UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA

CARLOS AUGUSTO BERGFELD MORI

Estudo comparativo do efeito de autoaquecimento em
transistores FinFET e SOI UTBB

São Paulo
2018

CARLOS AUGUSTO BERGFELD MORI

Estudo comparativo do efeito de autoaquecimento em
transistores FinFET e SOI UTBB

Dissertação apresentada à Escola
Politécnica da Universidade de São
Paulo para obtenção do título de Mestre em
Ciências.

Área de concentração: Microeletrônica.

Orientador: Prof. Dr. João Antonio Martino
Coorientadora: Prof^a. Dra. Paula Ghedini Der
Agopian

São Paulo

2018

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Mori, Carlos Augusto Bergfeld

Estudo comparativo do efeito de autoaquecimento em transistores FinFET e SOI UTBB / C. A. B. Mori -- versão corr. -- São Paulo, 2018.
87 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos.

1.Microeletrônica 2.Transistores 3.Autoaquecimento I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Sistemas Eletrônicos II.t.

Nome: MORI, Carlos Augusto Bergfeld

Título: Estudo comparativo do efeito de autoaquecimento em transistores FinFET e SOI UTBB

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Aprovado em: 09 / 02 / 2018

Banca Examinadora

Prof. Dr. João Antonio Martino

Instituição: Escola Politécnica da Universidade de São Paulo

Julgamento: _____

Prof. Dr. Victor Sonnenberg

Instituição: Faculdade de Tecnologia de São Paulo

Julgamento: _____

Prof. Dr. Gilson Wirth

Instituição: Universidade Federal do Rio Grande do Sul

Julgamento: _____

AGRADECIMENTOS

À Prof.^a Dra. Paula, pelas contribuições à estruturação, elaboração e compreensão do trabalho.

Ao Prof. Dr. Martino, pelo auxílio e orientação ao longo do desenvolvimento dos estudos.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pela concessão da bolsa de mestrado sob o processo 2015 e pelo apoio financeiro para a execução desta pesquisa.

À Escola Politécnica da Universidade de São Paulo, pela disponibilização da sua infraestrutura e laboratórios para pesquisa.

RESUMO

MORI, Carlos Augusto Bergfeld. **Estudo comparativo do efeito de autoaquecimento em transistores FinFET e SOI UTBB**. 2017. 88. Dissertação (Mestrado) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2017.

Devido às dimensões cada vez mais reduzidas dos transistores e a utilização de novos materiais com baixa condutividade térmica, o desempenho de transistores avançados é afetado pelo autoaquecimento. Dispositivos sob os efeitos de autoaquecimento sofrem um aumento da sua temperatura, fazendo com que a mobilidade seja reduzida, além de comprometer a confiabilidade e gerar atrasos de sinal, trazendo impactos na eficiência de circuitos analógicos, bem como afetando o desempenho de circuitos digitais. Apesar da relevância do fenômeno, muitos estudos não o levam em consideração devido à dificuldade de sua verificação, uma vez que os métodos utilizados para transistores avançados requerem estruturas ou equipamentos especiais, que são raramente disponíveis. Dessa forma, três novas técnicas são desenvolvidas neste trabalho com o objetivo de viabilizar o estudo do efeito utilizando estruturas convencionais e medidas em corrente contínua: (i) a condutância de saída média; (ii) o método da assinatura na eficiência do transistor; (iii) a estimativa da resistência térmica utilizando somente medidas em corrente contínua. Os dois primeiros métodos são focados em uma análise qualitativa do autoaquecimento, permitindo uma verificação preliminar eficiente da presença e relevância do efeito, enquanto o terceiro método permite a extração da resistência térmica a partir do inverso da eficiência do transistor utilizando um processo iterativo, consequentemente possibilitando a obtenção do aumento da temperatura do canal devido ao autoaquecimento, com boa precisão e maior simplicidade em relação aos métodos disponíveis na literatura (com erro máximo menor que 6% para transistores de múltiplas portas em relação ao método de medidas pulsadas). Com essas técnicas, são feitas comparações da elevação de temperatura do canal entre transistores de múltiplas portas (também chamados de FinFET ou transistores 3D) e transistores de silício sobre isolante com camada de silício e óxido enterrado extremamente finos (SOI UTBB), usando simulações tridimensionais para obter condições similares de potência. Em dispositivos com menores comprimentos de canal, os FinFETs apresentaram temperaturas cerca de 60 K acima dos UTBBs.

Palavras-chave: autoaquecimento, transistores, SOI, UTBB, FinFET.

ABSTRACT

MORI, Carlos Augusto Bergfeld. **Comparative study of the self-heating effect in FinFET and SOI UTBB transistors**. 2017. 88. Dissertação (Mestrado) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2017.

Due to the reduction of devices' dimensions and the use of new materials with low thermal conductivity, self-heating affects the performances of advanced transistors. Devices under self-heating effects suffer an increase of their temperature, causing mobility reduction, besides compromising reliability and generating signal delays, bringing impacts to the efficiency of analog circuits, and affecting the performance of digital circuits. Despite the relevance of the phenomenon, many studies do not consider it, given the difficulty to assess it, since the methods used for advanced transistors require special structures or equipment, which are rarely available. Hence, three new techniques are developed in this work, with the objective of permitting the study of the effect utilizing conventional structures and direct current measurements: (i) the mean output conductance method; (ii) the signature in the transistor efficiency method; (iii) the thermal resistance estimative using only direct current measurements. The first two methods are focused on a qualitative analysis of the self-heating, allowing an efficient preliminary verification of the presence and relevance of the effect, while the last allows the extraction of the thermal resistance from the inverse of the transistor efficiency through an iterative process, consequently making it possible to obtain the temperature rise in the channel due to the self-heating with a good precision and greater simplicity when compared to other methods available in the literature (with maximum error smaller than 6% for multiple gate transistors when compared to the pulsed method). With these techniques, comparisons between multiple gate transistors (also known as FinFET or 3D transistors) and silicon-on-oxide with ultra-thin body and buried oxide (SOI UTBB) are performed, utilizing three-dimensional simulations to obtain similar power conditions. In devices with smaller channel length, FinFETs presented temperatures approximately 60 K above the UTBBs.

Keywords: self-heating, transistors, SOI, UTBB, FinFET.

LISTA DE FIGURAS

Figura 1.1 – Resumo das tecnologias empregadas em nós passados e previsões para nós futuros.....	14
Figura 2.1 – Relação de capacitâncias entre porta e substrato em um dispositivo (a) MOS convencional ou SOI parcialmente depletado e (b) SOI totalmente depletado	18
Figura 2.2 – Transistor SOI UTBB.....	20
Figura 2.3 – Transistor FinFET de porta dupla.....	21
Figura 2.4 – Transistor FinFET de porta tripla.....	21
Figura 2.5 – FinFET SOI com múltiplas aletas	23
Figura 2.6 – Perfil de cargas em um transistor em saturação com (a) $V_{DS} = V_{GT}$ e (b) $V_{DS} > V_{GT}$	23
Figura 2.7 – Curvas de corrente de dreno em função da tensão de dreno, destacando o efeito de modulação do comprimento de canal na saturação	24
Figura 2.8 – Exemplo do método da segunda derivada para extração da tensão de limiar.....	27
Figura 2.9 – Método da extrapolação em saturação para extração da tensão de limiar	28
Figura 2.10 – Características experimentais (linhas cheias) e modeladas (símbolos) da função $I_D/g_m(Y)$ para vários comprimentos de porta entre 28 nm e 1 μm	31
Figura 2.11 – (a) Medidas de corrente realizadas com diferentes larguras de pulsos de tensão aplicados; (b) Comparação entre correntes medidas de forma pulsada e de forma contínua; (c) extração da resistência térmica a partir da temperatura estimada do canal.....	34
Figura 2.12 – (a) Caracterização da condutância de saída em função da frequência; (b) determinação da sensibilidade da corrente de dreno à temperatura externa	36
Figura 3.1 – Fluxograma da rotina de simulação em MATLAB	37
Figura 3.2 – Corrente de dreno em função da tensão de dreno simulada de um dispositivo pMOS para diferentes resistências térmicas	38
Figura 3.3 – Corrente de dreno de um dispositivo pMOS simulada em função da tensão de porta para múltiplas tensões de dreno sem resistência térmica e com $R_{th} = 50 K/mW$	39

Figura 3.4 – Condutância de saída média de um dispositivo pMOS simulada em função da tensão de porta para múltiplas tensões de dreno (a) sem resistência térmica e (b) com $R_{th} = 50 \text{ K/mW}$	40
Figura 3.5 – Ênfase na região de saturação das curvas de condutância de saída (a) sem resistência térmica e (b) com $R_{th} = 50 \text{ K/mW}$	41
Figura 3.6 – Influência do parâmetro λ na condutância de saída média	42
Figura 3.7 – Ênfase na região de saturação das curvas de condutância de saída de transistores FinFET de dimensões (a) grandes e (b) pequenas	44
Figura 3.8 – Comparação do efeito de autoaquecimento em dispositivos com três diferentes comprimentos de canal ($L = 130 \text{ nm}$, $L = 250 \text{ nm}$ e $L = 1 \text{ }\mu\text{m}$)	45
Figura 4.1 – Eficiência do transistor em função da corrente de dreno sem autoaquecimento (linhas tracejadas) e com (linhas cheias) para múltiplas sobretensões de condução, com destaque para a região de saturação	48
Figura 4.2 – Esquematização da influência do autoaquecimento na eficiência do transistor.....	49
Figura 4.3 – Comparação do método da assinatura na eficiência do transistor entre dispositivos com diferentes λ , $R_{th} = 50 \text{ K/mW}$ e (a) baixo $ V_{GT} $, (b) $ V_{GT} $ intermediário e (c) $ V_{GT} $ alto.....	51
Figura 4.4 – Comparação da eficiência do transistor em função da tensão de dreno para dispositivos com diferentes resistências térmicas e λ	52
Figura 4.5 – Verificação experimental da assinatura na eficiência do transistor em dispositivo com $L = 1 \text{ }\mu\text{m}$, $W_{eff} = 145 \text{ nm}$ e (a) $ V_{GT} = 0,1 \text{ V}$, (b) $ V_{GT} = 0,85 \text{ V}$ e (c) $ V_{GT} = 1,2 \text{ V}$	53
Figura 4.6 – Verificação experimental da assinatura na eficiência do transistor em dispositivo com $L = 250 \text{ nm}$, $W_{eff} = 145 \text{ nm}$ e (a) $ V_{GT} = 0,1 \text{ V}$, (b) $ V_{GT} = 0,45 \text{ V}$ e (c) $ V_{GT} = 1,2 \text{ V}$	53
Figura 4.7 – Verificação experimental da assinatura na eficiência do transistor em dispositivo com $L = 130 \text{ nm}$, $W_{eff} = 145 \text{ nm}$ e (a) $ V_{GT} = 0,1 \text{ V}$, (b) $ V_{GT} = 0,45 \text{ V}$ e (c) $ V_{GT} = 1,2 \text{ V}$	54
Figura 4.8 – Aplicação experimental do método eficiência do transistor em função da tensão de dreno para dispositivos com diferentes comprimentos de canal	54
Figura 5.1 – Fluxograma do método para estimativa da resistência térmica	61

Figura 5.2 – Estrutura gerada para o transistor FinFET (a) tridimensional (b) corte evidenciando W_{fin} (c) corte evidenciando L (d) representação esquemática das dimensões.....	63
Figura 5.3 – Estimativa da temperatura no canal a partir das curvas sem autoaquecimento em múltiplas temperaturas e uma curva com autoaquecimento a temperatura ambiente para (a) $L = 150$ nm, (b) $L = 300$ nm e (c) $L = 500$ nm	64
Figura 5.4 – Extração da resistência térmica a partir da temperatura estimada no canal em função da potência aplicada ao dreno.....	65
Figura 5.5 – Comparação das resistências térmicas extraídas a partir do método proposto com o método tradicional pulsado em função do comprimento de canal do dispositivo FinFET	66
Figura 5.6 – Comparação das temperaturas estimadas no canal dos dispositivos através do método tradicional pulsado e do método proposto para diferentes comprimentos de canal	66
Figura 6.1 – Estrutura gerada para o transistor UTBB (a) tridimensional (b) corte evidenciando W_{UTBB} (c) corte evidenciando L_{UTBB} (d) desenho esquemático com as principais dimensões utilizadas	69
Figura 6.2 – Estimativa da temperatura no canal a partir das curvas sem autoaquecimento em múltiplas temperaturas e uma curva com autoaquecimento a temperatura ambiente para (a) $L = 150$ nm, (b) $L = 300$ nm e (c) $L = 500$ nm	70
Figura 6.3 – Extração da resistência térmica a partir da temperatura estimada no canal em função da potência aplicada ao dreno para dispositivo UTBB	71
Figura 6.4 – Comparação das temperaturas de operação entre dispositivos FinFET e UTBB na situação de potência elevada obtida a partir do método de extração pulsado	72
Figura 6.5 – Comparação entre dispositivos FinFET e UTBB utilizando o método da assinatura na eficiência do transistor, com sobretensão de condução de 1 V e (a) $L = 150$ nm, (b) $L = 300$ nm e (c) $L = 500$ nm.....	72
Figura 6.6 – Comparação das temperaturas de operação entre dispositivos FinFET e UTBB na situação de potência elevada obtida a partir do método de extração proposto	74

LISTA DE SÍMBOLOS

I_D	Corrente de dreno [A]
I_{Dsat}	Corrente de dreno em saturação [A]
P	Potência elétrica [W]
g_m	Transcondutância [A/V]
g_D	Condutância de saída [A/V]
μ_{eff}	Mobilidade dos portadores de baixo campo e à temperatura de referência [$m^2/(V.s)$]
μ_{eff}	Mobilidade efetiva dos portadores [$m^2/(V.s)$]
μ_{sup}	Mobilidade dos portadores na interface superior [$m^2/(V.s)$]
μ_{lat}	Mobilidade dos portadores na interface lateral [$m^2/(V.s)$]
C_{ox}	Capacitância do óxido de porta [F/m^2]
C_{BOX}	Capacitância do óxido enterrado [F/m^2]
C_{Si}	Capacitância da camada de silício [F/m^2]
C_D	Capacitância da região de depleção [F/m^2]
t_{ox}	Espessura do óxido de porta [m]
t_{BOX}	Espessura do óxido enterrado [m]
t_{Si}	Espessura da camada de silício [m]
W	Largura do canal do transistor [m]
W_{fin}	Largura da aleta de transistores FinFET [m]
h_{fin}	Altura da aleta de transistores FinFET [m]
W_{eff}	Largura efetiva de transistores FinFET [m]
L	Comprimento do canal do transistor [m]
V_{GS}	Tensão entre porta e fonte [V]
V_t	Tensão de limiar [V]
V_{GT}	Sobretensão de condução [V]
V_{DS}	Tensão entre dreno e fonte [V]
n	Fator de corpo
Φ_{MS}	Diferença de função trabalho entre metal e semiconductor [V]
ϕ_s	Potencial de superfície entre silício e porta [V]
ϕ_F	Potencial de Fermi [V]
ϕ_{S1}	Potencial de superfície entre silício e porta em um dispositivo de

	duas portas [V]
ϕ_{S2}	Potencial de superfície entre silício e óxido enterrado em um dispositivo de duas portas [V]
V_A	Tensão Early [V]
λ	Constante de ajuste do efeito de modulação do comprimento de canal [V^{-1}]
θ	Fator de degradação da mobilidade por campo elétrico lateral [V^{-1}]
c	Coeficiente de degradação da mobilidade com a temperatura
T	Temperatura de operação [K]
T_0	Temperatura de referência [K]
R_{th}	Resistência térmica [K/W]
F	Fator relacionado ao autoaquecimento na eficiência do transistor em saturação
G	Inverso de F
T_m	Série de Taylor de ordem m
M	Fator independente do autoaquecimento no inverso da eficiência do transistor em saturação

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVOS	15
1.2	ESTRUTURA DO TRABALHO	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	TRANSISTORES AVANÇADOS	17
2.1.1	Camada de silício e óxido enterrado ultrafinos	19
2.1.2	FinFET	20
2.2	EFEITO DE MODULAÇÃO DO COMPRIMENTO DE CANAL	23
2.3	TENSÃO DE LIMIAR	25
2.4	TRANSCONDUTÂNCIA	29
2.5	EFICIÊNCIA DO TRANSISTOR	29
2.6	MOBILIDADE	29
2.7	CONDUTÂNCIA DE SAÍDA.....	32
2.8	AUTOAQUECIMENTO	32
3	MÉTODO DA CONDUTÂNCIA DE SAÍDA MÉDIA.....	37
3.1	DESENVOLVIMENTO DO MÉTODO E SIMULAÇÕES	38
3.2	VERIFICAÇÃO EXPERIMENTAL.....	43
3.2.1	Características dos dispositivos	43
3.2.2	Resultados	44
3.2.3	Análise	45
4	MÉTODO DA ASSINATURA NA EFICIÊNCIA DO TRANSISTOR	47
4.1	DESENVOLVIMENTO DO MÉTODO E SIMULAÇÕES	48
4.2	VERIFICAÇÃO EXPERIMENTAL.....	52
4.2.1	Resultados	52
4.2.2	Análise	55
5	ESTIMATIVA DA RESISTÊNCIA TÉRMICA UTILIZANDO MEDIDAS EM CORRENTE CONTÍNUA	57

5.1	LINEARIZAÇÃO DO INVERSO DA EFICIÊNCIA DO TRANSISTOR.....	57
5.2	EXTRAÇÃO DOS PARÂMETROS C , M E R_{TH}	59
5.3	SIMULAÇÕES NUMÉRICAS.....	62
5.4	RESULTADOS	64
5.5	DISCUSSÕES.....	67
6	COMPARAÇÃO ENTRE DISPOSITIVOS FINFET E SOI UTBB	68
6.1	SIMULAÇÃO DOS DISPOSITIVOS.....	68
6.2	MÉTODO BASEADO NA EXTRAÇÃO PULSADA	70
6.3	MÉTODO DA ASSINATURA NA EFICIÊNCIA DO TRANSISTOR.....	72
6.4	MÉTODO NUMÉRICO A PARTIR DO INVERSO DA EFICIÊNCIA DO TRANSISTOR	73
6.5	DISCUSSÃO	74
7	CONCLUSÕES	77
	PUBLICAÇÕES GERADAS.....	80
	BIBLIOGRAFIA.....	81
	APÊNDICE A – ROTINA EM MATLAB PARA ESTIMATIVA DA RESISTÊNCIA TÉRMICA A PARTIR DE MEDIDAS EM CORRENTE CONTÍNUA	85

1 INTRODUÇÃO

Desde a criação do primeiro circuito integrado (KILBY, 1959), a indústria da microeletrônica busca a redução dos transistores, causando tanto um aumento na capacidade de processamento dos circuitos como a possibilidade de uma maior integração de funções. Os rápidos avanços nessa indústria em seus primeiros anos foram determinantes para o seu futuro pelas próximas décadas, graças às observações feitas por Moore (1965). O padrão de aumento exponencial do número de transistores em um circuito integrado (CI) observado ao longo de seis anos tornou-se a famosa lei de Moore, que governaria a indústria por mais de 50 anos. Entretanto, o que inicialmente era uma simples observação empírica, logo viria a ser uma profecia auto realizada, isto é, houve um enorme esforço por parte da indústria para se manter dentro das projeções de Moore, incluindo cooperação entre concorrentes para estipulação de metas e desafios a serem vencidos com a criação da *“International Roadmap for Semiconductors”* (WALDROP, 2016).

