

HAMILTON KLIMACH

**MODELO DO DESCASAMENTO (*MISMATCH*)
ENTRE TRANSISTORES *MOS***

**FLORIANÓPOLIS
2008**

UNIVERSIDADE FEDERAL DE SANTA CATARINA

**PROGRAMA DE PÓS-GRADUAÇÃO
EM ENGENHARIA ELÉTRICA**

**MODELO DO DESCASAMENTO (*MISMATCH*)
ENTRE TRANSISTORES *MOS***

Tese submetida à
Universidade Federal de Santa Catarina
como parte dos requisitos para a
obtenção do grau de Doutor em Engenharia Elétrica.

HAMILTON KLIMACH

Florianópolis, março de 2008.

MODELO DO DESCASAMENTO (*MISMATCH*) ENTRE TRANSISTORES *MOS*

Hamilton Klimach

‘Esta Tese foi julgada adequada para obtenção do Título de Doutor em Engenharia Elétrica, Área de Concentração em *Circuitos e Sistemas Integrados*, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.’

Carlos Galup-Montoro, Doutor
Orientador

Kátia Campos de Almeida, Doutora
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

Banca Examinadora:

Carlos Galup-Montoro, Doutor
Presidente

Márcio Cherem Schneider, Doutor
Co-orientador

Jacobus W. Swart, Doutor

Fabiano Fruett, Doutor

Fernando Silveira, Doutor

Dedico esta obra à minha filha Sofia.

AGRADECIMENTOS

Gostaria de agradecer aos Professores Carlos Galup-Montoro e Márcio Cherem Schneider pela excelente orientação recebida durante meu doutoramento, pela forma generosa como fui acolhido em seu laboratório, e principalmente pelo exemplo valioso, percebido no convívio diário com estes dois brilhantes profissionais da pesquisa científica. Gostaria também de agradecer aos amigos e colegas de laboratório Pablo Dutra, Maurício Camacho, Luis Spiller e William Prodanov, com quem dividi tantos momentos, e que tornaram minha estada em Florianópolis mais agradável e enriquecedora, à secretária do departamento Nazide Martins, sempre prestativa, e a todos os colegas com quem convivi no Laboratório de Circuitos Integrados da UFSC.

Também, gostaria de agradecer a todos os colegas do Departamento de Engenharia Elétrica da UFRGS, pelo suporte que me foi concedido durante o afastamento, e em especial, aos amigos Eric Fabris e Altamiro Susin, que me foram fundamentais no sucesso desta empreitada.

Agradeço também aos Professores Jacobus Swart, Fabiano Fruett e Fernando Silveira, por participarem da banca de defesa e pelas críticas e sugestões que enriqueceram este trabalho. Ainda, agradeço ao Professor Sergio Bampi pelos excelentes comentários e sugestões ao texto, aos Professores André Pasa e Oscar Gouveia, e a Alfredo Arnaud.

Agradeço ainda à CAPES, pelo apoio financeiro, e à MOSIS pela fabricação gratuita dos circuitos prototipados.

Finalmente, gostaria de agradecer profundamente aos meus pais, Alberto e Silvia, responsáveis por muito do que sou hoje, e especialmente à minha esposa Vera, pelo incansável apoio durante este período, e por ter enriquecido tanto a minha vida ao longo dos anos. E, principalmente, por ter trazido ao mundo aquela que é a pessoa mais importante da minha vida e que tanto me faz crescer, através da sua espontaneidade e do seu amor de filha, que com apenas quatro anos, me ensina mais que quarenta de existência.

Resumo da Tese apresentada à UFSC como parte dos requisitos necessários para a obtenção do grau de Doutor em Engenharia Elétrica.

MODELO DO DESCASAMENTO (*MISMATCH*) ENTRE TRANSISTORES *MOS*

Hamilton Klimach

Março/2008

Orientador: Carlos Galup-Montoro, Doutor.

Co-orientador: Márcio Cherem Schneider, Doutor.

Área de Concentração: Circuitos e Sistemas Integrados.

Palavras-chave: descasamento, modelo compacto, MOSFET, projeto analógico.

Número de Páginas: 191.

Diversos modelos teóricos para o descasamento entre dispositivos na tecnologia *MOS* foram propostos desde a década de '80, sendo que geralmente estes pecam ou pela simplicidade, sendo válidos apenas sob condições de operação específicas, ou por resultarem em expressões muito complexas, o que torna necessário o uso de pesados recursos computacionais. Esta tese propõe uma abordagem inovadora para a modelagem do descasamento dos transistores de efeito de campo de *porta* isolada (*MOSFETs*), chegando a resultados melhores e mais abrangentes que outras propostas já publicadas. Para tanto, as variações microscópicas na corrente que flui pelo dispositivo, resultado das flutuações na concentração de dopantes na região ativa, são contabilizadas levando-se em conta a natureza não-linear do transistor. O resultado é um modelo compacto que prevê o descasamento com grande exatidão e de forma contínua, em todas as condições de operação do transistor, da inversão fraca à forte, e da região linear à saturação, necessitando apenas dois parâmetros de ajuste. Duas versões de circuitos de teste foram desenvolvidas e implementadas em diversas tecnologias, como forma de se obter suporte experimental para o modelo. A versão mais avançada possibilita a caracterização elétrica, de forma totalmente automática, de um grande número de dispositivos. O uso deste modelo substitui com vantagens a tradicional simulação *Monte Carlo*, que exige grandes recursos computacionais e consome muito tempo, além de oferecer uma excelente ferramenta de projeto manual, como é demonstrado através do desenvolvimento de um conversor digital-analógico, cujo resultado experimental corroborou a metodologia empregada.

Abstract of Thesis presented to UFSC as a partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering.

MISMATCH MODEL FOR *MOS* TRANSISTORS

Hamilton Klimach

March/2008

Advisor: Carlos Galup-Montoro, PhD.

Co-advisor: Márcio Cherem Schneider, PhD.

Area of Concentration: Integrated Circuits and Systems.

Keywords: mismatch, matching, compact model, MOSFET, analog design.

