta U_2 = -ev_f + (C_{ff}^{-1} - 2C_{if}^{-1}) \frac{e^2}{2}. \quad (\text{A.12})$$

A variação de energia total, causada pelo tunelamento, é dada pela soma das variações em cada etapa:

$$\Delta U = \Delta U_1 + \Delta U_2 = -e(v_f - v_i) + (C_{ii}^{-1} - 2C_{if}^{-1} + C_{ff}^{-1}) \frac{e^2}{2}, \quad (\text{A.13})$$

$\Delta U$  em (A.13) é igual a  $\Delta F$  em (4.1).

Uma vez que o sistema tunela a partir de estados de energia mais alta para estados de energia mais baixa, o tunelamento só ocorrerá se  $\Delta U < 0$ . A condição em que  $\Delta U = 0$  é utilizada para calcular as regiões estáveis do sistema (HADLEY; LIENTSCHNIG; LAI, 2002).

A variação de energia eletrostática é função apenas da diferença entre as tensões nodais iniciais (para os nós envolvidos no tunelamento) e de um termo que envolve as capacitâncias intrínseca e relativa destes nós. Esta variação, portanto, não depende do estado de carga do sistema. Assim, não é necessário calcular o estado de carga do sistema para cada evento de tunelamento, o que impacta diretamente na eficiência da simulação.

## A.2 Cálculo das Probabilidades de Ocupação dos Estados

Considere a forma matricial da Equação Mestra do tunelamento mono-elétron:

$$\begin{bmatrix} \frac{\partial P_1(t)}{\partial t} \\ \frac{\partial P_2(t)}{\partial t} \\ \vdots \\ \frac{\partial P_n(t)}{\partial t} \end{bmatrix} = \begin{bmatrix} -\sum_{i \neq 1} \Gamma_{i1} & \Gamma_{12} & \cdots & \Gamma_{1n} \\ \Gamma_{21} & -\sum_{i \neq 2} \Gamma_{i2} & \cdots & \Gamma_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \Gamma_{n1} & \Gamma_{n2} & \cdots & -\sum_i \Gamma_{in} \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ \vdots \\ P_n(t) \end{bmatrix}, \quad (\text{A.14})$$

$$\dot{\mathbf{p}} = \Gamma \mathbf{p}, \quad (\text{A.15})$$

em que  $\Gamma \in \mathbb{R}^{n \times n}$  é a matriz de transição, cujos elementos são as taxas de transição entre os estados.  $n$  é o número de estados considerados e o termo  $\mathbf{p}$  é o vetor de probabilidades de estado:

$$\mathbf{p} = [P_1(t), P_2(t), \dots, P_n(t)]^T. \quad (\text{A.16})$$

Para o caso estacionário,  $\dot{\mathbf{p}}$  na equação (A.15) é igual a zero. Assim,  $0 = \Gamma \mathbf{p}$ , é um sistema de equações lineares, cuja solução é possível através de uma variedade de algoritmos numéricos. O caso transiente, no entanto, resulta em um sistema de equações diferenciais ordinárias homogêneas de primeira ordem. Este sistema pode ser resolvido através da integração formal da equação (A.15), que implicará no cálculo da exponencial da matriz  $\Gamma$ , conforme (A.17) (WASSHUBER, 2001).

$$\mathbf{p}(t) = e^{\Gamma t} \mathbf{p}(0), \quad (\text{A.17})$$

na qual a exponencial matricial é dada pela série de Taylor:

$$e^{\Gamma t} = \sum_{k=0}^{\infty} \frac{(\Gamma t)^k}{k!}, \quad (\text{A.18})$$

em que  $k \in \mathbb{N}$ .