Atualmente, avalia-se que a lei tenha chegado ao seu fim (SIMONITE, 2016), não necessariamente no sentido de que inovações tecnológicas para redução dos transistores sejam impossíveis, mas sim que não seja possível seguir mais o seu ritmo exponencial. Dessa forma, tendo em vista as demandas do mercado para o aumento tanto do consumo doméstico, com as grandes promessas da Internet das Coisas (MORGAN, 2014), bem como das pesquisas em inteligência artificial e aprendizado de máquina (BUCK, 2017), é evidente que a micro e nanoeletrônica ainda tem um longo caminho a percorrer.

De forma a permitir a contínua redução das dimensões dos transistores, novas tecnologias foram introduzidas ao longo dos anos, de modo a controlar melhor as não-idealidades, que passam a ser cada vez mais relevantes para os nós menores. Iwai (2013) apresenta um resumo das tecnologias mais recentes utilizadas, bem como das principais frentes de pesquisa sendo atualmente desenvolvidas para aplicações futuras. A Figura 1.1 ilustra essas tecnologias, bem como os respectivos nós associados.

Figura 1.1 – Resumo das tecnologias empregadas em nós passados e previsões para nós futuros

Fonte: adaptado de Iwai (2013).

Evidencia-se na Figura 1.1 que as duas principais tecnologias atualmente no mercado e que devem ainda competir em nós futuros são a de silício sobre isolante com corpo e óxido enterrados ultrafinos, em inglês “*Silicon On Insulator*” (SOI) e “*Ultra Thin Body and Buried oxide*” (UTBB), e o transistor de efeito de campo vertical de porta tripla (FinFET).

Um problema crescente em tecnologias avançadas é o efeito de autoaquecimento, em inglês “*self-heating effect*” (SHE), primeiramente observado por McDaid et al. (1989) para uma tecnologia SOI. O efeito em si consiste no aumento da temperatura interna do dispositivo em função da dificuldade de dissipação do calor gerado devido à aplicação de potência elétrica ao seu canal. Assim, como o óxido enterrado em dispositivos SOI tem baixa condutividade térmica, a dissipação do calor gerado no canal é comprometida, portanto, esses dispositivos apresentam temperaturas de operação mais elevadas do que tecnologias tradicionais. Apesar de não ser um problema imediato para aplicações digitais, devido às altas frequências de chaveamento empregadas que levam a um efeito de autoaquecimento desprezível a nível de dispositivo (MAKOVEJEV, 2012), o autoaquecimento pode comprometer o desempenho de aplicações analógicas e a extração de parâmetros dos transistores, uma vez que normalmente ela é feita em condições de corrente contínua.

Consequentemente, todas as etapas de projeto subsequentes podem ser negativamente influenciadas pela imprecisão durante a caracterização dos dispositivos.

Apesar da extensa caracterização feita em transistores das tecnologias SOI UTBB e FinFET, como feitas por Nascimento (2017), Macambira (2017) e Caparroz (2017), dificilmente levam-se em conta os efeitos do autoaquecimento durante o processo da extração de parâmetros, sendo normalmente conduzidos estudos específicos para a caracterização desse fenômeno, como feito por Makovejev (2012), Takahashi et al. (2013) e Jin et al. (2016). Possíveis explicações para a ausência de comentários sobre o autoaquecimento em estudos não direcionados a ele são a dificuldade em conduzir medidas para sua caracterização, e a inexistência de metodologias capazes de verificar de forma simples e eficiente (por exemplo, utilizando somente medidas em corrente contínua) a sua existência em dispositivos. Atualmente, para realizar esta caracterização, estruturas especiais, como transistores com quatro contatos de porta, ou com contatos especiais para caracterização em radiofrequência, bem como aparelhos capazes de operar com períodos abaixo de cinco nanossegundos são necessários, evidenciando que o desenvolvimento de uma metodologia simples capaz de identificar a intensidade do efeito de autoaquecimento se faz necessária tanto em âmbito nacional como internacional.

1.1 OBJETIVOS

O objetivo principal deste trabalho é a comparação entre dispositivos SOI UTBB e FinFET do ponto de vista do autoaquecimento, verificando a influência das dimensões geométricas dos dispositivos, como comprimento e largura de canal, e também estabelecendo paralelos entre os efeitos do autoaquecimento sobre parâmetros elétricos dos transistores de ambas as tecnologias.

Além disso, estabelece-se como objetivo secundário o desenvolvimento de novas metodologias para a caracterização do autoaquecimento, uma vez que todas as técnicas atualmente empregadas apresentam alto grau de complexidade na sua implementação, comprometendo a disseminação do estudo desse efeito. A princípio, não é estritamente necessário que todos os métodos desenvolvidos resultem em valores numéricos da resistência térmica, mas que permitam eventualmente a comparação entre dispositivos.

1.2 ESTRUTURA DO TRABALHO

O trabalho foi estruturado de forma a apresentar sete capítulos, descritos a seguir:

- a) Capítulo 1 – Introdução: apresenta a motivação histórica para o desenvolvimento desse trabalho e os objetivos estabelecidos;
- b) Capítulo 2 – Fundamentação teórica: revisa os princípios de funcionamento dos transistores a serem estudados, bem como parâmetros elétricos de interesse ao estudo e considerações sobre o autoaquecimento;
- c) Capítulo 3 – Método da condutância de saída média: apresenta e verifica o primeiro método desenvolvido para a caracterização da intensidade do autoaquecimento;
- d) Capítulo 4 – Método da assinatura na eficiência do transistor: desenvolve e testa o segundo método desenvolvido para a caracterização da intensidade do autoaquecimento;
- e) Capítulo 5 – Estimativa numérica da resistência térmica: parte dos resultados obtidos no método da assinatura na eficiência do transistor com o intuito de desenvolver uma rotina capaz de estimar a resistência térmica de um dispositivo qualquer a partir de medidas em corrente contínua;
- f) Capítulo 6 – Comparação entre dispositivos FinFET e SOI UTBB: utiliza os resultados de simulações para comparar diretamente os efeitos das diferentes geometrias sob o ponto de vista de autoaquecimento.
- g) Capítulo 7 – Conclusões: resume os desenvolvimentos e principais resultados do trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Para que se tenha uma compreensão ampla do fenômeno de autoaquecimento e suas consequências para o funcionamento dos transistores, primeiramente será conduzida uma revisão dos mecanismos básicos de funcionamento dos transistores, seguida por um estudo mais aprofundado dos principais parâmetros afetados durante a operação em altas temperaturas.

2.1 TRANSISTORES AVANÇADOS

Tendo em vista que, devido aos efeitos de canal curto, os tradicionais dispositivos planares perdem o controle pleno das cargas no canal conforme as dimensões são reduzidas, e que técnicas de escalamento tradicionais não são mais possíveis, pois levam ao surgimento de efeitos indesejados (como por exemplo uma maior corrente de porta devido ao efeito de tunelamento), novas tecnologias foram introduzidas.

Nessa seção, as duas principais tecnologias abordadas ao longo do trabalho serão estudadas, tendo como objetivo o uso das suas características de tensão e corrente. De forma geral, ambas as tecnologias podem ser adaptadas dentro das equações tradicionais para dispositivos MOS, como apresentado por Colinge (2004) nas formas das equações (1) e (2) para dispositivos operando em triodo e saturação, respectivamente.

$$I_D = \mu_{\text{eff}} \cdot C_{\text{ox}} \cdot (W/L) \cdot [(V_{\text{GS}} - V_t) \cdot V_{\text{DS}} - n \cdot V_{\text{DS}}^2 / 2] \quad (1)$$

$$I_D = \mu_{\text{eff}} \cdot C_{\text{ox}} \cdot (W/L) \cdot (V_{\text{GS}} - V_t)^2 / (2 \cdot n) \quad (2)$$

onde I_D é a corrente de dreno, μ_{eff} é a mobilidade efetiva dos portadores, C_{ox} é a capacitância do óxido de porta, W é a largura do canal, L é o comprimento do canal, V_{GS} é a tensão aplicada entre porta e fonte, V_t é a tensão de limiar, V_{DS} é a tensão aplicada entre dreno e fonte e n é o fator de corpo.

Para que ocorra a transição da operação em triodo para saturação, é necessário que a tensão entre dreno e fonte atinja pelo menos o valor da diferença entre as tensões de porta e limiar (também chamada de sobretensão de condução, V_{GT}) dividida pelo fator de corpo, isto é

$$V_{\text{DS}} \geq (V_{\text{GS}} - V_t) / n = V_{\text{GT}} / n \quad (3)$$

O fator de corpo n é primeiramente empregado como uma constante que simplifica a aproximação de forma linear da tensão de limiar em função da tensão de fonte (tomando-se o substrato como referência). A equação (4) introduz outra constante, α , inicialmente obtida ao se fazer o processo matemático de linearização da expressão de tensão de limiar em dispositivos MOS convencionais em função da tensão de fonte.

$$n = 1 + \alpha \quad (4)$$

Uma outra interpretação que pode ser dada para a constante α é como uma relação das capacitâncias vistas a partir da camada de inversão. A Figura 2.1 apresenta esquematicamente a composição das capacitâncias para dispositivos MOS convencionais e SOI.

Figura 2.1 – Relação de capacitâncias entre porta e substrato em um dispositivo (a) MOS convencional ou SOI parcialmente depletado e (b) SOI totalmente depletado

Fonte: adaptado de Colinge (2004)

Na Figura 2.1 são introduzidas as diversas capacitâncias relacionadas com dispositivos MOS: C_{ox} é a capacitância do óxido de porta, C_D é a capacitância da região de depleção, C_{Si} é a capacitância da camada de silício em dispositivos totalmente depletados e C_{BOX} é a capacitância do óxido enterrado. São apresentados também os potenciais de superfície, relativos tanto à interface entre silício e óxido de

porta (Φ_s , Φ_{s1}), como à interface entre camada de silício e óxido enterrado (Φ_{s2}). Dessa forma, α assume os seguintes valores:

$$\alpha_{\text{conv}} = C_D / C_{\text{ox}} \quad (5)$$

$$\alpha_{\text{SOI,depl}} = (C_{\text{Si}} \cdot C_{\text{BOX}}) / [C_{\text{ox}} \cdot (C_{\text{Si}} + C_{\text{BOX}})] \quad (6)$$

$$\alpha_{\text{SOI,acc2}} = C_{\text{Si}} / C_{\text{ox}} \quad (7)$$

em que α_{conv} é a relação de capacitâncias para dispositivos MOS convencionais ou dispositivos SOI parcialmente depletados, $\alpha_{\text{SOI,depl}}$ é para dispositivos SOI totalmente depletados, e $\alpha_{\text{SOI,acc2}}$ é para dispositivos SOI de camada fina com a segunda interface em acumulação. Deve-se notar que a polarização da segunda interface pode agir de forma a isolar C_{Si} de C_{BOX} , caso ela esteja em acumulação, por isso $\alpha_{\text{SOI,acc2}}$ é análogo ao dispositivo convencional, substituindo-se C_D por C_{Si} .

Como consequência das relações apresentadas, considerando-se que em um dispositivo totalmente depletado a espessura da camada de silício é menor do que a profundidade da região de depleção e que $C_{\text{BOX}} < C_D$ devido à menor permissividade do óxido de silício, observa-se a relação

$$\alpha_{\text{SOI,depl}} < \alpha_{\text{conv}} < \alpha_{\text{SOI,acc2}} \quad (8)$$

Naturalmente, dadas as diferentes geometrias dos dispositivos a serem estudados, adaptações deverão ser feitas aos elementos das equações apresentadas acima.

2.1.1 Camada de silício e óxido enterrado ultrafinos

O avanço das tecnologias SOI foi atingido através do afinamento tanto da camada de silício do canal como do óxido enterrado, resultando na tecnologia UTBB, ilustrada pelo dispositivo na Figura 2.2.

Na Figura 2.2, são introduzidas as espessuras do óxido (t_{ox}), da camada de silício (t_{si}) e do óxido enterrado (t_{BOX}), em inglês “*Buried Oxide*” (BOX). Usualmente, a interface entre óxido de porta e camada de silício é chamada de primeira interface, enquanto que a interface entre camada de silício e óxido enterrado é a segunda interface. Apresenta-se também outra modificação da tecnologia, o plano de terra, posicionado imediatamente abaixo do óxido enterrado. Trata-se de uma região de alta dopagem, implantada com o intuito de minimizar a extensão da região de depleção abaixo do óxido enterrado, e assim controlar a capacitância associada ao óxido

enterrado, e consequentemente a tensão de limiar (FENOUILLET-BERANGER et al., 2009).

Figura 2.2 – Transistor SOI UTBB

Fonte: adaptado de Almeida (2012).

A tecnologia UTBB surge como um avanço dentro da tecnologia SOI, apresentando como principal diferencial a redução das espessuras t_{si} e t_{BOX} para valores menores do que 20 nm, de forma a permitir um melhor controle das cargas no canal sem a necessidade de grandes modificações no processo de fabricação, uma vez que a redução das espessuras da camada de silício e do óxido enterrado garantem uma redução dos efeitos de canal curto (COLINGE, 2008).

Para a análise das características de tensão e corrente, utilizam-se os resultados descritos em Colinge (2004) a partir da abordagem desenvolvida por Lim e Fossum (1984). Como o trabalho visa o estudo do autoaquecimento em condições usuais de operação, são considerados apenas os casos em que a segunda interface está totalmente depletada ou em acumulação. Assim, dentro das equações (1) e (2), basta considerar o fator de corpo como apresentado em (4), com α dado por (6) ou (7), dependendo da condição da segunda interface.

2.1.2 FinFET

Dentre os transistores de efeito de campo de múltiplas portas (MuGFET – “Multiple Gate Field Effect Transistor”), o FinFET é um dos primeiros dispositivos

desenvolvidos a não apresentar uma estrutura planar, isto é, as regiões de canal, fonte e dreno não são mais simplesmente implantadas no corpo do dispositivo, mas sim são construídas em uma aleta (“*fin*”), justificando dessa forma o seu nome. Originalmente, o termo era utilizado somente para se referir a dispositivos de apenas duas portas, como apresentado na Figura 2.3, entretanto, atualmente é utilizado também para dispositivos de porta tripla, como ilustrado na Figura 2.4.

Figura 2.3 – Transistor FinFET de porta dupla

Fonte: adaptado de Colinge (2008).

Figura 2.4 – Transistor FinFET de porta tripla

Fonte: o autor (2017).

Na Figura 2.4 foram destacadas as principais dimensões dos dispositivos SOI FinFET, o comprimento de canal (L), a largura da aleta (W_{fin}) e a altura da aleta (h_{fin}). Tanto os dispositivos de porta dupla como de porta tripla foram representados como dispositivos SOI, como os que serão estudados nesse trabalho; contudo, deve-se

destacar que a sua fabricação é possível em substratos convencionais, como feito pela Intel (2011). Enfatiza-se ainda que a única diferença significativa do ponto de vista de construção entre os casos de porta dupla e tripla é a espessura do óxido de porta na parte superior. Tendo em vista a ambiguidade do termo, define-se que neste trabalho o termo FinFET será empregado somente para dispositivos de porta tripla, a menos que seja especificado o contrário.

A partir das definições geométricas desses transistores, nota-se que algumas adaptações se fazem necessárias ao modelo tradicional de corrente, especialmente no que diz respeito à largura do transistor. Kilchytska et al. (2004) executou um estudo detalhado da influência das dimensões nos parâmetros de transistores FinFET, verificando que dispositivos com aletas mais estreitas ($W_{fin} < 85$ nm) apresentavam melhores características elétricas que as mais largas, uma vez que a camada de silício da aleta exibe um melhor acoplamento eletrostático entre porta e canal. Esse estudo também considera que, para se estimar a largura efetiva (W_{eff}) de um dispositivo de porta tripla, pode-se utilizar a aproximação indicada na equação (9)

$$W_{eff} \cong W_{fin} + 2.h_{fin} \quad (9)$$

ou seja, considera-se que tanto as duas interfaces laterais como a superior funcionam como se fossem canais independentes. Devido a menor mobilidade da interface superior, se W_{fin} for suficientemente estreito, ele pode ser considerado desprezível.

Daugé et al. (2004) analisa o comportamento da mobilidade nas diferentes interfaces dos FinFETs, e relata a significativa diferença nas mobilidades das interfaces laterais e superior. Assim, para se considerar a mobilidade efetiva em dispositivos de porta tripla sem alterar a forma geral das equações (1) e (2), é necessário realizar uma média ponderada das contribuições de cada interface para a mobilidade total do transistor na forma

$$\mu_{eff} = (W_{fin}.\mu_{sup} + 2.h_{fin}.\mu_{lat})/(W_{fin} + 2.h_{fin}) \quad (10)$$

onde μ_{sup} e μ_{lat} são as mobilidades das interfaces superior e lateral, respectivamente. Rudenko et al. (2008) constata que essa aproximação não é precisa para aletas relativamente estreitas, possivelmente devido a deformações da rede cristalina do material entre as interfaces laterais e superior; todavia, para os propósitos desse trabalho, ela é suficiente para estabelecer as comparações necessárias e será dessa forma utilizada.