Number of Pages: 191

Many mismatch models were proposed for the *MOS* devices since the '80s, but they use either too simple approaches, being restricted to specific operating conditions, or too complex expressions, only useful through hard computational resources. This thesis proposes a new approach for *MOSFETs* mismatch modeling, presenting better and more general results than that found in preceding articles. In this approach, the microscopic variations of the drain current, caused by random doping fluctuation inside the channel region, are integrated along the channel, considering the main transistor nonlinearities. It results in a compact model that accurately predicts mismatch, continuously over any transistor operating condition, from weak to strong inversion, and from linear to saturation region, and only needing two fitting parameters. Two versions of a test chip were developed and fabricated in many technologies to give experimental support to this model. The most advanced of them makes the automated electrical characterization possible for a huge number of devices. This model can surpass the traditional *Monte Carlo* simulation method with advantages, and can also be used as a hand-design tool, as demonstrated here through the design of a digital-to-analog converter.

SUMÁRIO

AGRADECIMENTOS	vii
SUMÁRIO.....	xiii
Lista de Figuras	xv
Lista de Tabelas	xxi
1 INTRODUÇÃO.....	1
1.1 Circuitos eletrônicos analógicos e digitais	1
1.2 Projeto de circuitos integrados	3
1.3 Descasamento entre transistores MOS	4
1.4 Causas e tipos de descasamento	7
1.5 Técnicas de redução dos efeitos globais.....	9
1.6 Técnicas de redução dos efeitos locais.....	12
1.7 Estrutura do texto	13
2 MODELOS DE DESCASAMENTO	15
2.1 Revisão bibliográfica.....	15
2.2 O modelo de Pelgrom.....	20
2.3 Considerações sobre modelagem.....	24
3 O MODELO DE DESCASAMENTO PROPOSTO	27
3.1 Considerações gerais	27
3.2 O modelo ACM.....	29
3.3 Distribuição de Poisson	30
3.4 Flutuação na concentração de dopantes	32
3.5 Variações de condutância no canal e seus efeitos	33
3.6 Um modelo compacto para o descasamento entre MOSFETs	36
3.6.1 Introdução.....	37
3.6.2 Um modelo consistente para as flutuações na corrente de <i>dreno</i>	39
3.6.3 Um modelo para a flutuação na concentração de dopantes.....	40
3.6.4 O modelo de descasamento em termos de nível de inversão	44
3.6.5 Resultados experimentais	46
3.6.6 Conclusões.....	55
3.6.7 Apêndice.....	56
3.7 Efeitos secundários que afetam o descasamento	57
3.8 Efeito da saturação na velocidade dos portadores	57
3.8.1 Inclusão na corrente de <i>dreno</i>	58
3.8.2 Desvio na corrente de <i>dreno</i> devido a flutuações locais	61
3.8.3 Inclusão no modelo de descasamento.....	64
3.9 Depleção no substrato.....	67
3.9.1 Efeito da polarização <i>porta-substrato</i>	67
3.9.2 Efeito da polarização <i>dreno-fonte</i>	69
3.10 Depleção na porta.....	72
3.10.1 Efeito médio na depleção de <i>porta</i>	72

3.10.2 Flutuações na depleção de <i>porta</i>	76
3.11 Efeitos de rugosidade de borda	77
3.12 Efeitos da corrente específica I_S	79
4 CIRCUITO DE TESTE	81
4.1 Objetivos	81
4.2 Caracterização de descasamento	82
4.3 Aspectos de projeto	83
4.4 Quantidade de dispositivos	85
4.5 Formas de polarização do dispositivo	85
4.6 Formas de seleção do dispositivo	86
4.7 Circuito de teste – versão 1	88
4.8 Circuito de teste – versão 2	90
4.8.1 Geometria e leiaute	95
4.8.2 Especificação e projeto	98
4.8.3 Seleção pela <i>porta</i>	99
4.8.4 Seleção pelo <i>dreno</i>	100
4.8.5 Registrador de deslocamento	102
4.8.6 <i>Buffers</i> de entrada e saída	103
4.8.7 Fabricação e resultados	104
4.9 Conclusões	126
5 PROJETO DE UM CONVERSOR D/A M-2M	127
5.1 Objetivos	127
5.2 O divisor de corrente baseado em MOSFETs	128
5.3 O conversor D/A M-2M	131
5.4 Metodologia de Projeto	132
5.4.1 Determinação de L	133
5.4.2 Determinação da área	134
5.5 Implementação	137
5.6 Caracterização e resultados	138
5.7 Conclusões	140
6 CONCLUSÕES	143
Anexo A – Consistência de um modelo de descasamento	145
A.1 Consistência de associação paralela	146
A.2 Consistência de associação série	147
A.3 Consistência dos modelos apresentados	148
A.3.1 Verificação do modelo proposto	148
A.3.2 Verificação do modelo de Pelgrom	149
Anexo B – Flutuação local na carga de depleção e seus efeitos sobre o campo elétrico e tensão de limiar	153
B.1 Campo elétrico e carga de depleção equivalente	153
B.2 Tensão de limiar	157
Anexo C – Publicações	159
REFERÊNCIAS	161

Lista de Figuras

Figura 1.1: Comportamento estatístico do atraso de dois ramos de distribuição de sinal de <i>clock</i> em processo de 0,35 μ m, causado pelo descasamento dos transistores. O histograma (a) resulta da simulação do atraso de 200 amostras, para um único inversor de dimensão mínima e com $V_{DD}=1.8V$. Flutuações do atraso acumulado em duas cadeias de 4 inversores em (b), cuja distribuição se amplia à medida que o sinal avança na cadeia (ΔT_1 : 2 inversores e ΔT_2 : 4 inversores). A tabela (c) apresenta a distribuição do atraso, normalizada pelo atraso médio de um inversor unitário (σ_T/T), para diferentes dimensões de transistores e valores de tensão de alimentação. Observa-se que a dispersão do atraso aumenta à medida que a tensão de alimentação e o tamanho dos transistores diminuem, os dois grandes desafios na evolução da tecnologia moderna.	5
Figura 1.2: Percentual de aproveitamento (<i>yield</i>) de ADCs fabricados, em função do desvio-padrão na tensão de <i>off-set</i> (descasamento) do par de transistores de entrada (extraído de [10]).	6
Figura 1.3: Simulação <i>Monte Carlo</i> da tensão de <i>off-set</i> de um amplificador operacional <i>Miller CMOS</i> . O histograma apresenta a distribuição desta tensão sobre 1000 amostras, em intervalos de 0,5 mV. O desvio-padrão calculado é 2,1 mV. A curva tracejada é a sua aproximação Gaussiana.	6
Figura 1.5: O casamento é melhor se os transistores forem idênticos.	10
Figura 1.6: O casamento é melhor se os transistores estiverem próximos.	10
Figura 1.7: A distribuição centróide-comum (b) integra os efeitos de gradientes ao longo da área do circuito, fazendo com que, em média, os dois transistores sejam igualmente afetados.	11
Figura 1.8: Deve-se implementar elementos ou dispositivos <i>dummy</i> ao redor dos dispositivos casados que se encontram na extremidade (a) do conjunto (b), fazendo com que estes tenham as mesmas condições de contorno durante a fabricação (c). ...	11
Figura 1.9: O casamento é melhor se a corrente nos transistores tiver a mesma direção.	11
Figura 1.10: Fluxo de projeto, incluindo o modelo de descasamento (extraído de [10]). ...	13
Figura 2.1: Medida do desvio-padrão da tensão de limiar (σ_{AVT}), em função da área do transistor para transistores <i>NMOS</i> , e o parâmetro A_{VT} de diversas gerações de processos (extraído de [10]).	23
Figura 3.1: Simulação da distribuição de dopantes e potenciais no substrato e no polissilício de um MOSFET com 50nm de comprimento de canal.	32
Figura 3.2: Representação da região ativa de um transistor <i>MOS</i>	35
Figura 3.3: Divisão de um transistor em três elementos em série: (a) circuito equivalente com transistores, e (b) equivalente para pequenas perturbações (ou pequenos-sinais).	40
Figura 3.4: Modelo capacitivo do canal do <i>MOSFET</i> para a análise de descasamento. A tensão nos terminais é considerada constante.	41

Figura 3.5: Corte longitudinal da região ativa do <i>MOSFET</i> , onde as flutuações nas cargas de inversão e depleção, causadas pelas flutuações na concentração de dopantes, estão representadas de forma exagerada.....	42
Figura 3.6: Circuito para ensaio experimental - M_{REF} é um transistor de referência, enquanto M_i e M_{i+1} são os transistores sob teste. I_B ($V_B, V_D'=V_D$) é uma fonte de corrente (tensão) para polarização.....	47
Figura 3.7: Descasamento normalizado da corrente para o arranjo <i>grande</i> NMOS. A junção substrato-fonte foi mantida a zero volt. As curvas do modelo resultam de (3.6.19)...	49
Figura 3.8: Descasamento normalizado da corrente para o arranjo <i>médio</i> NMOS. A junção substrato-fonte foi mantida a zero volt. As curvas do modelo resultam de (3.6.19)...	49
Figura 3.9: Descasamento normalizado da corrente para o arranjo <i>grande</i> PMOS. A junção substrato-fonte foi mantida a zero volts. As curvas do modelo resultam de (3.6.19).	50
Figura 3.10: Descasamento normalizado da corrente para o arranjo <i>médio</i> PMOS. A junção substrato-fonte foi mantida a zero volts. As curvas do modelo resultam de (3.6.19).	50
Figura 3.11: Descasamento normalizado da corrente medido (méd.) na região linear ($V_{DS}=20\text{mV}$), em função do nível de inversão, para os arranjos <i>grande</i> , <i>médio</i> e <i>pequeno</i> NMOS, sob duas tensões de substrato-fonte. As curvas do modelo (mod.) resultam de (3.6.19).....	52
Figura 3.12: Descasamento normalizado da corrente medido (méd.) na saturação ($V_{DS}=2\text{V}$), em função do nível de inversão, para os arranjos <i>grande</i> , <i>médio</i> e <i>pequeno</i> NMOS, sob duas tensões de substrato-fonte. As curvas do modelo (mod.) resultam de (3.6.19).	52
Figura 3.13: Descasamento normalizado da corrente medido (méd.) nas regiões linear e de saturação, em função do nível de inversão, para os arranjos <i>estrito</i> e <i>curto</i> NMOS, sob duas tensões de substrato-fonte. As curvas do modelo (mod.) resultam de (3.6.19).	54
Figura 3.14: Descasamento normalizado da corrente medido (méd.) nas regiões linear e de saturação, em função do nível de inversão, para os arranjos <i>grande</i> , <i>médio</i> e <i>pequeno</i> PMOS, sob duas tensões de substrato-fonte. As curvas do modelo resultam de (3.6.19).	55
Figura 3.15: Descasamento normalizado da corrente medido (méd.) nas regiões linear e de saturação, em função do nível de inversão, para os arranjos <i>estrito</i> e <i>curto</i> PMOS, sob duas tensões de substrato-fonte. As curvas do modelo (mod.) resultam de (3.6.19).	55
Figura 3.16: Representação do transistor <i>MOS</i> através de dois transistores, em série com a região onde ocorre a flutuação (a), e seu equivalente para pequenas variações (b)....	61
Figura 3.17: Representação do transistor <i>MOS</i> através de dois transistores em série (a), e seu equivalente para pequenas variações (b).....	62
Figura 3.18: Efeito da saturação na velocidade dos portadores (ESVP) sobre o descasamento, em função do nível de inversão (i_f) e da condição de saturação (V_{DS} variando entre 20 mV e 2 V), para um transistor NMOS na tecnologia <i>TSMC 0.35</i> . Os símbolos maiores representam um transistor com $L = 2\lambda$ (canal mínimo) e os menores, um transistor com $L = 8\lambda$	66
Figura 3.19: Efeito da modulação da região de depleção (EMRD), em função do nível de inversão (i_f) e da condição de saturação (V_{DS} variando entre 20 mV e 2 V), de um transistor NMOS. EMRD resulta da razão entre a eq. (3.9.17) e a eq. (3.9.16). Nesta curva utilizou-se parâmetros extraídos da tecnologia <i>TSMC 0.35</i>	71
Figura 3.20: Modelo do capacitor <i>MOS</i> , alterado para incorporar o efeito de depleção que ocorre no poli-silício da <i>porta</i>	73