Uma última consideração sobre a modelagem de corrente feita por Colinge (2008) é que, para se atingirem níveis de corrente mais elevados, é necessário o uso

de configurações com múltiplas aletas em paralelo, com periodicidade dada pelo passo (P), como ilustrado na Figura 2.5. Portanto, a expressão de corrente deve ser multiplicada pelo número de aletas (N_{fin}) para obter a corrente total de um dispositivo.

Figura 2.5 – FinFET SOI com múltiplas aletas

Fonte: adaptado de Colinge (2008).

2.2 EFEITO DE MODULAÇÃO DO COMPRIMENTO DE CANAL

Uma vez atingido o regime de saturação, a representação usual de corrente equacionada em (2) se torna independente da tensão aplicada entre dreno e fonte. Contudo, esse modelo é bastante simplificado, e traria problemas nas análises de autoaquecimento, pois, como será discutido posteriormente, esse efeito é mais pronunciado em situações com alta polarização de dreno, exigindo consequentemente modelos mais específicos nesse regime. O efeito de modulação do comprimento de canal é um dos mais básicos a ser considerado, mas que mesmo assim permite ajustes bastante precisos do comportamento da corrente de dreno em saturação. Seu princípio básico é ilustrado na Figura 2.6.

Figura 2.6 – Perfil de cargas em um transistor em saturação com (a) $V_{DS} = V_{GT}$ e (b) $V_{DS} > V_{GT}$

Fonte: adaptado de Foty (1997).

De forma simplificada, quando se aplica uma tensão de dreno equiparável à sobretensão de condução o canal do dispositivo é pinçado junto à região de dreno, como pode ser visto na Figura 2.6(a), uma vez que a depleção imposta pela polarização de dreno se sobrepõe à formação de canal induzida pela tensão de porta. Para tensões de dreno mais elevadas, o ponto de pinçamento é deslocado em direção à fonte, o que significa que o comprimento de canal real do dispositivo é diminuído. A Figura 2.6(b) ilustra essa diminuição, chamando de L' o deslocamento total do ponto de pinçamento. Deve-se notar que, como modelado por Baum e Beneking (1970), esse deslocamento é dependente principalmente das polarizações aplicadas e da dopagem do substrato. Como consequência dessa diminuição da largura efetiva de canal, obtém-se uma elevação da corrente de saturação do dispositivo, como apresentado na Figura 2.7.

Figura 2.7 – Curvas de corrente de dreno em função da tensão de dreno, destacando o efeito de modulação do comprimento de canal na saturação

Fonte: adaptado de Sedra e Smith (2010).

Além da inclinação, ou condutância de saída (g_D), positiva observada em saturação devido a esse fenômeno, a Figura 2.7 introduz dois parâmetros relacionados ao efeito, a tensão Early (V_A) e uma constante de ajuste (λ). Sedra e Smith (2010) notam que a tensão Early é diretamente proporcional ao comprimento de canal, assim, para tecnologias menores, nota-se uma inclinação mais acentuada (isto é, maior g_D) das curvas de corrente em saturação.

Foty (1997) descreve diversas formas de quantizar esse efeito com diferentes graus de complexidade. Apesar das formas mais complexas serem mais precisas e fidedignas ao fenômeno físico envolvido, tendo em vista a análise e aplicação desse efeito a uma modelagem do autoaquecimento, é mais interessante o uso do modelo

obtido de forma puramente matemática. Inicialmente, nota-se que o comportamento em saturação do dispositivo pode ser ajustado por uma reta de coeficiente angular a , descrito em (11).

$$a = (I_D - I_{Dsat})/V_{DS} \quad (11)$$

onde I_D representa a corrente total do dispositivo e I_{Dsat} é a corrente de saturação calculada de acordo com a equação (2). A fim de aproveitar este coeficiente, constrói-se a expressão (12), somando e subtraindo I_{Dsat} de um mesmo lado da igualdade $I_D = I_D$.

$$I_D = I_{Dsat} + (I_D - I_{Dsat}) \quad (12)$$

Aplicar (11) em (12) resultaria em $I_D = I_{Dsat} + a.V_{DS}$, mas, se for assumido que $a = I_{Dsat}.\lambda$, com λ utilizado de forma a garantir o ajuste matemático correto, finalmente tem-se a equação (13).

$$I_D = I_{Dsat}.(1 + \lambda V_{DS}) \quad (13)$$

2.3 TENSÃO DE LIMIAR

Tradicionalmente, a tensão de limiar é definida como “[...] tensão a partir da qual todo incremento de tensão acima desta será utilizada para a formação da camada de inversão [...]” (MARTINO, PAVANELLO e VERDONCK, 2003). Essa situação é conhecida como inversão forte, e é atingida quando o potencial elétrico na superfície (Φ_S) entre óxido e semiconductor é o dobro do potencial de Fermi (Φ_F) do canal, ou seja, $\Phi_S = 2.\Phi_F$.

Para dispositivos SOI UTBB totalmente depletados devido à pequena espessura da camada de silício, é necessário levar-se em conta o acoplamento entre os potenciais da primeira e da segunda interface (Φ_{S1} e Φ_{S2} , respectivamente). Como já discutido, serão utilizados somente os modos de operação com a segunda interface em acumulação ($\Phi_{S2} \approx 0$) e depleção ($0 < \Phi_{S2} < 2.\Phi_F$), portanto, de acordo com Colinge (2004), pode-se considerar que a tensão de limiar da primeira interface seja a indicada na equação (14).

$$V_{t1} = \Phi_{MS} - Q_{ox1}/C_{ox1} + (1 + C_{Si}/C_{ox1}).2.\Phi_F - C_{Si}/C_{ox1}.\Phi_{S2} - Q_{depl}/(2.C_{ox1}) \quad (14)$$

sendo Φ_{MS} a diferença de função trabalho entre material de porta e semiconductor do canal, Q_{ox1} a densidade de carga no óxido da primeira interface, C_{ox1} a capacitância por unidade de área do óxido da primeira interface, C_{Si} a capacitância por unidade de

área da camada de silício e Q_{depl} a densidade de cargas de depleção. No caso de um dispositivo tipo n totalmente depletado, Q_{depl} será dado por (15).

$$Q_{\text{depl}} = -q.N_A.t_{\text{Si}} \quad (15)$$

onde q é a carga do elétron ($\approx 1,6.10^{-19}$ C) e N_A é a concentração de impurezas aceitadoras na camada de silício. Caso a camada de silício utilizada seja suficientemente fina ($t_{\text{Si}} < 10$ nm), efeitos quânticos passam a afetar significativamente a tensão de limiar do dispositivo, como apresenta Ernst et al. (2003).

Para dispositivos de múltiplas portas como o FinFET, Poiroux et al. (2005) observa que o uso da definição de inversão forte não representa bem a mudança do regime de sublimiar para condução normal dos dispositivos, e troca-a por uma definição da condição de limiar em função da densidade de cargas de inversão (Q_{inv})

$$Q_{\text{inv}} = -2.C_{\text{ox}}.k.T/q \quad (16)$$

sendo k a constante de Boltzmann. A partir dessa definição, Poiroux calcula a tensão de limiar para dispositivos de múltiplas portas

$$V_t = \Phi_{\text{MS}} + k.T/q.\ln(2.C_{\text{ox}}.k.T/(q^2.n_i.W_{\text{fin}})) + \hbar^2.\pi^2/(2.q.m^*.W_{\text{fin}}^2) \quad (17)$$

em que n_i é a concentração de portadores intrínseca, $\hbar$ é a constante de Dirac e m^* é a massa de confinamento do portador na direção transversal. O último termo da expressão se refere aos efeitos quânticos, que passam a ser significativos para larguras de aleta inferiores a 7 nm.

Para esse trabalho em particular, é interessante ainda analisar a influência da temperatura na tensão de limiar, tanto para dispositivos SOI UTBB como FinFET. Uma possibilidade para esse estudo seria derivar as expressões (14) e (17) em função da temperatura, como feito por Groeseneken et al. (1990) no caso da primeira, resultando em

$$dV_t/dT = d\Phi_F/dT = 8,63.10^{-5}.(\ln(N_A) - 38,2 - 1,5.(1 + \ln(T))) \quad (18)$$

A variação da tensão de limiar com a temperatura permanece aproximadamente constante para as faixas de temperatura usuais, permanecendo na faixa de 0,7 mV/K a 0,8 mV/K entre as temperaturas de 25°C e 200°C, de acordo com Groeseneken.

Para dispositivos FinFET não há uma dedução particular dessa variação, contudo, experimentos feitos por Pavanello et al. (2007) indicam que uma tendência linear também é observada na relação entre tensão de limiar e temperatura. Dessa forma, como sugerido por Arora (1993), pode-se simplesmente assumir que

$$V_t \approx V_{t0} + dV_t/dT.(T-T_0) \quad (19)$$

sendo V_t a tensão de limiar medida à temperatura T , V_{t0} a tensão de limiar medida à temperatura T_0 , com dV_t/dT constante e determinado empiricamente através de medidas de V_t em diversas temperaturas.

Para a extração da tensão de limiar nesse trabalho, será empregado o método da segunda derivada, descrito inicialmente por Wong et al. (1987) e revisado por Ortiz-Conde et al. (2002). Como em um primeiro momento não se deseja observar a interação do autoaquecimento com a tensão de limiar, tensões tanto de porta como de dreno baixas são desejáveis, configuração essa facilmente obtida em polarizações de triodo; portanto, o pré-requisito para a aplicação dessa técnica é atendido.

O método consiste em realizar a segunda derivada da curva de corrente de dreno em função da tensão de porta, de tal forma que se obtém uma curva que apresentará um valor máximo em função da tensão de porta. A tensão de porta para a qual ocorre esse máximo será a tensão de limiar, extraída sem a influência das resistências série de fonte e dreno, que podem ser significativas em dispositivos de menores dimensões. A Figura 2.8 apresenta um exemplo de extração com ênfase a dois valores de tensão, $V_{GS} = 0,54$ V e $V_t = 0,5$ V, sendo que o primeiro se refere ao máximo encontrado antes da filtragem da curva, e o segundo após. Portanto, como observado por Ortiz-Conde, o uso da segunda derivada é bastante dependente de filtragens adequadas. Neste trabalho, será empregado o método descrito por Savitzky e Golay (1964), particularmente útil para suavização de curvas ruidosas.

Figura 2.8 – Exemplo do método da segunda derivada para extração da tensão de limiar

Fonte: adaptado de Ortiz-Conde et al. (2002).

Entretanto, em um estudo mais aprofundado do autoaquecimento, extrações avançadas de alguns parâmetros em saturação se fazem necessários, e, para tanto, utilizaremos outro método apresentado por Ortiz-Conde et al. (2002), a extrapolação na região de saturação. Para este método, levanta-se a curva da raiz da corrente de dreno em função da tensão de porta e, realizando uma extrapolação da região linear da curva, obtém-se o valor da tensão de limiar. Normalmente a polarização de dreno é variada junto com a tensão de porta, garantindo dessa forma a operação em saturação; entretanto, como o autoaquecimento será fortemente influenciado por ambas as tensões, aqui será optado pelo uso de uma tensão de dreno constante e suficientemente elevada (no mínimo o valor máximo adotado como tensão de porta), de modo a minimizar os parâmetros sendo variados. A Figura 2.9 ilustra um processo de extração com tensão de dreno variando junto com a de porta.

Figura 2.9 – Método da extrapolação em saturação para extração da tensão de limiar

Fonte: adaptado de Ortiz-Conde et al. (2002).

2.4 TRANSCONDUTÂNCIA

A definição mais tradicional da transcondutância (g_m) pode ser encontrada em Sedra e Smith (2010)

$$g_m \triangleq \partial I_D / \partial V_{GS} \quad (20)$$

Ou seja, a transcondutância representa o controle que a tensão de porta tem sobre a corrente de dreno. Pode ser utilizada para extrair um dos parâmetros mais importantes do transistor, a mobilidade, como descrito em Martino, Pavanello e Verdonck (2003); todavia, para nós tecnológicos mais recentes, é possível que essa técnica apresente resultados imprecisos, uma vez que a degradação da mobilidade pode ser significativa antes que a transcondutância atinja seu valor máximo. Dessa forma, Colinge (2004) recomenda o uso da técnica apresentada por Ghibaudo (1988), a qual emprega a razão entre a transcondutância e a raiz quadrada da corrente de dreno, de forma a eliminar a influência da tensão de porta na degradação da mobilidade.

2.5 EFICIÊNCIA DO TRANSISTOR

Parâmetro geralmente relacionado com o desempenho analógico, a eficiência do transistor “representa a amplificação obtida (g_m) dividida pela energia fornecida para se atingir essa amplificação (I_D)” (COLINGE, 2004, tradução nossa). Normalmente analisa-se seu comportamento em função do grau de inversão, assim, na inversão fraca

$$g_m/I_D = q/(nkT) \quad (21)$$

enquanto que, para inversão forte

$$g_m/I_D = [2 \cdot \mu \cdot C_{ox} \cdot W/L / (n \cdot I_D)]^{1/2} \quad (22)$$

Consequentemente, esse parâmetro é usualmente analisado indiretamente em função da tensão de porta, uma vez que ela será a principal responsável por determinar a condição de inversão do dispositivo.

2.6 MOBILIDADE

Dentro do contexto de autoaquecimento e dispositivos avançados, dois diferentes mecanismos de degradação da mobilidade são particularmente relevantes:

a variação com a temperatura e a degradação com o campo elétrico lateral. Apesar de modelos físicos serem capazes de descrever os principais fenômenos que afetam a mobilidade, parâmetros de ajuste empíricos se fazem necessários, como descrito por Caughey e Thomas (1967), Sze (1981), Chain et al. (1997), e Colinge (2004).

Em relação à variação com a temperatura, uma modelagem normalmente adotada é a descrita por Sze (1981) e apresentada na equação (23)

$$\mu = \mu_0.(T_0/T)^c \quad (23)$$

em que μ é a mobilidade medida à temperatura T (em Kelvin), μ_0 é a mobilidade medida à temperatura T_0 (em Kelvin), e c é o coeficiente de degradação da mobilidade com a temperatura. Almeida (2012) reporta que valores usuais do fator c se situam na faixa entre 1,5 e 2,4 em dispositivos de planares e entre 0,8 a 1,2 em dispositivos de múltiplas portas. Além disso, o autor também observa que uma modelagem mais acurada é obtida se for considerada uma relação linear entre o fator c e a temperatura, isto é

$$c = a + b.T \quad (24)$$

em que a e b são parâmetros empíricos e T é a temperatura (em Kelvin). Nascimento (2017) verifica ainda que transistores FinFET tipo n apresentam um menor fator c do que dispositivos p , porém são mais susceptíveis a variações de temperatura.

A extração do fator c pode ser feita rearranjando-se (23) e aplicando logaritmo em ambos os lados da expressão, obtendo

$$c(T) = \log(\mu/\mu_0)/\log(T_0/T) \quad (25)$$

Como o fator c varia em função da temperatura, idealmente ele deve ser extraído em diferentes temperaturas, tomando-se sempre uma mesma como referência (por exemplo $T_0 = 300$ K), de modo a se obter os coeficientes a e b propostos em (24). (24)

Para a análise dos efeitos do campo elétrico transversal, considera-se o modelo adaptado de Colinge (2004)

$$\mu = \mu_0/(1+\theta.V_{GT}) \quad (26)$$

no qual μ_0 é a mobilidade extraída para baixo campo e θ é o fator de degradação da mobilidade por campo elétrico vertical.

Um método recentemente proposto por Henry et al. (2016) utiliza a função $Y = I_D.(g_m)^{-1/2}$ e o inverso da eficiência do transistor como forma de determinar θ e μ_0 , realizando-se o ajuste da expressão

$$I_D/g_m = n.kT/q + Y.\beta^{-1/2}.(1 + \theta.Y.\beta^{-1/2}) \quad (27)$$

sendo n o fator de corpo, k a constante de Boltzmann, T a temperatura, q a carga fundamental e $\beta = W.C_{ox}.\mu_0/L$ o fator de ganho do transistor. Para que a expressão seja válida e resulte na parábola esperada, V_{DS} deve ser mantido constante e em valores baixos (de modo a valer a aproximação linear I_D), enquanto V_{GS} é variado. Um conjunto de curvas experimentais e os respectivos ajustes obtidos através do modelo é apresentado na Figura 2.10.

Figura 2.10 – Características experimentais (linhas cheias) e modeladas (símbolos) da função $I_D/g_m(Y)$ para vários comprimentos de porta entre 28 nm e 1 μm .

Fonte: adaptado de Henry et al. (2016).

Assim, combinando-se ambos os efeitos, obtém-se o modelo a ser empregado ao longo desse trabalho

$$\mu_{eff} = \mu_0.(T_0/T)^c / (1 + \theta.V_{GT}) \quad (28)$$

notando-se que, para esse caso, μ_0 deve ser a mobilidade obtida à temperatura T_0 e também com baixo campo transversal. Assume-se que θ independe da temperatura, e que o uso de V_{DS} baixo no processo de extração de μ_0 não irá provocar mudanças significativas em função do autoaquecimento. Caso a degradação da transcondutância ocorra somente em função do campo elétrico transversal, ou seja, quando a condução se dá pela primeira interface, pode-se substituir μ_{eff} e μ_0 por g_m e g_{m0} em (28), sendo g_m a transcondutância para um valor fixo de sobretensão de condução na temperatura T e g_{m0} o valor na mesma sobretensão à temperatura T_0 .

2.7 CONDUTÂNCIA DE SAÍDA

Pode-se definir a condutância de saída (g_D) de forma diferencial, ou seja

$$g_D \triangleq \partial I_D / \partial V_{DS} \quad (29)$$

para V_{GS} mantido constante. Considerar a corrente de saturação como definida na equação (2) resultaria em uma condutância de saída nula; contudo, se for considerado o acréscimo do efeito de modulação do comprimento de canal como apresentado na equação (13), a condutância se torna positiva e proporcional a λ .