Figura 3.21: Simulação de (3.10.1), para: $V_{FB} = -1$ V, $C'_{ox} = 445$ nF/cm ² , $N_a = 2 \times 10^{17}$ cm ⁻³ , para dois valores de N_p : 5×10^{18} cm ⁻³ (marca x) e 5×10^{19} cm ⁻³ (sem marca).....	75
Figura 3.22: Efeito da rugosidade das bordas do poli-silício da <i>porta</i> , sobre a definição da região do canal (extraído de [58]).....	78
Figura 4.1: Polarização do transistor <i>NMOS</i> sob teste (a) e curvas que devem ser obtidas no procedimento de caracterização elétrica, visando a estimativa do descasamento (b). 83	83
Figura 4.2: Polarização do transistor sob teste (a) direta e (b) com um transistor de referência (V_B e V'_B são uma mesma fonte).	86
Figura 4.3: Estratégias de estímulo em tensão do dispositivo.....	86
Figura 4.4: Exemplo de transistores de teste dispostos em pares (<i>differential pooling</i>) e polarizados por um transistor de referência. A seleção do par a ser medido é interna ao circuito, e realizada pelo terminal de <i>porta</i>	87
Figura 4.5: Microfotografia da primeira versão do circuito de teste, implementado na tecnologia <i>AMIS 1,5 μm</i> . É formado por cinco arranjos de geometrias diferentes: (a) grande, (b) curto, (c) pequeno, (d) médio e (e) estreito. Os arranjos são replicados em transistores <i>NMOS</i> e <i>PMOS</i>	88
Figura 4.6: Diagrama de dois grupos (N e P) de transistores de teste.....	89
Figura 4.7: Sistema proposto para caracterização de descasamento de transistores <i>MOS</i> . 92	92
Figura 4.8: Conceito da topologia proposta em <i>complete pooling</i> (apenas tipo N desenhado), incluindo as chaves de seleção de <i>dreno</i> e de <i>porta</i>	92
Figura 4.9: Estratégia de caracterização da corrente de dreno em <i>differential pooling</i> , em dois momentos, com o emprego da reversão das UEMs, visando à redução do efeito do erro do descasamento existente entre os canais do equipamento, sobre o descasamento entre os dispositivos a ser medido	93
Figura 4.10: Topologia proposta em <i>differential pooling</i> (apenas transistores tipo N apresentados). As chaves de seleção de <i>drenos</i> , que possibilitam o acesso a cada um dos arranjos de transistores idênticos (na vertical), foram implementadas para permitir o uso da técnica <i>force-sense</i> . Esta técnica também é utilizada no acesso aos terminais de <i>fonte</i> . Em cada arranjo, um par específico pode ser selecionado pela chave de <i>porta</i> (S_{Gi}). As chaves S_{Bj} permitem a seleção do transistor de referência de interesse.....	94
Figura 4.11: Proposta de disposição dos transistores de teste em grupos de 6x6 pares.....	96
Figura 4.12: Proposta de planta-baixa para o circuito.	97
Figura 4.13: Proposta de leiaute de um grupo <i>PMOS</i> , mostrando a disposição dos pares de transistores de teste e conexões. O leiaute do grupo <i>NMOS</i> é análogo.	98
Figura 4.14: Chaveamento do transistor de teste pela <i>porta</i>	99
Figura 4.15: Chaveamento do transistor de teste pelo <i>dreno</i>	100
Figura 4.16: Chaveamento do transistor de referência.....	102
Figura 4.17: Circuito esquemático do registrador tipo <i>latch</i>	103
Figura 4.18: Microfotografia da segunda versão do circuito de teste, implementado na tecnologia <i>TSMC 0.35</i> , em uma pastilha com 1,5 mm de lado. Cada conjunto <i>NMOS</i> e <i>PMOS</i> é formado por 9 grupos de geometrias diferentes, cada um composto por 36 pares de transistores idênticos.	105
Figura 4.19: Medidas $I_D \times V_{DS}$ (x - cima) e $\sigma(I_D) \times V_{DS}$ (o - baixo) de um grupo <i>NMOS</i> 04x04 na tecnologia <i>TSMC 0.35</i> . As medidas foram realizadas com V_{DS} variando entre 20 mV e 2 V, para os níveis de inversão 0,01; 0,1; 1; 10; 100 e 1000 (de baixo para cima). As linhas contínuas representam a estimativa teórica de I_D , calculada através do modelo <i>ACM</i> , e do descasamento médio absoluto $\sigma(I_D)$, calculado através do modelo aqui apresentado (os parâmetros de modelo I_{SQ} , N_{oi} e B_{ISQ} foram extraídos experimentalmente da tecnologia).....	108

- Figura 4.20: Descasamento dos grupos 16x16 (12 μ m / 8 μ m - cima), 16x04 (12 μ m / 2 μ m - centro) e 16x01 (12 μ m / 0,5 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.35*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000..... 110
- Figura 4.21: Descasamento dos grupos 04x16 (3 μ m / 8 μ m - cima), 04x04 (3 μ m / 2 μ m - centro) e 04x01 (3 μ m / 0,5 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.35*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. No grupo 04x16 NMOS o nível de inversão 0,01 não está representado..... 111
- Figura 4.22: Descasamento dos grupos 01x16 (0,75 μ m / 8 μ m - cima), 01x04 (0,75 μ m / 2 μ m - centro) e 01x01 (0,75 μ m / 0,5 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.35*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. Nos grupos 01x16 e 01x04 o nível de inversão 0,01 não está representado. 112
- Figura 4.23: Descasamento dos grupos 16x16 (4,8 μ m / 3,2 μ m - cima), 16x04 (4,8 μ m / 0,8 μ m - centro) e 16x01 (4,8 μ m / 0,2 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.18*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. No grupo 16x16 o nível de inversão 0,01 não está representado. 117
- Figura 4.24: Descasamento dos grupos 04x16 (1,2 μ m / 3,2 μ m - cima), 04x04 (1,2 μ m / 0,8 μ m - centro) e 04x01 (1,2 μ m / 0,2 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.18*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. No grupo 04x16 e no 04x04 PMOS o nível de inversão 0,01 não está representado. 118
- Figura 4.25: Descasamento dos grupos 01x16 (0,3 μ m / 3,2 μ m - cima), 01x04 (0,3 μ m / 0,8 μ m - centro) e 01x01 (0,3 μ m / 0,2 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.18*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. Nos grupos 01x16 e 01x04 o nível de inversão 0,01 não está representado. No grupo 01x16 PMOS o nível de inversão 0,1 também não está representado. 119
- Figura 4.26: Descasamento medido dos grupos NMOS 16x16 em *TSMC 0.35* (esq.) e *TSMC 0.18* (dir.), com (o) e sem (+) o uso da reversão dos canais. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. 120
- Figura 4.27: Corrente de “fuga” I_{OFF} , obtida de cada uma das nove geometrias com todos os 36 pares de transistores desativados (*off state*) através de $V_{GS} = 0V$. Resultados da tecnologia *TSMC 0.35* para o conjunto NMOS (cima) e PMOS (baixo). 121
- Figura 4.28: Corrente de “fuga” I_{OFF} , obtida de cada uma das nove geometrias com todos os 36 pares de transistores desativados (*off state*) através de $V_{GS} = 0V$. Resultados da tecnologia *TSMC 0.18* para o conjunto NMOS (cima) e PMOS (baixo). 122
- Figura 4.29: Curvas $I_D \times V_{DS}$ para os níveis de inversão 1, 10, 100 e 1000 (de baixo para cima), de todos os transistores do grupo NMOS 04x04, na tecnologia *TSMC 0.35*. 123