No contexto do autoaquecimento, a importância desse parâmetro está relacionada à observação primeiramente feita por McDaid et al. (1989) de que a condutância de saída na saturação pode se tornar negativa caso o autoaquecimento seja suficientemente intenso. Além disso, como Tenbroek et al. (1996) nota, quando o autoaquecimento é significativo, a condutância de saída se torna uma função da frequência de medida, sendo mais elevada em frequências maiores.

2.8 AUTOAQUECIMENTO

Autoaquecimento é o acúmulo de calor que o dispositivo adquire em seu canal quando está em condições de operação específicas. Devido às dimensões cada vez mais reduzidas e a utilização de novos materiais com baixa condutividade térmica, o desempenho dos transistores avançados é cada vez mais afetado pelo autoaquecimento. Transistores pertencentes à tecnologia SOI são termicamente isolados do substrato através do óxido enterrado. Como resultado, a dissipação do excesso de calor gerado no interior do dispositivo é menos eficiente, o que resulta numa elevação substancial da temperatura do dispositivo, como destacam McDaid et al. (1989) e Jomaah et al. (1995).

Partindo-se do princípio utilizado por Jomaah, Ghibaudo e Balestra (1995) de que o autoaquecimento age de maneira uniforme ao longo do canal com a passagem de corrente, pode-se definir a resistência térmica (R_{th}) de um dispositivo em equilíbrio térmico como sendo

$$R_{th} = (T - T_0) / P \quad (30)$$

em que T_0 é a temperatura inicial do dispositivo, T é a temperatura atingida em função do autoaquecimento e P é a potência elétrica aplicada ao canal do dispositivo, como definida em (31)

$$P = V_{DS}.I_D \quad (31)$$

Uma grande dificuldade em se caracterizar o autoaquecimento através da resistência térmica é a natureza dinâmica do fenômeno. O aumento da temperatura afeta a tensão de limiar no sentido de reduzi-la, como notado pelos valores negativos de dV_t/dT obtidos por Groeseneken et al. (1990). Essa redução resulta em um aumento de corrente, como evidenciado pelas equações (1) e (2), consequentemente aumentando ainda mais a temperatura. Por outro lado, competindo com esse efeito, tem-se a redução da mobilidade com o aumento da temperatura, como evidenciado pela equação (23). Tendo em vista a proporcionalidade direta entre corrente e mobilidade, é evidente que, sob essa perspectiva, o autoaquecimento reduz a corrente total. Combinando-se ambos os efeitos, a degradação da mobilidade com a temperatura é o fator predominante, que resulta na diminuição da corrente de dreno na região de saturação.

Como consequência da dificuldade em se estabelecer analiticamente o ponto de equilíbrio entre aumento da temperatura e degradação de parâmetros, não há atualmente um método capaz de determinar a resistência térmica a partir de medidas em corrente contínua. No passado, Jomaah, Ghibaudo e Balestra (1995) propuseram um método de estimação a partir da análise da condutância de saída e da sensibilidade da corrente à temperatura; entretanto, como os dispositivos ainda apresentavam canal suficientemente longo, o efeito de modulação do comprimento de canal era desprezível, o que permitia atribuir os valores negativos de condutância de saída na saturação imediatamente ao autoaquecimento. Todavia, com a diminuição das dimensões dos transistores, esse fenômeno passa a se tornar significativo, possivelmente se sobrepondo ao autoaquecimento e tornando a condutância de saída positiva, inviabilizando o método como inicialmente proposto.

Para contornar o problema da separação entre modulação do comprimento de canal e autoaquecimento, foram propostas medidas em frequência. Em função da capacidade térmica dos materiais, se o intervalo de aplicação de potência for suficientemente pequeno, o aumento de temperatura será desprezível, o que permite a caracterização do dispositivo livre de autoaquecimento. Atualmente, todos os métodos para caracterização do autoaquecimento são baseados nesse fato; por conseguinte, são realizados ou em corrente alternada, ou em pulsos curtos, com frequências altas e ciclos de trabalho baixos, de forma a evitar o aquecimento dos dispositivos. Makovejev (2012) realiza uma revisão sistemática de diversos métodos

de caracterização em corrente alternada, capazes de estimar tanto a resistência como a capacidade térmica dos dispositivos. Em seu trabalho, fica evidente a dificuldade em caracterizar adequadamente dispositivos de menores dimensões, uma vez que a capacidade térmica é uma propriedade extensiva, obrigando o uso de pulsos curtos (abaixo das dezenas de nanossegundos) ou frequências elevadas (da ordem de GHz) para se eliminar completamente o autoaquecimento. Como consequência, equipamentos especiais capazes de gerar frequências dessa magnitude, bem como estruturas de porta, dreno e fonte especiais para evitar problemas de reflexão e efeitos parasitários se fazem necessários.

Dentre os métodos para estimativa da resistência térmica estudados por Makovejev, dois são os mais relevantes: a caracterização pulsada, ilustrada na Figura 2.11, e a caracterização em radiofrequência (RF), apresentada na Figura 2.12.

Figura 2.11 – (a) Medidas de corrente realizadas com diferentes larguras de pulsos de tensão aplicados; (b) Comparação entre correntes medidas de forma pulsada e de forma contínua; (c) extração da resistência térmica a partir da temperatura estimada do canal

Fonte: adaptado de Makovejev (2012).

A primeira baseia-se na ideia de que, caso sejam aplicados pulsos de tensão suficientemente curtos na porta dos dispositivos, obtém-se uma curva livre de autoaquecimento, como demonstrado ao reduzir-se a largura dos pulsos na Figura 2.11(a). Sabendo-se deste fato, é possível realizar a extração das curvas I_D em função de V_{DS} para diversas temperaturas livres de autoaquecimento, processo este ilustrado pelas curvas pontilhadas na Figura 2.11(b). Em seguida, é feita uma extração em temperatura ambiente em corrente contínua, garantindo que o autoaquecimento tenha tempo de agir totalmente no dispositivo e o equilíbrio térmico seja atingido, tal como feito na curva contínua na Figura 2.11(b). Os pontos de intersecção entre as curvas livres de autoaquecimento e a curva com autoaquecimento fornecem uma estimativa da temperatura no canal do transistor em função da potência aplicada, a partir da qual realiza-se uma regressão linear, de forma a extrair a resistência térmica, como feito na Figura 2.11(c).

Para a caracterização em RF, parte-se da demonstração de Rinaldi (2001), que relaciona parâmetros elétricos em uma rede de duas portas com a sua impedância térmica. Adaptando suas expressões para o caso do transistor MOS, obtemos a expressão

$$R_{th} = \Delta g_D / [\partial I_D / \partial T \cdot (V_{DS} \cdot g_{D,AF} + I_D)] \quad (32)$$

em que Δg_D é a diferença entre a condutância de saída medida em baixas e altas frequências, $g_{D,AF}$ é a condutância de saída em alta frequência, e $\partial I_D / \partial T$ é a sensibilidade da corrente à temperatura aplicada externamente ao dispositivo. As condutâncias podem ser extraídas de um gráfico como da Figura 2.12(a), em que se analisa a variação da condutância de saída em função da frequência; enquanto que a sensibilidade da corrente à temperatura pode ser obtida avaliando-se a taxa de variação no gráfico da corrente de dreno em função da temperatura externa para um dado conjunto de polarizações, como apresentado na Figura 2.12(b). Dessa forma, é possível calcular a resistência térmica de um transistor a partir dos parâmetros extraídos destas curvas.

Figura 2.12 – (a) Caracterização da condutância de saída em função da frequência; (b) determinação da sensibilidade da corrente de dreno à temperatura externa

Fonte: adaptado de Makovejev (2012).

3 MÉTODO DA CONDUTÂNCIA DE SAÍDA MÉDIA

O método da condutância de saída média ($\Delta I_D / \Delta V_{DS}$) foi desenvolvido como uma alternativa para determinar de maneira mais simples se um dispositivo sofre de autoaquecimento. Todos os métodos apresentados por Makovejev (2012) necessitam de estruturas e equipamentos especiais que permitam a execução de medidas em altas frequências, dificultando sua realização, uma vez que uma parcela considerável dos parâmetros de transistores é obtida através de medidas estáticas.

Tendo em vista a recursividade no equacionamento do efeito do autoaquecimento na corrente de transistores, torna-se inviável o desenvolvimento de expressões simplificadas capazes de considerar a resistência térmica nos dispositivos. Contudo, o uso de simulações com poucos fenômenos, como a dependência da mobilidade com a temperatura e o efeito de modulação de canal, permite uma melhor compreensão da ação do autoaquecimento nas curvas características dos transistores. Portanto, uma rotina baseada nas equações (1), (2), (13), (28) e (30) foi criada em MATLAB, com o intuito de identificar possíveis padrões que permitam a caracterização do autoaquecimento. A Figura 3.1 ilustra como foram combinadas as equações de modo a simular o efeito de autoaquecimento.

Figura 3.1 – Fluxograma da rotina de simulação em MATLAB

Fonte: o autor (2017)

3.1 DESENVOLVIMENTO DO MÉTODO E SIMULAÇÕES

Inicialmente, a fim de verificar o funcionamento das simulações, quatro curvas de corrente de dreno em módulo ($|I_D|$) em função da tensão de porta (V_{DS}) foram levantadas para um dispositivo pMOS convencional, cada uma apresentando uma diferente resistência térmica, conforme apresentado na Figura 3.2. Em um primeiro momento, desconsiderou-se o efeito da degradação da mobilidade devido ao campo transversal, isto é, adotou-se $\theta = 0 \text{ V}^{-1}$.

Os valores de R_{th} destas curvas foram escolhidos de forma a ressaltar algumas das características mais relevantes do autoaquecimento dentro do esquema de simulações adotado, não sendo necessariamente condizentes com valores observados experimentalmente. Assim, observa-se que dispositivos que possuam maior R_{th} apresentam menor corrente, e essa discrepância se torna mais significativa conforme se aumenta $|V_{DS}|$. Deve-se notar ainda que para os casos $R_{th} = 50 \text{ K/mW}$ e $R_{th} = 100 \text{ K/mW}$ não é possível detectar a presença do autoaquecimento através de uma análise direta das curvas de corrente, e particularmente para $R_{th} = 100 \text{ K/mW}$, observa-se um equilíbrio entre autoaquecimento e modulação do comprimento de canal.

Figura 3.2 – Corrente de dreno em função da tensão de dreno simulada de um dispositivo pMOS para diferentes resistências térmicas

Fonte: o autor (2017).

Portanto, torna-se evidente a necessidade de um método de detecção, de modo a permitir a identificação de casos nos quais o autoaquecimento é relevante,

porém ainda não suficientemente forte de modo a resultar na condutância de saída negativa. Tendo em vista que essas curvas não trazem imediatamente mais informações para a detecção do fenômeno, fazem-se as curvas $|I_D|$ por V_{GS} com múltiplos valores de V_{DS} da Figura 3.3 para dois diferentes valores de R_{th} .

Figura 3.3 – Corrente de dreno de um dispositivo pMOS simulada em função da tensão de porta para múltiplas tensões de dreno sem resistência térmica e com $R_{th} = 50$ K/mW

Fonte: o autor (2017).

Na situação tradicional, visualizada na Figura 3.3(a), tem-se que curvas com maior $|V_{DS}|$ apresentam maior nível de corrente em triodo, como esperado da equação (1), e também em saturação, em função do efeito de modulação do comprimento de canal. Ao se acrescentar o autoaquecimento nas simulações, como apresentado na Figura 3.3(b), nota-se uma mudança na ordem esperada, com $|I_D|$ da curva $|V_{DS}| = 2$ V menor do que da curva $|V_{DS}| = 1,6$ V, que por sua vez está abaixo da curva $|V_{DS}| = 1,2$ V. Essa mudança de ordem será utilizada como base para o desenvolvimento de um detector do efeito de autoaquecimento, após alguns refinamentos metodológicos, de forma a permitir uma melhor comparação de resultados.

Uma possibilidade para enfatizar a mudança de ordem entre curvas com diferentes V_{DS} é analisar a diferença das correntes, isto é, $I_{D2} - I_{D1} = \Delta I_D$, sendo I_{D2} a corrente a uma polarização V_{D2} e I_{D1} a corrente a uma polarização V_{D1} . Para criar uma normalização entre os resultados obtidos, divide-se a diferença das correntes pela diferença das tensões, isto é, $(I_{D2} - I_{D1}) / (V_{DS2} - V_{DS1}) = \Delta I_D / \Delta V_{DS}$. No caso limite, quando ΔV_{DS} tende a zero, o resultado obtido seria a condutância de saída (g_D),

entretanto, como a princípio a proposta desse método é apenas de detecção, e não de uma estimativa numérica da resistência térmica, serão utilizados passos relativamente largos de V_{DS} (da ordem de 100 mV). Em função dessa aproximação de passos largos, será adotado o nome condutância de saída média ($\Delta I_D / \Delta V_{DS}$) para descrever esse método. A Figura 3.4 introduz as curvas características obtidas através da aplicação desse método às curvas simuladas, com e sem a adição do autoaquecimento.

Figura 3.4 – Condutância de saída média de um dispositivo pMOS simulada em função da tensão de porta para múltiplas tensões de dreno (a) sem resistência térmica e (b) com $R_{th} = 50 \text{ K/mW}$

Fonte: o autor (2017).

Como a região de sublimar não apresenta efeito significativo de autoaquecimento, optou-se pelo uso da sobretensão de condução como abscissa para esses gráficos, enfatizando melhor as regiões de interesse. A comparação entre a Figura 3.4(a) e a Figura 3.4(b) permite estabelecer um primeiro critério na detecção do autoaquecimento: caso $\Delta I_D / \Delta V_{DS}$ se torne negativo, o efeito está presente. Todavia, esse critério por si só não traz inovações, pois, considerando que $\Delta I_D / \Delta V_{DS}$ é uma aproximação de g_D , dizer que seu valor é negativo equivale ao caso clássico de condutância de saída negativa. A fim de verificar se é possível extrair mais informações dessas curvas, verifica-se mais detalhadamente seu comportamento com o transistor operando em saturação na Figura 3.5.

Figura 3.5 – Ênfase na região de saturação das curvas de condutância de saída (a) sem resistência térmica e (b) com $R_{th} = 50 \text{ K/mW}$

Fonte: o autor (2017).

O comportamento parabólico observado na Figura 3.5(a) é facilmente explicado utilizando as expressões de condutância de saída média, corrente em saturação (2), efeito de modulação do comprimento de canal (13), e assumindo que a mobilidade efetiva seja independente da tensão de dreno, conforme demonstrado em (33).

$$\begin{aligned}\Delta I_D / \Delta V_{DS} &= (I_{D2} - I_{D1}) / (V_{DS2} - V_{DS1}) \\ &= \mu_{\text{eff}} \cdot C_{\text{ox}} \cdot (W/L) \cdot V_{GT}^2 / 2n \cdot (1 + \lambda \cdot V_{DS2} - 1 - \lambda \cdot V_{DS1}) / (V_{DS2} - V_{DS1}) \\ &= \mu_{\text{eff}} \cdot C_{\text{ox}} \cdot (W/L) \cdot \lambda \cdot V_{GT}^2 / 2n\end{aligned}\quad (33)$$

A Figura 3.5(b), apesar de não apresentar uma tendência polinomial clara como no caso anterior, possui três pontos característicos dentro da saturação: inflexão (isto é, o ponto em que ocorre a mudança da concavidade, encontrado como sendo um zero da segunda derivada da função), máximo e zero. Estes pontos podem ser utilizados de forma a definir a intensidade com que o autoaquecimento age para certas polarizações; assim, definem-se as regiões de autoaquecimento desprezível (entre $V_{GT} = 0$ e inflexão), fraco (entre inflexão e máximo), moderado (entre máximo e zero), e forte (após zero).

Contudo, deve-se destacar que essas regiões não dizem respeito à temperatura do dispositivo de forma direta, mas sim do quanto o autoaquecimento afeta a condutância de saída do dispositivo em questão. Para demonstrar melhor esse conceito, realizou-se a simulação de dispositivos de mesmas configurações que os anteriores, porém com diferentes valores do parâmetro do efeito de modulação do comprimento de canal (λ), como apresentado na Figura 3.6.

Figura 3.6 – Influência do parâmetro λ na condutância de saída média

Fonte: o autor (2017).

Com o aumento do valor de λ , a tendência parabólica da condutância de saída média antes do ponto de inflexão torna-se mais pronunciada, elevando os valores atingidos pela condutância, conseqüentemente deslocando os três pontos característicos para valores maiores de $|V_{GT}|$ e aumentando os valores dos pontos de inflexão e máximo. Para ilustrar os problemas desse último fato, comparam-se as curvas com $\lambda = 0,1 \text{ V}^{-1}$ e $\lambda = 0,01 \text{ V}^{-1}$. Na primeira, a região de autoaquecimento desprezível se dá para $|V_{GT}| < 0,3 \text{ V}$, enquanto que na segunda ela ocorre para o intervalo $|V_{GT}| < 0,1 \text{ V}$. Entretanto, considerando que incrementos no valor de λ resultam em maiores valores de corrente absoluta e, por conseguinte, em um maior aumento de temperatura devido ao autoaquecimento, seria inicialmente esperado que o dispositivo com $\lambda = 0,1 \text{ V}^{-1}$ apresentasse autoaquecimento significativo para menores valores de $|V_{GT}|$ quando comparado com o dispositivo com $\lambda = 0,01 \text{ V}^{-1}$.

Portanto, apesar de ser possível determinar qual dispositivo possui um maior λ , torna-se inviável a comparação do efeito de autoaquecimento entre dispositivos que possuam diferentes valores de λ . Não obstante, destaca-se a capacidade desse método de ampliar a compreensão da ação do autoaquecimento no funcionamento de transistores MOS, permitindo uma melhor visualização do quanto esse efeito oculta a modulação do comprimento de canal.

3.2 VERIFICAÇÃO EXPERIMENTAL

A fim de verificar se as aproximações realizadas no desenvolvimento teórico e simulações básicas estão de acordo com a realidade, medidas experimentais foram realizadas em alguns dispositivos FinFET de porta tripla.

3.2.1 Características dos dispositivos

Os dispositivos medidos são FinFETs de porta tripla fabricados no imec em substratos SOI de orientação (100), óxido enterrado com espessura de 150 nm, porta composta por uma camada de 100 nm de silício policristalino sobre 5 nm de nitreto de titânio (TiN), altura da aleta (h_{fin}) de 65 nm, canal com dopagem natural da lâmina ($N_a = 1.10^{15} \text{ cm}^{-3}$), dielétrico de porta composto por 2 nm de dióxido de háfnio (HfO_2)

e 1 nm de dióxido de silício (SiO_2), apresentando cinco aletas em paralelo, e possuem diversas configurações de comprimento de canal (L) e largura de aleta (W_{fin}).

3.2.2 Resultados

As curvas de condutância de saída médias obtidas para dois dispositivos de diferentes comprimentos de canal são apresentadas na Figura 3.7.

Figura 3.7 – Ênfase na região de saturação das curvas de condutância de saída de transistores FinFET de dimensões (a) grandes e (b) pequenas

Fonte: o autor (2017).

Além da comparação para transistores com diferentes valores de W e L , e consequentemente diferentes densidades de corrente, outra possibilidade é a

comparação de uma sequência de dispositivos com múltiplos comprimentos de canal, como apresentado na Figura 3.8.

Figura 3.8 – Comparação do efeito de autoaquecimento em dispositivos com três diferentes comprimentos de canal ($L = 130 \text{ nm}$, $L = 250 \text{ nm}$ e $L = 1 \text{ }\mu\text{m}$)

Fonte: o autor (2017).

3.2.3 Análise

O dispositivo de grandes dimensões da Figura 3.7(a) foi escolhido de forma a apresentar níveis de corrente de dreno da mesma ordem de grandeza do dispositivo de pequenas dimensões da Figura 3.7(b), permitindo dessa forma a comparação de resultados em dois diferentes casos de densidade de corrente, uma vez que o autoaquecimento deve ser mais pronunciado em dispositivos com maior densidade. Essas curvas experimentais estão de acordo com os modelos teóricos, ou seja, o dispositivo de grandes dimensões apresenta efeito de autoaquecimento desprezível, enquanto que, para certas polarizações, o dispositivo de menores dimensões possui os três pontos característicos do autoaquecimento.

Entretanto, para curvas com tensões de dreno mais baixas ($V_{DS1} = -1,5 \text{ V}$ e $V_{DS1} = -1,7 \text{ V}$), observa-se a ausência dos pontos de inflexão e máximo esperados, saltando-se imediatamente para a região de autoaquecimento forte mesmo em polarizações de porta baixas, o que não era esperado dentro dos modelos inicialmente adotados. Uma explicação possível, dadas as dimensões do dispositivo e a polarização aplicada a ele, seria a presença dos efeitos de canal curto, trazendo os dois primeiros pontos característicos para valores positivos da sobretensão de

condução. Outra possibilidade é a intensificação do efeito de portadores quentes para essas polarizações de dreno, atingindo o nível de autoaquecimento forte mesmo para baixas correntes.

Na Figura 3.8, as três linhas separadoras indicam os pontos característicos do autoaquecimento no dispositivo com menor comprimento de canal ($L = 130$ nm). Como a corrente é inversamente proporcional ao comprimento, caso sejam mantidas condições de polarização similares entre dispositivos, o autoaquecimento agirá de maneira mais pronunciada nos mais curtos, fato possível de ser verificado nas curvas apresentadas. Os dispositivos de 130 nm e 250 nm apresentam os três pontos e, no caso do dispositivo de 130 nm, o valor mais negativo de condutância média de saída, enquanto que o de 1 μm não apresenta claramente nenhum dos pontos.

Todavia, com o acréscimo do dispositivo de 250 nm nessa análise, nota-se a dificuldade em determinar critérios de comparação entre transistores nas regiões de autoaquecimento fraco e moderado, uma vez que a localização dos pontos característicos é dependente do efeito de modulação do comprimento de canal. Um resultado positivo a ser observado é a concordância entre resultados esperados no que diz respeito ao efeito de modulação do comprimento de canal, isto é, os valores mais elevados de máximo e inflexão do dispositivo de 130 nm apontam para um efeito mais pronunciado, como seria esperado, uma vez que dispositivos mais curtos costumam sofrer mais intensamente com esse efeito.

4 MÉTODO DA ASSINATURA NA EFICIÊNCIA DO TRANSISTOR

Outro parâmetro que pode ser utilizado para a caracterização do autoaquecimento em transistores MOS é a eficiência do transistor (g_m/I_D), apesar de não ser um resultado imediato da forma mais tradicional dessa expressão, apresentada na equação (22). Entretanto, partindo-se de

$$\partial \log(I_D) / \partial V_{GS} = (1/I_D) \cdot (\partial I_D / \partial V_{GS}) = g_m / I_D \quad (34)$$

é possível obter uma relação mais clara com o autoaquecimento. Utilizando a expressão da corrente em saturação da equação (2), e considerando que λ e n independem de V_{GS} , tem-se de (34) que

$$\begin{aligned} g_m / I_D &= \partial \log[\mu \cdot C_{OX} \cdot W/L \cdot V_{GT}^2 \cdot (1 + \lambda V_{DS}) / 2n] / \partial V_{GS} \\ &= \partial \log(\mu) / \partial V_{GS} + \partial \log(V_{GT}^2) / \partial V_{GS} + \partial \log[C_{OX} \cdot W/L \cdot (1 + \lambda V_{DS}) / 2n] / \partial V_{GS} \\ &= \partial \log(\mu) / \partial V_{GS} + \partial \log(V_{GT}^2) / \partial V_{GS} \end{aligned} \quad (35)$$

Expandindo a mobilidade de acordo com a equação (28), considerando a temperatura T da equação (30) e inicialmente desprezando a relação da tensão de limiar com a temperatura, (35) resulta em

$$\begin{aligned} g_m / I_D &= \partial \log(\mu_0 \cdot T_0^c / (1 + \theta \cdot V_{GT})) / \partial V_{GS} - c \cdot \partial \log(R_{th} \cdot P + T_0) / \partial V_{GS} + \partial \log(V_{GT}^2) / \partial V_{GS} \\ &= -1 / (1 + \theta \cdot V_{GT}) \cdot \partial V_{GT} / \partial V_{GS} - c / (1 + R_{th} \cdot P + T_0) \cdot R_{th} \cdot V_{DS} \cdot \partial I_D / \partial V_{GS} + 2 / V_{GT} \\ &= 2 / V_{GT} - \theta / (1 + \theta \cdot V_{GT}) - c \cdot R_{th} \cdot V_{DS} \cdot g_m / (R_{th} \cdot P + T_0) \end{aligned} \quad (36)$$

Evidencia-se assim que, excluído o autoaquecimento, a eficiência do transistor na saturação independe de V_{DS} . Manipulando-se ambos os lados de (36), obtém-se finalmente a equação (37).

$$\begin{aligned} g_m / I_D + c \cdot R_{th} \cdot V_{DS} \cdot g_m / (R_{th} \cdot P + T_0) \cdot I_D / I_D &= 2 / V_{GT} - \theta / (1 + \theta \cdot V_{GT}) \\ \Leftrightarrow g_m / I_D \cdot [(1 + c) \cdot R_{th} \cdot P + T_0] / (R_{th} \cdot P + T_0) &= (-\theta \cdot V_{GT} + 2 + 2 \cdot \theta \cdot V_{GT}) / [V_{GT} \cdot (1 + \theta \cdot V_{GT})] \\ \Leftrightarrow g_m / I_D &= \{(R_{th} \cdot P + T_0) / [(1 + c) \cdot R_{th} \cdot P + T_0]\} \cdot \{1 / V_{GT} + 1 / [V_{GT} \cdot (1 + \theta \cdot V_{GT})]\} \end{aligned} \quad (37)$$

De modo a simplificar a expressão, o termo relacionado ao autoaquecimento será representado pela equação (38), resultando na expressão reduzida da eficiência do transistor em (39).

$$F = (R_{th} \cdot P + T_0) / [(1 + c) \cdot R_{th} \cdot P + T_0] \quad (38)$$

$$g_m / I_D = F \cdot \{1 / V_{GT} + 1 / [V_{GT} \cdot (1 + \theta \cdot V_{GT})]\} \quad (39)$$

Dessa forma, nota-se claramente a influência do autoaquecimento na eficiência do transistor, considerando-se sua ação na mobilidade. F possui dependência com a tensão de dreno, dada a relação de potência da equação (31), e com o autoaquecimento, como evidenciado pela resistência térmica, enquanto que os termos

restantes da equação (39) não apresentam relação imediata com a tensão de dreno; ou seja, se o autoaquecimento for desprezível, a eficiência do transistor independe da tensão de dreno.

Para melhor compreender o comportamento dessa expressão, simulações numéricas e medidas experimentais utilizando os mesmos dispositivos da seção anterior foram realizadas.

4.1 DESENVOLVIMENTO DO MÉTODO E SIMULAÇÕES

Tendo em vista a relação (37), uma variável que poderia ser considerada como abscissa para os gráficos de eficiência do transistor seria a potência dissipada no canal conforme se aumenta em módulo a tensão de dreno. Entretanto, considerando que pequenas quedas na corrente de dreno devido ao autoaquecimento podem ser ocultadas pelo aumento da tensão de dreno, torna-se mais interessante o uso direto da corrente, variada unicamente através de incrementos (em módulo) da tensão de dreno, com tensão de porta fixa. A Figura 4.1 apresenta os resultados obtidos para simulações feitas em condições similares às da seção 3.2.1, acrescentando-se o efeito de degradação da mobilidade pelo campo elétrico transversal, com e sem autoaquecimento e três diferentes condições de polarização de porta.

Figura 4.1 – Eficiência do transistor em função da corrente de dreno sem autoaquecimento (linhas tracejadas) e com (linhas cheias) para múltiplas sobretensões de condução, com destaque para a região de saturação

Fonte: o autor (2017).

Comparando-se em cada um dos três casos a curva com autoaquecimento e a sem, nota-se claramente a evolução do efeito, reduzindo a corrente total e degradando mais a eficiência do transistor conforme a sobretensão de condução é aumentada. Em particular, na região de saturação, nota-se o deslocamento do patamar, chegando a ocorrer a inversão do seu sentido caso o efeito seja suficientemente pronunciado. Para melhor esclarecer os desdobramentos possíveis do efeito e tentar identificar e classificar certos comportamentos, elaborou-se a Figura 4.2.

Figura 4.2 – Esquematização da influência do autoaquecimento na eficiência do transistor

Fonte: o autor (2017).

A partir da expressão (38), e levando-se em conta a interação entre autoaquecimento e corrente de saturação quando se varia a tensão de dreno, é possível explicar cada uma das cinco situações apresentadas na Figura 4.2.

Primeiramente, na situação (1), desprezando-se o autoaquecimento (tanto por resistência térmica ou potência aplicada desprezível), tem-se a formação de um patamar devido à independência da eficiência com $|V_{DS}|$ e ao aumento de corrente pelo efeito de modulação do comprimento de canal ao aumentar $|V_{DS}|$ em saturação. Caso esse último efeito também seja desprezível, o patamar de saturação seria reduzido a único ponto, tendo em vista que nem a corrente, nem a eficiência do transistor seriam afetadas por variações em $|V_{DS}|$.

Considerando agora dispositivos cuja resistência térmica não seja desprezível, a intensidade do autoaquecimento nessas curvas pode ser relativamente controlada através de $|V_{GT}|$. Apesar de não ser uma relação de proporcionalidade direta, o aumento de $|V_{GT}|$ irá aumentar a corrente, consequentemente aumentando a

temperatura do dispositivo. Nota-se na expressão (38) que o aumento do produto $R_{th}.P$ faz com que F diminua, causando assim a degradação do patamar da eficiência do transistor, o que justifica o comportamento observado em (2), no qual se classifica o efeito do autoaquecimento como fraco.

Caso a potência aplicada seja aumentada ainda mais através de $|V_{GT}|$, o efeito do autoaquecimento pode ser suficientemente intenso ao ponto de se cancelar com o efeito de modulação do comprimento de canal, ou seja, fazendo com que a corrente em saturação permaneça inalterada com o aumento de $|V_{DS}|$. Esse fato não compromete a diminuição de F pois, apesar da corrente permanecer constante, a potência aplicada continua aumentando com $|V_{DS}|$, resultando, portanto, na situação ilustrada em (3), em que se classifica o efeito de autoaquecimento como moderado.

Novos incrementos em $|V_{GT}|$ após o ponto de equilíbrio entre efeitos resultarão na diminuição da corrente de dreno com o aumento de $|V_{DS}|$. Essa diminuição por sua vez irá resultar na inversão de sentido da eficiência do transistor observada em (4) e (5). Por ambas representarem as situações em que se tem condutância de saída negativa, o efeito de autoaquecimento nelas é classificado como forte.

Destaca-se ainda que a distinção entre as situações (4) e (5) é feita devido ao fato de (5) representar uma situação limite. Aumentar indefinidamente o produto $R_{th}.P$ na expressão (38) fará com que F tenda a $1/(1+c)$ de maneira assintótica, ou seja, quanto mais se aumenta $|V_{GT}|$, maior será o intervalo de $|V_{DS}|$ necessário para que seja possível observar alguma mudança significativa em F e, portanto, na eficiência do transistor.

Uma verificação necessária para o uso desse método é o seu uso para comparação entre dispositivos. Como visto na seção 3.1, o efeito de modulação do comprimento de canal é crítico para a comparação do autoaquecimento entre dispositivos quando se utiliza o método da condutância de saída média, uma vez que afeta diretamente o comportamento da corrente em saturação. Assim, os gráficos da Figura 4.3 foram elaborados a partir de simulações de dispositivos similares, alterando-se somente a intensidade do efeito de modulação do comprimento de canal através do parâmetro λ , de forma a testar exclusivamente a ação desse efeito nas comparações.

Figura 4.3 – Comparação do método da assinatura na eficiência do transistor entre dispositivos com diferentes λ , $R_{th} = 50 \text{ K/mW}$ e (a) baixo $|V_{GT}|$, (b) $|V_{GT}|$ intermediário e (c) $|V_{GT}|$ alto

Fonte: o autor (2017).

Na Figura 4.3(a), nota-se que se λ for pequeno, mesmo para baixas polarizações, configura-se imediatamente a situação de autoaquecimento moderado, sem a passagem através das regiões de efeito desprezível e fraco. Esse fato evidencia que as classificações estabelecidas não dizem respeito exclusivamente à temperatura do dispositivo, mas sim às interações entre temperatura e modulação do comprimento de canal. Entretanto, da equação (37) nota-se que a influência de λ na eficiência do transistor é pequena, ou seja, os problemas na classificação surgem em função do uso da corrente de dreno como abscissa. Essa conclusão é reforçada em todos os gráficos da Figura 4.3, pois neles a variação de g_m/I_D na saturação para os diferentes valores de λ permanece bastante próxima. Por conseguinte, para análises comparativas, torna-se interessante o uso da tensão de dreno como abscissa dos gráficos, conforme ilustrado na Figura 4.4.

Como previsto pelas equações, a variação no parâmetro da modulação do comprimento de canal não afeta a eficiência do transistor caso o autoaquecimento seja nulo ($R_{th} = 0 \text{ K/mW}$). Todavia, ao inserir a resistência térmica, a potência aplicada passa a ser significativa na degradação da eficiência do transistor, portanto, dispositivos que apresentam maior λ terão maior potência e, consequentemente, maior degradação da eficiência do transistor devido ao autoaquecimento.

Figura 4.4 – Comparação da eficiência do transistor em função da tensão de dreno para dispositivos com diferentes resistências térmicas e λ

Fonte: o autor (2017).

Assim, cria-se uma adaptação ao método da assinatura na eficiência do transistor, permitindo o seu uso como ferramenta comparativa direta entre dispositivos. Entretanto, essa adaptação não substitui completamente o uso da corrente como abscissa, uma vez que ela não permite uma análise da interação entre autoaquecimento e modulação do comprimento de canal.

4.2 VERIFICAÇÃO EXPERIMENTAL

Medidas experimentais foram realizadas com o intuito de verificar se as considerações e aproximações realizadas ao longo do desenvolvimento do método são válidas também para dispositivos avançados. Todos os resultados foram obtidos utilizando transistores FinFET de porta tripla como os descritos na seção 3.2.1.

4.2.1 Resultados

Da Figura 4.5 a Figura 4.7, aplica-se o método da assinatura na eficiência do transistor em dispositivos de mesma largura de aleta ($W_{fin} = 15$ nm), mesma altura da aleta ($h_{fin} = 65$ nm), três diferentes comprimentos de canal ($L = 1$ μ m na Figura 4.5, $L = 250$ nm na Figura 4.6 e $L = 250$ nm na Figura 4.7) e polarização entre dreno e fonte variando de 0 V a -2 V. Como os comprimentos de canal são variados, a corrente

de dreno nas abscissas é normalizada por esse parâmetro, de forma a melhorar a comparação entre resultados.

Figura 4.5 – Verificação experimental da assinatura na eficiência do transistor em dispositivo com $L = 1 \mu\text{m}$, $W_{\text{eff}} = 145 \text{ nm}$ e (a) $|V_{\text{GT}}| = 0,1 \text{ V}$, (b) $|V_{\text{GT}}| = 0,85 \text{ V}$ e (c) $|V_{\text{GT}}| = 1,2 \text{ V}$

Fonte: o autor (2017).

Figura 4.6 – Verificação experimental da assinatura na eficiência do transistor em dispositivo com $L = 250 \text{ nm}$, $W_{\text{eff}} = 145 \text{ nm}$ e (a) $|V_{\text{GT}}| = 0,1 \text{ V}$, (b) $|V_{\text{GT}}| = 0,45 \text{ V}$ e (c) $|V_{\text{GT}}| = 1,2 \text{ V}$

Fonte: o autor (2017).

Figura 4.7 – Verificação experimental da assinatura na eficiência do transistor em dispositivo com $L = 130$ nm, $W_{\text{eff}} = 145$ nm e (a) $|V_{\text{GT}}| = 0,1$ V, (b) $|V_{\text{GT}}| = 0,45$ V e (c) $|V_{\text{GT}}| = 1,2$ V

Fonte: o autor (2017).

Aplica-se ainda o método modificado da assinatura na eficiência do transistor para verificar o seu desempenho em comparações entre os três dispositivos utilizados acima, como ilustrado na Figura 4.8.

Figura 4.8 – Aplicação experimental do método eficiência do transistor em função da tensão de dreno para dispositivos com diferentes comprimentos de canal

Fonte: o autor (2017).