Figura 4.30: Corrente de <i>dreno</i> (I_D), medida dos grupos geométricos 16x16 (a, d), 04x04 (b, e), e 01x01 (c, f), operando em saturação sob duas condições de inversão: $i_f = 1$ (a, b, c) e $i_f = 100$ (d, e, f). Cada coluna sólida representa a corrente de um dos transistores que formam a matriz do grupo. Resultados obtidos na tecnologia TSMC 0.35.	125
Figura 5.1: Forma geral do divisor de corrente.	128
Figura 5.2: Dois transistores NMOS idênticos e sob mesma polarização formam um divisor de corrente por 2.	129
Figura 5.3: Equivalência entre um transistor MOS e uma associação série-paralelo de transistores.	130
Figura 5.4: Decomposição de uma associação paralela de transistores MOS (fig. 5.2) em seu equivalente, através de associações série-paralelo.	130
Figura 5.5: Diagrama esquemático do conversor D/A de 8 bits, cuja relação binária das parcelas de corrente é obtida através de uma rede, composta por associações série-paralelo de transistores MOS (rede M-2M). O valor digital, a ser convertido em analógico, é programado em um registrador de deslocamento.	131
Figura 5.6: Descasamento de corrente normalizado de um transistor com $10\ \mu\text{m} \times 10\ \mu\text{m}$, sob dois níveis de inversão diferentes (20 e 2000), usando o modelo apresentado com os parâmetros extraídos de um lote da tecnologia TSMC 0.35. As linhas tracejadas correspondem a $N_{oi} = 2 \times 10^{12}\ \text{cm}^{-2}$ e $B_{ISQ} = 0,4\ \text{\%}\text{-}\mu\text{m}$. As linhas sólidas correspondem a $N_{oi} = 3,5 \times 10^{12}\ \text{cm}^{-2}$ e $B_{ISQ} = 0,9\ \text{\%}\text{-}\mu\text{m}$. A tensão <i>dreno-fonte</i> (V_{DS}) varia logaritmicamente entre 2 V e 20 mV.	136
Figura 5.7: Microfotografia dos conversores fabricados: DAC0 (esq.) e DAC1 (dir.). A área superior é a rede M-2M, onde se observa claramente a matriz de 4×10 transistores, cercada pelo anel de guarda e <i>dummies</i> . Sob essa área vê-se claramente os 8 registradores, as chaves de acionamento da rede e dois capacitores de desacoplamento de alimentação. No DAC0, ao lado esquerdo dos registradores, observa-se o buffer do sinal <i>Ck</i>	137
Figura 5.8: Desvio-padrão do erro medido das 20 amostras de DAC0 (esq.) e DAC1 (dir.), para todos os dados de entrada, e normalizado para 1 LSB. As medidas foram realizadas sob os níveis de inversão 20 e 2000, e com $K_I = 1,9$	139
Figura 5.9: Amostras de DAC0 (sup.) e DAC1 (inf.) que apresentaram os valores mínimo e máximo de erro medido, sob os dois níveis de inversão extremos, 20 (esq.) e 2000 (dir.). As medidas estão normalizadas para 1 LSB e foram realizadas com $K_I = 1,9$	140
Figura A.1: Associações paralelo e série de transistores MOS.	145
Figura A.2: Equivalente paralelo de um transistor MOS e seus circuitos de polarização e de variações.	146
Figura A.3: Equivalente série de um transistor MOS e seus circuitos de polarização e de variações.	147
Figura A.4: Inconsistência do modelo de Pelgrom para inversão forte (SI) e fraca (WI), nas regiões linear e saturada.	150
Figura A.5: Inconsistência do modelo de Pelgrom para diferentes níveis de inversão em saturação.	150
Figura B.1: Corte vertical da região ativa do MOSFET, onde ficam evidentes a espessura do óxido (t_{ox}) e o substrato. A partir da interface óxido-substrato ($y = 0$), ocorre a região de depleção, que se estende até a profundidade y_d . Em uma pequena fatia da região de depleção (Δy) há uma flutuação na sua carga, causada pela flutuação local na concentração de dopantes. A dimensão Δx representa uma pequena porção do	

comprimento do canal. A dimensão não apresentada corresponde à largura W do canal.	154
Figura B.2: Comportamento do campo elétrico E e do potencial ϕ , no sentido da profundidade do substrato. As linhas tracejadas correspondem à situação da concentração de dopantes homogênea, e as linhas sólidas, à situação onde há uma flutuação na carga de depleção na posição y	154

Lista de Tabelas

Tabela 4.1: Dimensões das estruturas de teste (em μm), implementadas na versão 1.....	89
Tabela 4.2: Dimensões dos transistores de teste, implementados na versão 2. Sem parênteses estão as dimensões em μm para a tecnologia <i>TSMC 0.35</i> ($\lambda = 0,25\mu\text{m}$), e entre parênteses, para a <i>TSMC 0.18</i> ($\lambda = 0,10\mu\text{m}$).....	96
Tabela 4.3: Principais parâmetros da tecnologia <i>TSMC 0.35</i> , obtidos no site <i>www.mosis.org</i> , para a rodada T3BM (LO_EPI). O valor de n foi estimado através de simulações com o modelo BSIM 3V3 fornecido no mesmo site. A corrente específica por quadrado - I_{SQ} - foi calculada através dos parâmetros da tecnologia.	99
Tabela 4.4: Número do pino de acesso, nome e significado, referentes ao encapsulamento <i>DIP28</i> utilizado. A letra, entre parênteses, informa se o pino deve ser utilizado para estímulo (E) ou medida (M).	105
Tabela 4.5: Principais parâmetros das tecnologias <i>TSMC 0.35</i> (rodada T52F MM_EPI) e <i>TSMC 0.18</i> (rodada T55U MM_NON-EPI), utilizadas na fabricação do circuito. ...	107
Tabela 4.6: Valores dos parâmetros do modelo de descasamento utilizados nas figuras 4.20 a 4.22. Foram calculados através de ajuste feito pelo método de mínimos quadrados, aplicado às curvas de descasamento obtidas experimentalmente da rodada T52F, utilizada na fabricação do circuito de teste na tecnologia <i>TSMC 0.35</i>	113
Tabela 4.7: Valores dos parâmetros de descasamento, referentes ao modelo de Pelgrom, obtidos através de três versões da literatura técnica do processo <i>CMOS AMS 0,35 μm</i>	115
Tabela 4.8: Valores dos parâmetros do modelo de descasamento utilizados nas figuras 4.23 a 4.25. Foram calculados através de ajuste feito pelo método de mínimos quadrados, aplicado às curvas de descasamento obtidas experimentalmente da rodada T55U, utilizada na fabricação do circuito de teste na tecnologia <i>TSMC 0.18</i>	120
Tabela 4.9: Corrente específica (I_{SQ}) e fatores geométricos (dL e dW), obtidos através de ajuste por mínimos quadrados das curvas experimentais na tecnologia <i>TSMC 0.35</i>	124
Tabela 4.10: Corrente específica (I_{SQ}) e fatores geométricos (dL e dW), obtidos através de ajuste por mínimos quadrados das curvas experimentais na tecnologia <i>TSMC 0.18</i>	125
Tabela 5.1: Comprimento mínimo do canal, em função do erro máximo aceitável.....	134

1 INTRODUÇÃO

O presente capítulo discorre sobre as particularidades envolvidas no desenvolvimento dos circuitos integrados (CIs) atuais, visando situar o leitor no tema a ser abordado.

1.1 Circuitos eletrônicos analógicos e digitais

Por circuito eletrônico analógico, entende-se o circuito onde a informação contida em um sinal que trafega pelo mesmo é representada através de uma faixa contínua de valores de tensão ou corrente. Em contraposição aos circuitos analógicos, tem-se os chamados circuitos digitais, caracterizados pelo fato de essa informação ser codificada por um número finito de níveis (informação discretizada), representados através de valores estanques de tensão ou corrente.

Pelo aspecto do modo de operação, pode-se estabelecer a diferença entre os circuitos digitais e analógicos pelo fato de que, para os primeiros, os elementos “ativos”¹ (fundamentalmente os transistores e seus derivados) geralmente operam apenas nas condições de menor dissipação de potência (corrente alta e tensão baixa ou o contrário – na tecnologia bipolar essas duas condições são chamadas *saturação* e *corte*, enquanto que na tecnologia *MOS*² são chamadas *linear* e *corte*), passando pela situação de maior dissipação apenas transitoriamente. Já para os circuitos analógicos, esses elementos podem operar em qualquer região, situando-se normalmente em condições de maior dissipação (região *ativa direta*).

Historicamente, nota-se que desde os anos 60, com o surgimento dos circuitos digitais integrados, as aplicações antes dominadas pela eletrônica analógica vêm sendo

¹ O termo elemento “ativo” é aqui empregado significando o elemento que permite o controle sobre os parâmetros elétricos de um circuito, notadamente corrente ou tensão, conduzindo à amplificação.

² A sigla *MOS* significa metal-óxido-semicondutor e foi cunhada para designar a tecnologia de transistores de efeito de campo de *porta* isolada, referindo-se à seqüência de camadas originalmente empregada na sua fabricação.

progressivamente adaptadas ao domínio digital. Algumas das razões que fomentam esse avanço digital são a maior possibilidade de integração, maior adaptabilidade dos circuitos a alterações externas, menor exigência quanto aos processos de fabricação e maior facilidade de projeto.

Com o advento da tecnologia *MOS* (origem do *MOSFET*³) na indústria dos anos 70, e posteriormente da tecnologia *CMOS*⁴, esse avanço tornou-se ainda mais veloz, visto que essa tecnologia era mais adequada aos circuitos digitais e propiciava menores custos de fabricação, maior escala de integração e menor consumo de potência dos circuitos, permitindo a incorporação cada vez maior de funcionalidades em um único CI.

Ainda que a tecnologia digital tenha sobrepujado a analógica com grande margem de vantagens e em uma vasta gama de aplicações, esta última jamais será extinta, visto que é necessária alguma interface dos sistemas digitais com o mundo, que é analógico. Também, em alguns nichos de aplicações, a tecnologia analógica ainda apresenta uma melhor relação custo/benefício (*rádio-de-pilha*, por exemplo).

Todos esses elementos, alinhados à explosão de mercado para produtos eletrônicos que se iniciou nos anos 70-80, levaram à total predominância dos circuitos *CMOS* como a base tecnológica da indústria eletrônica que se tem na atualidade.

Nos anos 90, as necessidades mercadológicas de circuitos de aplicação específica (*ASICs – application-specific integrated circuits*) concomitantes a um mercado de alto volume de consumo, impulsionaram a incorporação dos circuitos analógicos e digitais de um mesmo sistema em um único CI, denominado *SoC (system-on-a-chip*, ou sistema-em-uma-pastilha). Por questões de custo de fabricação, e levando-se em conta que os circuitos digitais compõem hoje a maior parte de um sistema eletrônico, a tendência mundial foi de se optar pela adoção da tecnologia *CMOS* também para a fabricação das partes analógicas (exceção feita a algumas aplicações específicas, onde se adota a tecnologia *BiCMOS*, que compatibiliza dispositivos bipolares e *CMOS* na mesma pastilha). Estes circuitos incorporam sistemas completos analógico-digitais multifuncionais, operando de baixas a altas frequências, compostos de vários milhões de transistores em uma única pastilha *CMOS* (como exemplo, cita-se os circuitos receptores de TV em um único CI).