4.2.2 Análise

A Figura 4.5, a Figura 4.6 e a Figura 4.7 foram estruturadas de forma a apresentar três dentre os cinco casos discutidos na Figura 4.2. No item (a) das três figuras, tem-se uma situação de fraco autoaquecimento (caso (2)), como esperado ao se aplicar uma sobretensão de condução suficientemente pequena.

O item (b) das figuras foi configurado de forma a apresentar o caso (3), exigindo dessa forma a aplicação de diferentes tensões. Nos dispositivos de 130 nm e 250 nm, para que houvesse o balanço entre modulação do comprimento de canal e degradação da corrente devido ao autoaquecimento, bastou a aplicação de $|V_{GT}| = 0,45$ V, enquanto que no dispositivo de 1 μm foi necessário um valor mais elevado ($|V_{GT}| = 0,85$ V). O autoaquecimento em um dispositivo mais longo é consequência imediata da menor densidade de corrente, sendo necessária uma maior sobretensão de porta para que se atinja um efeito equiparável aos dispositivos menores. Assim, a necessidade de uma maior potência aplicada para se atingir a região de autoaquecimento moderado parece razoável.

Por fim, verificou-se no item (c) que a aplicação de $|V_{GT}| = 1,2$ V nos três dispositivos foi o bastante para que todos apresentassem condição próxima ao caso (5). Apesar desse fato, destaca-se a diferença entre as quedas de correntes dos dispositivos: tem-se uma queda de pouco menos que 1 $\mu\text{A} \cdot \mu\text{m}$ para o dispositivo de 1 μm , de aproximadamente 3 $\mu\text{A} \cdot \mu\text{m}$ para o dispositivo de 250 nm e cerca de 2 $\mu\text{A} \cdot \mu\text{m}$ para o dispositivo de 130 nm. Inicialmente seria esperado que dispositivos com menor comprimento de canal sofressem mais intensamente com o autoaquecimento, uma vez que apresentam maior densidade de corrente, fato esse observado quando se compara a corrente normalizada do dispositivo de 1 μm com os dois mais curtos. Contudo, o transistor de 250 nm apresenta uma queda ligeiramente maior da corrente do que o de 130 nm, ao contrário do que seria esperado em uma análise superficial. A explicação para esse fato está relacionada ao que foi observado na Figura 4.3, ou seja, o efeito de modulação do comprimento de canal pode ocultar parcialmente a queda de corrente relacionada ao autoaquecimento, dificultando a comparação direta entre dispositivos, como previsto na análise teórica.

Com o auxílio da Figura 4.8, torna-se finalmente possível realizar a comparação imediata entre os três dispositivos, concluindo-se claramente que, quanto menor o

comprimento de canal, maior será a degradação da eficiência do transistor na região de saturação, ou seja, maior será o efeito de autoaquecimento.

Portanto, apesar de nem todos os efeitos de segunda ordem (como por exemplo a variação da tensão de limiar com a temperatura) terem sido considerados durante o processo de modelagem do fenômeno, os resultados obtidos experimentalmente apresentam boa concordância com as previsões teóricas. Consequentemente, tendo em vista os bons resultados obtidos com o método da assinatura na eficiência do transistor, ele será adotado para o estudo de dispositivos avançados individualmente, e a sua versão modificada será empregada para a comparação entre dispositivos de diferentes tecnologias.

5 ESTIMATIVA DA RESISTÊNCIA TÉRMICA UTILIZANDO MEDIDAS EM CORRENTE CONTÍNUA

Apesar de o método da assinatura na eficiência do transistor fornecer uma ferramenta para comparação qualitativa do efeito de autoaquecimento entre diferentes dispositivos, seria de grande interesse a obtenção de uma estimativa numérica da resistência térmica, de modo a permitir uma aproximação da temperatura interna atingida no canal do dispositivo em operação, consequentemente possibilitando a correção de parâmetros elétricos extraídos, principalmente na operação em saturação. Portanto, neste capítulo será estudada a correlação entre a expressão da eficiência do transistor e a resistência térmica. Para verificar os resultados teóricos produzidos, serão geradas simulações numéricas com o software Sentaurus Device, da Synopsys.

5.1 LINEARIZAÇÃO DO INVERSO DA EFICIÊNCIA DO TRANSISTOR

Iniciando a análise pela função F (38), uma vez que ela representa diretamente o quanto a eficiência do transistor degrada em função do autoaquecimento, claramente não há linearidade do ponto de vista analítico. Entretanto, definindo uma função G como sendo o inverso de F , obtém-se

$$G(P) = 1/F = [(1 + c) \cdot R_{th} \cdot P + T_0] / (R_{th} \cdot P + T_0) = 1 + c \cdot R_{th} \cdot P / (R_{th} \cdot P + T_0) \quad (40)$$

que também não é, a rigor, linear. Entretanto, para uma região em que $R_{th} \cdot P$ seja suficientemente menor do que T_0 , uma aproximação linear aparenta ser viável.

Para testar essa hipótese, avalia-se a série de Taylor de ordem m (denotada como T_m) desta expressão em torno de um certo ponto de potência P_0

$$G(P) \approx T_m(P) = G(P_0) + \sum_{i=1}^m [G^{(i)}(P_0) \cdot (P - P_0)^i / i!]$$

em que $G^{(i)}$ é a i -ésima derivada de G , que será dada por

$$G^{(i)}(P) = (-1)^{i+1} \cdot c \cdot R_{th}^i \cdot T_0 / (R_{th} \cdot P + T_0)^{i+1}$$

Pode-se descrever então o coeficiente de P^m como sendo o m -ésimo termo de uma progressão geométrica de razão $r = -R_{th} \cdot P / (R_{th} \cdot P_0 + T_0)$ e termo inicial $a_0 = -c \cdot T_0 / (R_{th} \cdot P_0 + T_0)$ dividido por $m!$. Tomando $T_0 = 300$ K como temperatura de referência, considerando c de ordem unitária, e considerando que $R_{th} \cdot P_0$ e $R_{th} \cdot P$ sejam valores da ordem de 100 K, tem-se as aproximações

$$r = -R_{th} \cdot P / (R_{th} \cdot P_0 + T_0) \approx -100/400 = -0,25$$

$$a_0 = -c.T_0/(R_{th}.P_0 + T_0) \approx -300/400 = -0,75$$

Ou seja, $|a_1/1!|$ é da ordem de 0,1875, enquanto que $|a_2/2!|$ é da ordem de 0,02348, quase uma ordem de grandeza abaixo, justificando assim o uso de uma aproximação linear neste contexto. Portanto, desenvolvendo para $m = 1$

$$\begin{aligned} G(P) &\approx T_1(P) = G(P_0) + G^{(1)}(P_0).(P - P_0)^{1/1!} \\ &= 1 + c.R_{th}.P_0/(R_{th}.P_0 + T_0) + [c.R_{th}.(R_{th}.P_0 + T_0) - c.R_{th}.P_0.R_{th}]/(R_{th}.P_0 + T_0)^2.(P - P_0) \\ &= 1 + c.R_{th}.P_0/(R_{th}.P_0 + T_0) + c.R_{th}.T_0/(R_{th}.P_0 + T_0)^2.(P - P_0) \\ &= 1 + c.[R_{th}.P_0/(R_{th}.P_0 + T_0)]^2 + c.R_{th}.T_0.P/(R_{th}.P_0 + T_0)^2 \end{aligned} \quad (41)$$

Repetindo-se a análise numérica para avaliar a influência do termo constante na expressão (41), nota-se que

$$1 + c.[R_{th}.P_0/(R_{th}.P_0 + T_0)]^2 \approx 1 + (100/400)^2 = 1 + 0,0625 \approx 1$$

Considerando-se que é buscada uma aproximação de primeira ordem para a estimativa do fenômeno, será adotado que o termo independente da expressão é 1. Assim, obtém-se

$$1/F = G(P) \approx 1 + c.R_{th}.T_0.P/(R_{th}.P_0 + T_0)^2 \quad (42)$$

Além da análise da função F , é necessário avaliar também o termo dependente de V_{GT} em (39). Entretanto, como é possível variar P mantendo V_{GT} constante (desprezando-se variações na tensão de limiar em função da temperatura), e considerando que θ independe da temperatura de operação do dispositivo, pode-se assumir que todo o fator dependente de V_{GT} em (39) permanece constante caso seja variada somente a tensão aplicada ao dreno. Assim, para um V_{GT} fixo,

$$\{1/V_{GT} + 1/[V_{GT}.(1+\theta.V_{GT})]\}^{-1} = M \quad (43)$$

Invertendo a eficiência do transistor representada em (39), fica

$$I_D/g_m = M/F = M.G \approx M.[1 + c.R_{th}.T_0.P/(R_{th}.P_0 + T_0)^2] \quad (44)$$

em que I_D/g_m é função somente de P , e depende dos parâmetros c , M e R_{th} . P_0 será escolhido de forma a otimizar a aproximação linear, portanto, será adotado como sendo a média dos valores de P para os quais I_D/g_m melhor se aproxima do comportamento linear, enquanto que T_0 será tomada como a temperatura ambiente, isto é, 300 K.

5.2 EXTRAÇÃO DOS PARÂMETROS C, M E R_{th}

Somente com as informações contidas na equação (44), ainda não é possível determinar os três parâmetros desejados. Entretanto, utilizando-se os métodos de extração estudados na seção de revisão teórica, combinados com algumas aproximações numéricas, podem-se obter estimativas para os valores de c, M e R_{th} .

Primeiramente, para calcular M, uma opção é extrair o valor da tensão de limiar em saturação (utilizando, por exemplo, o método descrito na seção 2.3), em seguida obter a constante θ (como descrito na seção 2.6), e finalmente aplicar ambos os valores na equação (43), obtendo imediatamente uma estimativa para este parâmetro. Contudo, se for considerado que ambas as extrações podem introduzir fontes de erro adicionais, uma vez que ambas desconsideram a interação do autoaquecimento com o procedimento, pode ser de maior interesse utilizar uma aproximação direta, algo que é possível graças ao termo constante da equação (44). O termo independente obtido ao fazer o ajuste linear de I_D/g_m em função de P será uma aproximação razoável da constante M, que pode ser imediatamente extraída para o mesmo valor de V_{GT} a ser utilizado durante as próximas etapas.

Em seguida, como a expressão deduzida para I_D/g_m é válida somente para o transistor operando em saturação, é necessária a extração do fator c em valores de tensão de porta e dreno próximos aos utilizados para estimativa da resistência térmica. Dessa forma, tem-se um problema para a extração deste parâmetro a partir da razão entre mobilidades em diferentes temperaturas. Consequentemente, como sugerido na seção 2.6, pode-se utilizar o método de extração a partir de g_m com V_{GT} constante, uma vez que, em saturação, a definição do ponto de máximo em g_m não é clara, e não garante um valor diretamente proporcional à mobilidade. Além disso, os altos valores de V_{GT} e V_{DS} empregados implicam na impossibilidade de desprezar o efeito de autoaquecimento, ou seja, seria necessário realizar uma correção nos valores T e T_0 utilizados. Como resultado, conclui-se que a fórmula final para extração do fator c será dada por

$$c = \log(g_m(V_{GT,T})/g_{m0}(V_{GT,T0}))/\log((T_0 + R_{th}.P_{T0})/(T + R_{th}.P_T)) \quad (45)$$

em que $g_m(V_{GT,T})$ é a transcondutância calculada para V_{GT} na temperatura T, $g_{m0}(V_{GT,T0})$ é a transcondutância calculada para V_{GT} na temperatura T_0 , P_{T0} é a potência consumida pelo dispositivo na temperatura T_0 e com tensão $V_{GT,T0}$ aplicada, e P_T é a potência consumida pelo dispositivo na temperatura T e com tensão $V_{GT,T}$

aplicada. Deve-se notar que, apesar de $V_{GT,T}$ precisar ser igual a $V_{GT,T0}$ para que a aproximação seja válida, é interessante o uso de uma notação diferenciada, uma vez que V_t é função da temperatura, como observado em (19); por conseguinte, serão necessários valores distintos de V_{GS} para que a igualdade seja mantida constante, justificando assim essa notação. A equação (45) não pode ser diretamente resolvida ainda, uma vez que o valor de R_{th} ainda é desconhecido; entretanto, considerando a equação (44), temos um sistema com duas equações e somente dois parâmetros para serem ajustados, ou seja, o ajuste é possível.

Tendo em vista que a inserção de (45) em (44) resultará em uma função transcendental, não é possível obter uma expressão analítica para a extração da resistência térmica. Além desse problema, c possui uma dependência com os valores de temperatura utilizados, então o correto seria o uso de um valor $T + R_{th} \cdot P_T$ que melhor se aproximasse da temperatura real de operação do dispositivo. Assim, para a extração de c e R_{th} quando o autoaquecimento é significativo, propõe-se a seguinte metodologia numérica:

1. I_D em função de V_{DS} para um pequeno conjunto de V_{GS} próximos, permitindo a extração da I_D/g_m em função de P , com V_{GS} fixo e V_{DS} variando;
2. Obter curvas I_D em função de V_{GS} com $V_{DS,sat}$ fixo de modo a garantir operação em saturação para N temperaturas;
3. Extração da tensão de limiar em saturação e em função da temperatura;
4. Inicialização das grandezas a serem estimadas numericamente, c e R_{th} . Criar também uma variável auxiliar, R_{th0} , para determinação do fim do processo numérico;
5. Determinar o valor de V_{GT0} a ser mantido como referência durante o processo, condizente com os valores de V_{GS} e V_{DS} ;
6. Determinar o valor da potência $P_T = I_D(V_{GT0}, T) \cdot V_{DS,sat}$ para cada uma das N temperaturas;
7. Iniciar um laço a ser rodado enquanto $(R_{th} - R_{th0})/R_{th}$ for maior do que o erro desejado;
8. Atualizar R_{th0} com o valor de R_{th} ;
9. Estimar a temperatura real $T_{est} = T + R_{th} \cdot P_T$ para cada uma das N temperaturas;
10. Calcular c utilizando (45) para $N - 1$ temperaturas T_{est} , tomando uma delas como a temperatura T_{est0} de referência;

11. A partir de todas T_{est} , realizar uma regressão linear para estimar o comportamento de V_t com a temperatura, isto é, $V_t = a.T_{est} + b$;
12. Calcular a temperatura estimada média $T_{est,SHE} = T_0 + R_{th}.P_{med}$, sendo P_{med} a potência média dentro do intervalo de V_{DS} que garante regime em saturação;
13. Calcular $V_{GT,SHE} = V_{GS} - (a.T_{est,SHE} + b)$;
14. Determinar I_D/g_m e $P = V_{DS}.I_D$ para todos os pontos de operação em saturação;
15. Localizar o valor de T_{est} que melhor se aproxima de $T_{est,SHE}$ para obter c ;
16. Estimar M a partir da regressão linear de I_D/g_m em saturação;
17. Realizar o ajuste dos pontos em saturação de I_D/g_m ao modelo em função de P proposto em (44) utilizando os valores de c e M estimados para obter um novo valor de R_{th} ;
18. Retornar ao passo 7 caso a diferença entre R_{th} e R_{th0} não seja aceitável.

A Figura 5.1 apresenta um fluxograma resumido, contendo os principais passos apresentados na lista anterior.

Figura 5.1 – Fluxograma do método para estimativa da resistência térmica

5.3 SIMULAÇÕES NUMÉRICAS

A fim de verificar o funcionamento do método, será utilizada uma estrutura FinFET, inteiramente simulada no Sentaurus, da Synopsys. Primeiramente, é construída uma estrutura tridimensional com dimensões próximas as dos dispositivos experimentais estudados nos capítulos anteriores, isto é, $h_{fin} = 60$ nm e $W_{fin} = 20$ nm, L de 150 nm, 300 nm e 500 nm, óxido enterrado (t_{oxb}) de 150 nm, espessura equivalente do óxido de porta como 1,2 nm, dopagens de boro com concentração 10^{15} cm⁻³ no canal e de arsênio com concentração 10^{20} cm⁻³ na fonte e no dreno. Os contatos elétricos foram feitos de forma superficial ao redor da aleta no caso do contato de porta, abaixo do óxido enterrado para o contato de substrato, e nas extremidades da aleta para os contatos de fonte e dreno. Na Figura 5.2 exibe-se a estrutura gerada a partir dessas diretrizes, com os contatos representados como retângulos de borda rosa.

Já sob o ponto de vista de modelos físicos empregados, destaca-se que a mobilidade é o parâmetro que apresenta maior dependência com a temperatura, e, portanto, é crucial dentro da análise do autoaquecimento. Assim, será adotado um dos modelos mais completos da biblioteca do Sentaurus, o modelo de mobilidade unificado de Philips (*PhuMob*). Além disso, a influência do campo elétrico transversal na mobilidade foi incluída através do comando *Enormal*, representando de forma mais precisa o fenômeno que foi aproximado na equação (26). Finalmente, acrescenta-se aqui o uso da saturação da mobilidade devido ao alto campo (*HighFieldSaturation*), com a finalidade de verificar se o fato de ele não ter sido contabilizado nos modelos de primeira ordem pode gerar discrepâncias relevantes.

Conjuntamente aos modelos de mobilidade, alguns outros foram considerados em função das suas dependências com a temperatura, como o estreitamento da banda proibida (*EffectiveIntrinsicDensity*) e os efeitos de recombinação (*Recombination*). Como altas dopagens foram utilizadas nas regiões de fonte e dreno, faz-se necessário acrescentar o modelo estatístico de Fermi (*Fermi*) dentro da seção de modelos físicos, garantindo o funcionamento adequado das simulações nas regiões de fonte e dreno.

Figura 5.2 – Estrutura gerada para o transistor FinFET (a) tridimensional (b) corte evidenciando W_{fin} (c) corte evidenciando L (d) representação esquemática das dimensões

Fonte: o autor (2017).

Para a ativação do efeito de autoaquecimento, foi utilizado o modelo termodinâmico (*Thermodynamic*) em conjunto com o comando *Temperature* durante a etapa de resolução das equações, uma vez que ele acrescenta o gradiente de temperatura às expressões de densidade de corrente. Para suprimir o autoaquecimento, basta remover o comando *Temperature* dos termos a serem resolvidos. O acréscimo desse modelo também implica na declaração de um componente adicional na declaração do dispositivo, os contatos térmicos (*Thermodes*). No caso das simulações aqui realizadas, foram escolhidos os próprios

contatos elétricos de fonte, porta, dreno e substrato para atuarem como contatos térmicos, com sua temperatura fixa e igual a ambiente, e com resistência térmica de $10^{-4} \text{ cm}^2.\text{K/W}$.