³ *MOSFET* é a sigla de *metal-oxide-semiconductor field-effect-transistor*, e diz respeito ao transistor de efeito de campo fabricado na tecnologia *MOS*.

⁴ Tecnologia *MOS* complementar, onde são fabricados transistores tipo N e P no mesmo substrato. Atualmente é comum o uso da sigla *MOS* também com o significado de *CMOS*.

Entretanto, o progressivo escalamento da tecnologia *CMOS*, reduzindo suas dimensões geométricas a valores sub-micrométricos, bem como a redução na tensão de alimentação (que favorece a queda de consumo), fatores que tanto impulsionam o aumento da complexidade dos circuitos digitais, acabam por trazer novas dificuldades e desafios ao projeto dos circuitos analógicos.

Dentro do contexto apresentado percebe-se a grande importância da pesquisa e desenvolvimento de novas técnicas e ferramentas que possibilitem o projeto e a implementação de circuitos integrados em tecnologia *CMOS* de forma mais rápida e segura.

1.2 Projeto de circuitos integrados

O que genericamente se denomina *projetar* um circuito integrado pode ser traduzido como a tarefa de se idealizar, conceber, dimensionar, verificar, simular e desenhar um circuito para ser fabricado, a partir de um conjunto de especificações, que podem ser funcionais, elétricas ou tecnológicas⁵. Ainda, as tarefas de teste e validação são compreendidas como partes integrantes das atividades do projetista, após o circuito ter sido fabricado.

Uma das tarefas mais complexas durante esse processo é a de se conseguir idealizar uma configuração de circuito que coadune as diversas especificações entre si, de forma não conflitante. Ou que, ao menos, atenuar os conflitos entre as especificações de forma aceitável.

No campo dos circuitos digitais, geralmente busca-se soluções que otimizem as especificações de *velocidade* e *dissipação de potência*.

Dentre os circuitos analógicos, entretanto, o projetista tem de lidar com um grupo maior de parâmetros, o que torna seu trabalho bem mais complexo. Geralmente, pode-se listar esses parâmetros como sendo⁶: *ganho*, *linearidade*, *ruído*, *tensão de alimentação*, *excursão de sinal*, *velocidade*, *impedâncias de entrada e saída* e *dissipação de potência*. Este grupo de parâmetros forma o *octógono de projeto analógico*, apresentado em [1].

⁵ Aqui, especificação tecnológica é posta no sentido das restrições impostas pela tecnologia-alvo do projeto.

⁶ Dependendo da especificidade do projeto, outros parâmetros podem ser incluídos nessa lista, tendo-se destacado apenas os mais gerais descritos em [1].

Em conjunto com as métricas apresentadas, o projeto de um circuito integrado deve ainda conciliar outros dois fatores de desempenho: *área* (de silício ocupada pelo circuito) e *casamento* (*matching*) entre dispositivos. Menor área representa menor custo de fabricação, e maior casamento se reflete em maior repetibilidade e exatidão no comportamento dos circuitos fabricados, reduzindo as perdas do processo.

Conclui-se assim que o desenvolvimento e o melhoramento de ferramentas que permitam ao projetista a rápida e simples avaliação dos parâmetros apresentados, ainda na fase de concepção e dimensionamento de um circuito, são sempre bem-vindos, pois isso aumenta as garantias de sucesso e reduz os esforços (e custos) de um projeto.

1.3 Descasamento entre transistores MOS

As topologias de circuitos eletrônicos integrados usuais, digitais ou analógicos, baseiam-se em grande parte no conceito de similaridade comportamental (casamento) entre dispositivos identicamente fabricados. Ou seja, supõe-se que dois (ou mais) dispositivos que sejam identicamente desenhados, e fabricados simultaneamente no mesmo processo, apresentem um comportamento elétrico idêntico. Entretanto, uma vasta gama de tipos de flutuações e variações incontroláveis, presentes em todas as etapas do processo de fabricação de um circuito integrado, acrescenta um grau de diferença de comportamento aos dispositivos fabricados, denominado *descasamento*.

O descasamento entre dispositivos *MOS* tem sido estudado e modelado há mais de três décadas [2], [3], [4], [5], [6], [7], [8], [9], [10], tendo assumido papel de grande relevância ao longo dos últimos anos, dado o crescente impacto desse efeito no desempenho dos circuitos integrados, em função da progressiva redução das dimensões físicas (escalamento) e da tensão de alimentação que tem ocorrido nesta tecnologia.

Nos circuitos digitais, o descasamento entre transistores leva a flutuações no tempo de atraso entre *portas* lógicas (figura 1.1), ao aumento das correntes de fuga e ao aumento dos erros nos processos de escrita e leitura em memórias [9], [11], [12]. Todos esses fatores aumentam as perdas em um processo industrial de fabricação de circuitos integrados.

No caso dos circuitos analógicos, o impacto é bem mais severo, reduzindo a exatidão de conversores analógico-digitais (*ADC* – figura 1.2) e digital-analógicos (*DAC*),

degradando a estabilidade de fontes de referência e aumentando a tensão de *off-set* de amplificadores operacionais (figura 1.3), por exemplo [6], [7], [10], [13].