5.4 RESULTADOS

Primeiramente, como não há um parâmetro de ajuste exato da resistência térmica de cada transistor, uma vez que ela é consequência das propriedades físicas dos materiais utilizados, faz-se uma extração baseada em um dos métodos já consolidados na literatura. Tendo em vista que é possível simplesmente desconsiderar o efeito de autoaquecimento nas simulações, elimina-se o maior problema das medidas pulsadas, que é a incerteza se os pulsos aplicados são suficientemente curtos para se obter uma corrente livre de autoaquecimento. Assim, a partir da intersecção das curvas de corrente sem autoaquecimento em múltiplas temperaturas com uma única curva com autoaquecimento a temperatura ambiente, pode-se estimar a temperatura do dispositivo para uma dada potência, como ilustrado na Figura 5.3.

Figura 5.3 – Estimativa da temperatura no canal a partir das curvas sem autoaquecimento em múltiplas temperaturas e uma curva com autoaquecimento a temperatura ambiente para (a) $L = 150 \text{ nm}$, (b) $L = 300 \text{ nm}$ e (c) $L = 500 \text{ nm}$

(c)

Fonte: o autor (2017).

A partir da relação entre potência e temperatura, extrai-se a resistência térmica real de cada um dos dispositivos, como demonstrado na Figura 5.4.

Figura 5.4 – Extração da resistência térmica a partir da temperatura estimada no canal em função da potência aplicada ao dreno

Fonte: o autor (2017).

A Figura 5.5 compara os resultados da resistência térmica extraída segundo uma das metodologias tradicionais com os valores obtidos ao seguir os procedimentos propostos na seção 5.2.

Figura 5.5 – Comparação das resistências térmicas extraídas a partir do método proposto com o método tradicional pulsado em função do comprimento de canal do dispositivo FinFET

Fonte: o autor (2017).

Para reforçar a maior importância da ordem de grandeza, verificou-se a diferença entre as temperaturas estimadas a partir da equação (31) na condição de máxima potência dissipada nas simulações realizadas, isto é, quando são aplicadas as tensões $V_{DS} = 2,0$ V e $V_{GS} = 1,5$ V.

Figura 5.6 – Comparação das temperaturas estimadas no canal dos dispositivos através do método tradicional pulsado e do método proposto para diferentes comprimentos de canal

Fonte: o autor (2017)

5.5 DISCUSSÕES

Considerando a natureza do fenômeno, a ordem de grandeza da resistência térmica fica bem determinada, bastante similar àquela obtida por Makovejev (2012) após a realização da normalização em função da largura efetiva do dispositivo, o que se traduz em uma boa estimativa da temperatura no canal dos dispositivos. Os resultados também apresentam uma tendência coerente com o esperado do ponto de vista físico, isto é, dispositivos com maior comprimento de canal apresentam uma menor resistência térmica, e, conseqüentemente, um menor aumento de temperatura. Ademais, uma grande vantagem deste método é a sua capacidade de implementação em qualquer sistema capaz de realizar medidas em corrente contínua, ao contrário de todos os métodos existentes até o momento. Portanto, considerando que há uma consistência entre os valores extraídos e que o pior erro cometido na estimativa de temperatura está na faixa de 5,6%, este método também será útil para a comparação de diferentes dispositivos.

6 COMPARAÇÃO ENTRE DISPOSITIVOS FINFET E SOI UTBB

Finalmente, utilizando os métodos propostos ao longo do trabalho, será realizada uma análise comparativa entre as estruturas FinFET e SOI UTBB. Para tanto, serão utilizadas simulações tridimensionais realizadas no software Sentaurus Device, da Synopsys, para dispositivos de ambas as estruturas. Para que seja possível estabelecer critérios de comparação, alguns compromissos terão que ser feitos, uma vez que comparar o efeito de autoaquecimento em dispositivos que não estejam operando em condições ideais não faria sentido. Por exemplo, comparar um transistor UTBB otimizado com um FinFET de W_{fin} largo apenas para que se tenham dispositivos com mesma área de contato não seria efetivo, uma vez que este FinFET apresentaria um pior acoplamento eletrostático entre as portas, resultando em um dispositivo de baixo desempenho. Assim, a comparação será realizada para transistores de correntes similares, resultando em efeitos de autoaquecimento de mesma magnitude caso as resistências térmicas sejam equivalentes. Nesta etapa serão utilizados dispositivos de canal tipo n, a fim de reafirmar a validade dos métodos desenvolvidos independentemente das dopagens utilizadas.

6.1 SIMULAÇÃO DOS DISPOSITIVOS

Tendo em vista o critério de comparação estabelecido, uma aproximação inicial para o problema é realizar o projeto das estruturas com mesmo comprimento de canal ($L_{UTBB} = L_{FinFET}$), e largura equivalente do FinFET similar à largura total do UTBB ($W_{eff,FinFET} = W_{UTBB}$). Tendo em vista que as dimensões de largura e altura do FinFET são as mais críticas para o seu funcionamento, serão mantidas as utilizadas no capítulo anterior como sendo $h_{fin} = 60$ nm e $W_{fin} = 20$ nm, garantindo dessa forma um bom acoplamento. Tendo em vista a equação (9), torna-se necessária a adoção de um $W_{UTBB} = 60.2 + 20 = 140$ nm. Como as larguras se tornam fixas em função das restrições do FinFET, serão utilizados diferentes comprimentos de canal, de modo a utilizar um conjunto de dispositivos com diferentes resistências térmicas. A partir dessas dimensões, já é possível fazer uma primeira aproximação das características necessárias de ambos os dispositivos. Adotando outros parâmetros elétricos e térmicos similares ao do FinFET para o UTBB, serão utilizados espessura equivalente do óxido de porta como 1,2 nm (EOT), dopagens de arsênio com concentração 10^{20}

cm^{-3} na fonte e no dreno, óxido enterrado de 30 nm (t_{BOX}) e espessura de 10 nm do filme de silício (t_{Si}). Deve-se notar que a espessura do filme de silício acaba influenciando as dimensões dos contatos de fonte e dreno, resultando em uma área adicional destes contatos de $t_{\text{Si}} \cdot W_{\text{UTBB}} - h_{\text{fin}} \cdot W_{\text{fin}} = 10 \text{ nm} \cdot 140 \text{ nm} - 60 \text{ nm} \cdot 20 \text{ nm} = 200 \text{ nm}^2$, ou seja, um aumento de aproximadamente 17%. Para facilitar a convergência numérica e melhorar o comportamento do dispositivo reduzindo-se transições abruptas, será adotada uma região de dreno levemente dopada com arsênio com concentração de 10^{17} cm^{-3} e comprimento de 10 nm. Construindo o transistor UTBB projetado com as características acima no Sentaurus Device Editor, foram obtidas as estruturas apresentadas na Figura 6.1.

Figura 6.1 – Estrutura gerada para o transistor UTBB (a) tridimensional (b) corte evidenciando W_{UTBB} (c) corte evidenciando L_{UTBB} (d) desenho esquemático com as principais dimensões utilizadas

Fonte: o autor (2017).

Como todas as considerações feitas sobre os modelos físicos na seção 5.3 continuam válidas dentro deste contexto de comparação entre dispositivos, todos eles serão mantidos idênticos aos previamente mencionados.

6.2 MÉTODO BASEADO NA EXTRAÇÃO PULSADA

Aplicando a modificação do método pulsado utilizada na seção 5.4, obteve-se para dispositivos UTBB o resultado apresentado na Figura 6.3, o qual é resumido na Tabela 6.1, juntamente com os dados dos FinFETs extraídos na seção 5.4.

Figura 6.2 – Estimativa da temperatura no canal a partir das curvas sem autoaquecimento em múltiplas temperaturas e uma curva com autoaquecimento a temperatura ambiente para (a) $L = 150$ nm, (b) $L = 300$ nm e (c) $L = 500$ nm

Fonte: o autor (2017).

Figura 6.3 – Extração da resistência térmica a partir da temperatura estimada no canal em função da potência aplicada ao dreno para dispositivo UTBB

Fonte: o autor (2017).

Tabela 6.1 – Resultados da extração da resistência térmica a partir do método baseado na extração pulsada

L (nm)	$R_{th,FinFET}$ (K/W)	$R_{th,UTBB}$ (K/W)
150	$5,72 \cdot 10^5$	$3,26 \cdot 10^5$
300	$3,68 \cdot 10^5$	$1,99 \cdot 10^5$
500	$2,82 \cdot 10^5$	$1,29 \cdot 10^5$

Fonte: o autor (2017).

A partir destes dados de resistência térmica, faz-se a estimativa de temperatura de ambas as tecnologias a partir de (31) no pior caso, isto é, com máxima potência aplicada ($V_{GS} = 1,5$ V e $V_{DS} = 2,0$ V), de forma a evidenciar a máxima diferença entre temperaturas, resultando na Figura 6.4.

Figura 6.4 – Comparação das temperaturas de operação entre dispositivos FinFET e UTBB na situação de potência elevada obtida a partir do método de extração pulsado

Fonte: o autor (2017).

6.3 MÉTODO DA ASSINATURA NA EFICIÊNCIA DO TRANSISTOR

Para realizar as comparações, são feitos os gráficos de g_m/I_D em função de V_{DS} para os dispositivos operando com mesmo V_{GT} , como apresentado na Figura 6.5.

Figura 6.5 – Comparação entre dispositivos FinFET e UTBB utilizando o método da assinatura na eficiência do transistor, com sobretensão de condução de 1 V e (a) $L = 150 \text{ nm}$, (b) $L = 300 \text{ nm}$ e (c) $L = 500 \text{ nm}$

(c)

Fonte: o autor (2017).

Como a comparação visual não permite estabelecer parâmetros objetivos para diferenciar as curvas, foi realizada uma regressão linear das regiões de saturação (ou seja, um ajuste da forma $y = a.x + b$), com os coeficientes apresentados na Tabela 6.2

Tabela 6.2 – Regressão linear das regiões de saturação da eficiência do transistor em função da tensão de dreno

L (nm)	FinFET		UTBB	
	a (V^{-2})	b (V^{-1})	a (V^{-2})	b (V^{-1})
150	-0,073	1,3	-0,058	1,4
300	-0,044	1,4	-0,030	1,5
500	-0,027	1,5	-0,015	1,6

Fonte: o autor (2017).

6.4 MÉTODO NUMÉRICO A PARTIR DO INVERSO DA EFICIÊNCIA DO TRANSISTOR

Aplicando o método descrito na seção 5.2 através do algoritmo do Apêndice A, foram obtidos os resultados sumarizados na Tabela 6.3.

Tabela 6.3 – Resultados da extração da resistência térmica a partir do método numérico

L (nm)	$R_{th,FinFET}$ (K/W)	$R_{th,UTBB}$ (K/W)
150	$5,49.10^5$	$2,20.10^5$
300	$2,48.10^5$	$1,09.10^5$
500	$1,34.10^5$	$0,56.10^5$

Fonte: o autor (2017).

Aplicando os valores de resistência térmica obtidos em (31) para $V_{GS} = 1,5 \text{ V}$ e $V_{DS} = 2,0 \text{ V}$, foram obtidas as temperaturas apresentadas na Figura 6.6.

Figura 6.6 – Comparação das temperaturas de operação entre dispositivos FinFET e UTBB na situação de potência elevada obtida a partir do método de extração proposto

Fonte: o autor (2017)

6.5 DISCUSSÃO

Para sintetizar todos os resultados obtidos ao longo deste capítulo, elaborou-se a Tabela 6.4 de modo a facilitar as discussões.

Tabela 6.4 – Resumos dos resultados obtidos para os diferentes métodos para dispositivos FinFET e UTBB

L (nm)	FinFET			UTBB		
	Pulsado R_{th} (K/W)	Proposto R_{th} (K/W)	g_m/I_D a (V^{-2})	Pulsado R_{th} (K/W)	Proposto R_{th} (K/W)	g_m/I_D a (V^{-2})
150	$5,72 \cdot 10^5$	$8,51 \cdot 10^5$	-0,073	$3,26 \cdot 10^5$	$3,89 \cdot 10^5$	-0,058
300	$3,68 \cdot 10^5$	$3,20 \cdot 10^5$	-0,044	$1,99 \cdot 10^5$	$1,46 \cdot 10^5$	-0,030
500	$2,82 \cdot 10^5$	$1,67 \cdot 10^5$	-0,027	$1,29 \cdot 10^5$	$1,25 \cdot 10^5$	-0,015

Fonte: o autor (2017)

Primeiramente, para o método baseado na técnica de extração pulsada, fica evidente que todos os dispositivos FinFET apresentaram uma resistência térmica mais elevada do que os dispositivos UTBB, sendo que o valor absoluto da diferença

aumenta para menores comprimentos de canal. Traduzindo esta análise em termos de temperaturas, a tendência é mantida, isto é, quanto menor o comprimento de canal, maior é a diferença entre a temperatura do FinFET em relação ao UTBB.

Para o método de análise da eficiência do transistor em função da tensão de dreno, apesar de claramente haver uma mudança significativa entre dispositivos de comprimentos diferentes, não é imediatamente evidente através da inspeção visual a mudança da inclinação na região de saturação para dispositivos de mesmo comprimento. Para resolver este problema, a solução implementada foi realizar o ajuste de uma reta e avaliar os coeficientes. A partir dessa ideia, verifica-se, ao comparar os coeficientes lineares (a) na Tabela 6.2, que o dispositivo FinFET apresenta uma maior declividade em módulo (coeficiente a), reforçando a hipótese de que este dispositivo sofre mais intensamente de autoaquecimento do que o UTBB.

Por fim, para a análise a partir da estimativa numérica utilizando o inverso da eficiência do transistor, obteve-se um resultado geral no mesmo sentido, isto é, o FinFET novamente apresentou piores características térmicas do que o UTBB.

Sob o ponto de vista físico, os resultados até aqui observados são coerentes, pois os dispositivos foram projetados de modo a ter contatos de porta, fonte e dreno com áreas similares, resultando em uma dissipação térmica semelhante através deles. Portanto, o diferencial fica em função do contato de substrato, e, neste caso, dois fatores desfavorecem o FinFET: seu óxido enterrado cinco vezes mais espesso ($t_{\text{BOX,FinFET}} = 150 \text{ nm} > t_{\text{BOX,UTBB}} = 30 \text{ nm}$), e a sua área de contato sete vezes menor ($L.W_{\text{fin}} = L.20 \text{ nm} < L.W_{\text{UTBB}} = L.140 \text{ nm}$). Assim, é de se esperar que o SOI FinFET apresente uma maior resistência térmica total, e, conseqüentemente, sofra um maior efeito de autoaquecimento.

Tomando o método baseado na técnica pulsada como referência para comparação com a análise da assinatura na eficiência do transistor, fica evidente que a relação das resistências térmicas não pode ser tomada como diretamente proporcional à degradação da eficiência na saturação, confirmando a ideia de que a técnica da assinatura serve somente para comparações de primeira ordem, com o intuito de verificar qual dispositivo possui uma maior ou menor resistência térmica. Além disso, se fatores adicionais como diferença de comprimento de canal não forem considerados, a análise pode não ser completamente precisa, como evidenciado ao comparar-se o dispositivo FinFET de $L = 300 \text{ nm}$ com o UTBB de $L = 150 \text{ nm}$. Neste caso, verifica-se pelo método baseado na técnica pulsada que o dispositivo FinFET é

pior, entretanto, a análise simples dos coeficientes no método da assinatura aparentaria indicar o contrário. Como já discutido, uma possível fonte para esse erro se dá pela desconsideração das diferentes degradações da mobilidade devido ao campo elétrico transversal.

Comparando agora o método baseado na técnica pulsada com a estimativa numérica feita a partir do inverso da eficiência do transistor, nota-se que a estimativa da resistência térmica para dispositivos FinFET apresenta um erro menor do que para os UTBBs. Uma explicação para este fenômeno é que o erro cometido no método proposto aparenta diminuir significativamente conforme aumenta a resistência térmica do dispositivo avaliado, e, como todos os FinFETs apresentam maior resistência térmica, o erro cometido para eles aparenta ser menor. Duas evidências que reforçam essa hipótese são o erro reduzido cometido para o FinFET de comprimento de canal de 150 nm, que foi o dispositivo com maior resistência térmica tanto pelo método tradicional quanto pelo proposto, e a similaridade entre os valores de resistência térmica obtidos para o FinFET de $L = 300$ nm com o UTBB de $L = 150$ nm, implicando em erros muito próximos para dois valores similares de resistência térmica de dispositivos inteiramente diferentes. Do ponto de vista do modelo do inverso da eficiência do transistor, o que pode explicar este aumento do erro para menores valores de resistência térmica seria o aumento da importância do fator M , o qual é estimado de forma bastante simplificada através de uma regressão linear. Um refinamento no seu cálculo pode levar a valores de erro reduzidos até mesmo para dispositivos com baixa resistência térmica.

7 CONCLUSÕES

Tendo como objetivo o estudo comparativo do efeito de autoaquecimento nas tecnologias SOI UTBB e FinFET, novas técnicas de caracterização do efeito foram desenvolvidas a partir da modelagem física do fenômeno. Simulações numéricas montadas a partir dos modelos tradicionais de parâmetros críticos ao efeito, como a mobilidade e a tensão de limiar, foram utilizadas como ferramentas para testar e comprovar as deduções matemáticas realizadas. Ademais, medidas experimentais foram feitas utilizando dispositivos FinFET canal p e comparadas aos resultados das simulações, corroborando os resultados obtidos a partir da teoria.

A técnica da condutância de saída média (método 1) baseia a detecção do autoaquecimento não somente em valores negativos da condutância de saída para uma dada polarização de porta fixa, como feito tradicionalmente, mas também analisando todo o seu comportamento quando essa polarização é variada em um intervalo que garanta a operação do dispositivo em saturação. Três pontos críticos na curva de condutância média de saída em função da tensão de porta são definidos de forma a permitir a identificação do autoaquecimento. Apesar de permitir a identificação do fenômeno em condições nas quais normalmente ele seria relevado, a metodologia não é a mais adequada para comparação entre dispositivos, só podendo ser aplicada quando a modulação de canal é desprezível. Caso o efeito de modulação do comprimento de canal seja significativo e diferente entre os dispositivos, as definições de intensidade do efeito não são comparáveis.

A partir das expressões básicas de corrente e transcondutância, foi possível concluir que transistores em saturação deveriam apresentar eficiência constante em função da tensão de dreno, assumindo que o autoaquecimento não esteja presente. Caso contrário, demonstrou-se que quanto maior a tensão de dreno, maior será a degradação desse parâmetro. Verificou-se então que os gráficos da eficiência do transistor em função da corrente de dreno com tensão de porta constante e tensão de dreno variável apresentam uma assinatura única em dispositivos que sofrem com o autoaquecimento. Com essa assinatura, determinaram-se cinco classificações para a intensidade do autoaquecimento no dispositivo, que é uma análise mais completa do que o método anterior; porém, após uma inspeção mais detalhada, ficou claro que essa característica também dependeria da intensidade do efeito de modulação do comprimento de canal. Entretanto, esse problema pôde ser sanado utilizando a tensão

de dreno como eixo das abscissas (método 2). Apesar de ocultar informações sobre a relação entre a modulação do comprimento de canal e o autoaquecimento no dispositivo, tal modificação permite a comparação objetiva entre dispositivos. É necessário ainda levar-se em conta a diferença entre os parâmetros de degradação da mobilidade com o campo elétrico transversal, entretanto, considerando a menor correlação desse efeito com o autoaquecimento, a comparação torna-se mais simples.

Ao fazer a análise do inverso da eficiência do transistor (método 3), verificou-se que é possível obter não apenas uma forma de linearizar a expressão na saturação, mas também de extrair numericamente o valor da resistência térmica. Entretanto, para que essa extração ocorra corretamente, é necessário ter conhecimentos de dois valores, a degradação da mobilidade com o campo elétrico transversal e o coeficiente de degradação da mobilidade com a temperatura. O primeiro é obtido através de processos de extração em baixa polarização, portanto, não sofre de autoaquecimento. O segundo, por possuir correlação com a polarização de dreno, precisa ser extraído em saturação; contudo, como normalmente são necessárias polarizações mais elevadas para tanto, deve-se levar em conta o aumento de temperatura em função do autoaquecimento. Como resolução, propôs-se um novo método iterativo, que permite a extração não somente do coeficiente de degradação da mobilidade com a temperatura e da resistência térmica, mas também elimina a necessidade de extração da constante de degradação da mobilidade com campo elétrico transversal. Além destas vantagens, este é o primeiro método proposto capaz de estimar a resistência térmica de um transistor utilizando somente caracterizações em corrente contínua.

Finalmente, foi feita a comparação entre dispositivos SOI FinFET e SOI UTBB, utilizando um ambiente de simulações, de forma a garantir que somente a geometria dos dispositivos interferiria nas análises feitas, eliminando problemas de variabilidade de processo. Como parâmetro de comparação, adotou-se que ambos dispositivos deveriam ter correntes similares, de modo a manter a potência constante. Dentre todos os métodos de análise utilizados e desenvolvidos para esse estudo, houve unanimidade que dispositivos SOI FinFETs apresentam piores características térmicas do que os SOI UTBBs, e, portanto, sofrem mais de autoaquecimento.

Como propostas de passos futuros, colocam-se os seguintes itens: uma otimização do método numérico para reduzir o erro em dispositivos com menor resistência térmica; uma análise para dispositivos SOI FinFET de múltiplas aletas, uma vez que o aumento da área de contato pode trazer vantagens térmicas

significativas, equiparando-os possivelmente ao SOI UTBB; o estudo de FinFETs de corpo, considerando que a dissipação do calor através do contato de substrato aparenta ser um fator significativo para a redução da resistência térmica do SOI UTBB em relação ao SOI FinFET, e, portanto, a eliminação do óxido enterrado no caso do FinFET de corpo pode trazer vantagens térmicas significativas.

PUBLICAÇÕES GERADAS

MORI, C. A. B.; AGOPIAN, P. G. D. ; MARTINO, J. A. . Improvement for the thermal resistance extraction method related with self-heating effect. In: Microelectronics Students Forum 2016, 2016, Belo Horizonte. Microelectronics Students Forum 2016, 2016. v. 16. p. 1-3.

MORI, C. A. B.; AGOPIAN, P. G. D. ; MARTINO, J. A. . Influence of the dimensions on pFinFET devices towards the self-heating effect. In: Seminatec 2017 - XII Workshop on Semiconductions And Micro & Nano technology, 2017, São Paulo. Seminatec 2017 - XII Workshop on Semiconductions and Micro & Nano technology, 2017. v. 12. p. 1-2.

MORI, C. A. B.; AGOPIAN, P. G. D.; MARTINO, J. A. Simple method for detection of the self-heating signature. In: SYMPOSIUM ON MICROELECTRONICS TECHNOLOGY AND DEVICES, 32., 2017, Fortaleza.

MORI, C. A. B.; AGOPIAN, P. G. D.; MARTINO, J. A. New method for observing self-heating signature using transistor efficiency. In: IEEE SOI-3D-SUBTHRESHOLD MICROELECTRONICS TECHNOLOGY UNIFIED CONFERENCE, 2017, San Francisco.

MORI, C. A. B.; AGOPIAN, P. G. D.; MARTINO, J. A. Improvement of g_m/I_D method for detection of self-heating effects. In: ELECTROCHEMICAL SOCIETY MEETING, 232., 2018, Seattle. (aceito)

MORI, C. A. B.; AGOPIAN, P. G. D.; MARTINO, J. A. New method for self-heating estimation using only DC measurements. In: Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon, 2018, Granada.

BIBLIOGRAFIA

- ALMEIDA, L. M. Conceitos básicos sobre a tecnologia SOI. In: ALMEIDA, L. M. **Estudo de célula de memória dinâmica de apenas um transistor SOI de óxido enterrado ultrafino**. São Paulo: [s.n.], 2012. Cap. 2, p. 16-43.
- ARORA, N. D. **MOSFET Models for VLSI Circuit Simulation**. 1ª. ed. Viena: Springer-Verlag Wien, 1993.
- BAUM, G.; BENEKING, H. Drift velocity saturation in MOS transistors. **IEEE Transactions on Electron Devices**, v. 17, n. 6, p. 481-482, Junho 1970.
- BUCK, I. IBM Cloud to Offer AI Computing with Latest NVIDIA Pascal GPUs. **NVIDIA Blog**, 2017. Disponível em: <<https://blogs.nvidia.com/blog/2017/04/05/ibm-cloud-tesla-p100/>>. Acesso em: 18 Julho 2017.
- CAPARROZ, L. F. V. **Efeito da Radiação Em Transistores 3D Em Baixas Temperaturas**. Escola Politécnica da Universidade de São Paulo. São Paulo, p. 117. 2017.
- CAUGHEY, D. M.; THOMAS, R. E. Carrier mobilities in silicon empirically related to doping and field. **Proceedings of the IEEE**, v. 55, n. 12, p. 2192-2193, Dezembro 1967.
- CHAIN, K. et al. A MOSFET electron mobility model of wide temperature range (77–400 K) for IC simulation. **Semiconductor Science and Technology**, v. 12, n. 4, p. 355-358, Abril 1997.
- COLINGE, J. -P. **Silicon-On-Insulator Technology: Materials to VLSI**. 3ª. ed. Nova Iorque: Springer, 2004.
- COLINGE, J. -P. The SOI MOSFET: from Single Gate to Multigate. In: _____(ED.) **FinFETs and other Multi-Gate Transistors**. 1ª. ed. Nova Iorque: Springer, 2008. Cap. 1, p. 1-48.
- DAUGÉ, F. et al. Coupling effects and channels separation in FinFETs. **Solid-State Electronics**, v. 48, n. 4, p. 535-542, Abril 2004.
- ERNST, T. et al. Ultimately Thin Double-Gate SOI MOSFETs. **IEEE Transactions on Electron Devices**, v. 50, n. 3, p. 830-838, Março 2003.
- FENOUILLET-BERANGER, C. et al. FDSOI devices with thin BOX and ground plane integration for 32nm node and below. **Solid-State Electronics**, v. 53, n. 7, p. 730-734, 2009.
- FOTY, D. **MOSFET Modeling With SPICE: Principles and Practice**. 1. ed. Upper Saddle River: Prentice Hall, 1997.

- GHIBAUDO, G. New method for the extraction of MOSFET parameters. **Electronics Letters**, v. 24, n. 9, p. 543-545, Abril 1988.
- GROESENEKEN, G. et al. Temperature dependence of threshold voltage in thin-film SOI MOSFETs. **IEEE Electron Device Letters**, v. 11, n. 8, p. 329-331, Agosto 1990.
- HENRY, J. B. et al. New Y-function based MOSFET parameter extraction method from weak to strong inversion range. **Solid-State Electronics**, v. 123, p. 84-88, Setembro 2016. ISSN 0038-1101.
- INTEL. Intel Announces New 22 nm 3D Tri-gate Transistors. **Intel**, 2011. Disponível em: <<https://www.intel.com/content/www/us/en/silicon-innovations/standards-22nm-3d-tri-gate-transistors-presentation.html>>. Acesso em: 7 Julho 2017.
- IWAI, H. **Future of nano CMOS technology**. 28th Symposium on Microelectronics Technology and Devices (SBMicro 2013). Curitiba: [s.n.]. 2013. p. 1-10.
- JIN, M. et al. **Hot carrier reliability characterization in consideration of self-heating in FinFET technology**. 2016 IEEE International Reliability Physics Symposium (IRPS). Pasadena: [s.n.]. 2016. p. 2A-2-1-2A-2-5.
- JOMAAH, J. et al. **Impact of self-heating effects on the design of SOI devices versus temperature**. 1995 IEEE International SOI Conference Proceedings. Tucson: [s.n.]. 1995. p. 114-115.
- JOMAAH, J.; GHIBAUDO, G.; BALESTRA, F. Analysis and Modeling of Self-Heating Effects in Thin-Film SOI MOSFETs as a Function of Temperature. **Solid-State Electronics**, v. 38, n. 3, p. 615-618, Março 1995.
- KILBY, J. S. **Miniaturized electronic circuits**. US 3138743 A, 6 Fevereiro 1959.
- KILCHYTSKA, V. et al. **Perspective of FinFETs for analog applications**. Proceedings of the 30th European Solid-State Circuits Conference. Leuven: [s.n.]. 2004. p. 65-68.
- LIM, H. K.; FOSSUM, J. G. Transient drain current and propagation delay in SOI CMOS. **IEEE Transactions on Electron Devices**, v. 31, n. 9, p. 1251-1258, Setembro 1984.
- MACAMBIRA, C. N. **Estudo do ponto invariante com a temperatura (ZTC) em UTBB SOI nMOSFETs**. Escola Politécnica da Universidade de São Paulo. São Paulo, p. 85. 2017.
- MAKOVEJEV, S. **Characterisation of Thermal and Coupling Effects in Advanced Silicon MOSFETs**. Newcastle upon Tyne: Newcastle University, 2012.

MARTINO, J. A.; PAVANELLO, M. A.; VERDONCK, P. B. **Caracterização elétrica de tecnologia e dispositivos MOS**. 1ª. ed. São Paulo: Pioneira Thomson Learning, 2003.

MCDALD, L. J. et al. Physical origin of negative differential resistance in SOI transistors. **Electronics Letters**, v. 25, n. 13, p. 827-828, Junho 1989.

MOORE, G. E. Cramming more components onto integrated circuits. **Electronics**, v. 38, n. 8, p. 114-117, Abril 1965.

MORGAN, J. A Simple Explanation Of 'The Internet Of Things'. **Forbes**, 2014. Disponível em: <<https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#2eb9115b1d09>>. Acesso em: 18 Julho 2017.

NASCIMENTO, V. M. **Estudo do Ponto Invariante com a Temperatura (ZTC) em SOI-FinFETs Tensionados e Radiados**. Escola Politécnica da Universidade de São Paulo. São Paulo, p. 98. 2017.

ORTIZ-CONDE, A. et al. A review of recent MOSFET threshold voltage extraction methods. **Microelectronics Reliability**, v. 42, n. 4-5, p. 583-596, Abril-Maio 2002.

PAVANELLO, M. A. et al. Temperature Influences on FinFETs with Undoped Body. **ECS Transactions**, v. 6, n. 4, p. 211-216, 2007.

RINALDI, N. Small-signal operation of semiconductor devices including self-heating, with application to thermal characterization and instability analysis. **IEEE Transactions on Electron Devices**, v. 48, n. 2, p. 323-331, Fevereiro 2001.

RUDENKO, T. et al. Carrier Mobility in Undoped Triple-Gate FinFET Structures and Limitations of Its Description in Terms of Top and Sidewall Channel Mobilities. **IEEE Transactions on Electron Devices**, v. 55, n. 12, p. 3532-3541, Dezembro 2008.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627-1639, Julho 1964.

SEDRA, A. S.; SMITH, K. C. **Microelectronic Circuits**. 6ª. ed. Oxford: Oxford University Press, 2010.

SIMONITE, T. Moore's Law Is Dead. Now What? - MIT Technology Review. **MIT Technology Review**, 2016. Disponível em: <<https://www.technologyreview.com/s/601441/moores-law-is-dead-now-what/>>. Acesso em: 18 Julho 2017.

SZE, S. M. **Physics of Semiconductor Devices**. 2ª. ed. Nova Iorque: Wiley, 1981.

TAKAHASHI, T. et al. **Comparison of self-heating effect (SHE) in short-channel bulk and ultra-thin BOX SOI MOSFETs**: Impacts of doped well, ambient temperature, and SOI/BOX thicknesses on SHE. 2013 IEEE International Electron Devices Meeting. Washington, DC: [s.n.]. 2013. p. 7.4.1-7.4.4.

WALDROP, M. M. The chips are down for Moore's law: Nature News & Comment. **Nature**, 2016. Disponivel em: <<http://www.nature.com/news/the-chips-are-down-for-moore-s-law-1.19338>>. Acesso em: 18 Julho 2017.

WONG, H.-S. et al. Modeling of transconductance degradation and extraction of threshold voltage in thin oxide MOSFET's. **Solid-State Electronics**, v. 30, n. 9, p. 953-968, Setembro 1987.

APÊNDICE A – Rotina em MATLAB para estimativa da resistência térmica a partir de medidas em corrente contínua

```
function [Rth,c,F] = est_Rth( VGS, VDS, ID, T, VG_temp, VD_temp, ID_temp )

%Estimativa da resistência térmica a partir de VGT, VDS e ID
% VGS - (alta) tensão de porta para extração de Rth - L x 1
% VDS - tensão de dreno para extração de Rth - N x 1
% ID - corrente de dreno para extração de Rth - N x 1
% T - temperaturas - K x 1
% VG_temp - tensão de porta para extração de VT, alpha, c, theta - M x 1
% VD_temp - (alta) tensão de dreno para extração c - 1 x 1
% ID_temp - corrente de dreno para extração de c - M x K
```

Preparo dos dados de entrada

```
close all
M = length(VG_temp);
N = length(VDS);
K = length(T);

gm = diff(ID)./repmat(diff(VGS),[1,N]);

IDgm = ID(1:end-1,:)./gm;
gm_temp = diff(ID_temp)./repmat(diff(VG_temp),[1,K]);
```

Determinação de VT com a temperatura

```
div = 10;
w = floor(M/div);
S0normr = 1e10;
p0 = [0;0];
for j = 1:K
    for i = 1:M-w
        a = i:i+w;
        [p(:,j),S] = polyfit(VG_temp(a),sqrt(ID_temp(a,j)),1);
        if S.normr < S0normr && -p(2,j)/p(1,j) > .2
            I(:,j) = a;
            p0(:,j) = p(:,j);
            S0normr = S.normr;
        end
    end
end

VT = (-p0(2,:)./(p0(1,:)))';

figure
hold all
for i = 1:K
    plot(VG_temp,sqrt(ID_temp(:,i)))
```

```

plot(VG_temp(I(:,i)),polyval(p0(:,i),VG_temp(I(:,i))));
end

clear S S0normr a div i j p w I

```

Determinação da resistência térmica

```

Rth0 = -1;
Rth = 0;
count = 0;
VGT0 = 1;
P = 0;
P_temp = zeros(K,1);
c = zeros(K-1,1);
VGT = repmat(VG_temp,[1 K]) - repmat(VT',[M 1]);
[~,iVGT] = min(abs(VGT-VGT0));
for i = 1:K
    P_temp(i) = VD_temp*ID_temp(iVGT(i),i);
end

while abs((Rth-Rth0)/Rth0) > 1e-5
    Rth0 = Rth;

    Test = T+Rth0*P_temp;

    for i = 2:K
        c(i-1) = log(gm_temp(iVGT(i),i)./gm_temp(iVGT(1),1))./...
            log((T(1)+Rth0*P_temp(i))./Test(i))';
    end
    alpha = polyfit(Test,VT,1);

    Test_SHE = T(1) + Rth*mean(P);
    VGT_SHE = VGS - polyval(alpha,Test_SHE);
    [~,iVGT2] = min(abs(VGT_SHE-VGT0));
    [~,a] = min(abs(VGT_SHE(iVGT2) - VDS));
    a=a+5;
    P = VDS(a:end).*ID(iVGT2,a:end)';
    IDgm_sat = IDgm(iVGT2,a:end)';

    [~,ind] = min(abs(T-Test_SHE));
    if ind >= K
        ind = K - 1;
    end
    c = c(ind);

    F = polyfit(P,IDgm_sat,1);
    F = F(2);

    save('IDgm.mat','P','IDgm_sat');
    O = optimoptions(@lsqnonlin);
    O.TolX = 1e-15;
    O.TolFun = 1e-15;
    O.MaxFunEvals = 1e4;
    O.MaxIter = 1e4;
    x = lsqnonlin(@IDgmfun,[1e5 F c],[0 F*0.99 c*.95],...
        [1e7 F*1.01 c*1.05],O);
    Rth = x(1);

```



```
        count = count+1;  
        if count == 100  
            break;  
        end  
    end  
close all  
end
```