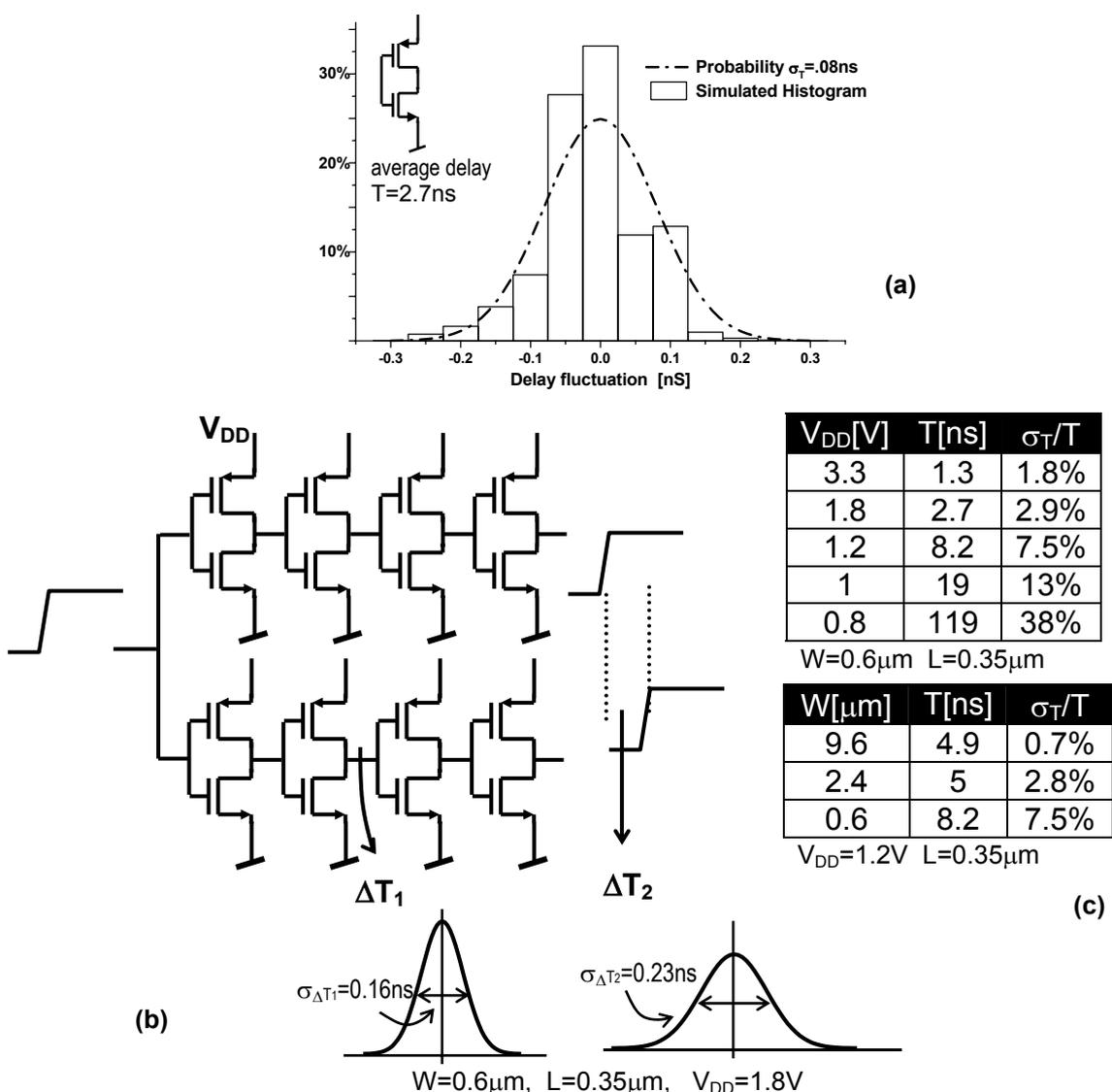


Figura 1.1: Comportamento estatístico do atraso de dois ramos de distribuição de sinal de *clock* em processo de $0,35\mu\text{m}$, causado pelo descasamento dos transistores. O histograma (a) resulta da simulação do atraso de 200 amostras, para um único inversor de dimensão mínima e com $V_{DD}=1.8\text{V}$. Flutuações do atraso acumulado em duas cadeias de 4 inversores em (b), cuja distribuição se amplia à medida que o sinal avança na cadeia (ΔT_1 : 2 inversores e ΔT_2 : 4 inversores). A tabela (c) apresenta a distribuição do atraso, normalizada pelo atraso médio de um inversor unitário (σ_T/T), para diferentes dimensões de transistores e valores de tensão de alimentação. Observa-se que a dispersão do atraso aumenta à medida que a tensão de alimentação e o tamanho dos transistores diminuem, os dois grandes desafios na evolução da tecnologia moderna.

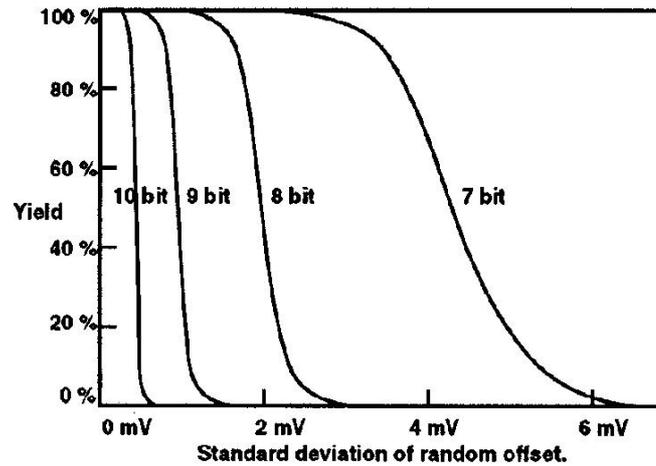


Figura 1.2: Percentual de aproveitamento (*yield*) de ADCs fabricados, em função do desvio-padrão na tensão de *off-set* (descasamento) do par de transistores de entrada (extraído de [10]).

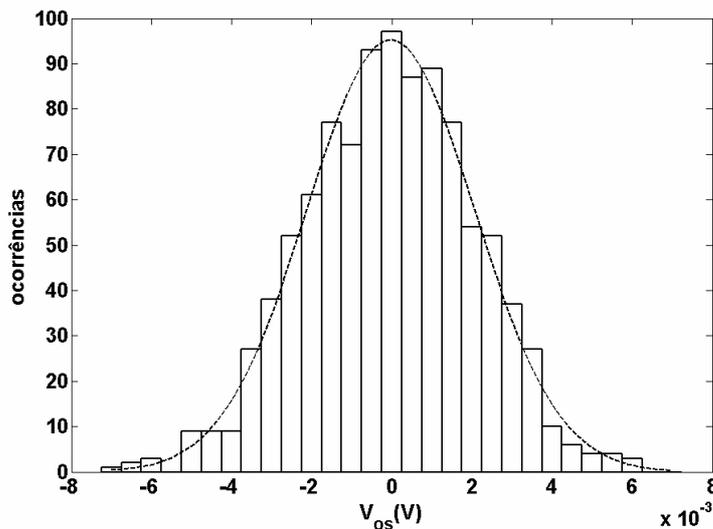


Figura 1.3: Simulação *Monte Carlo* da tensão de *off-set* de um amplificador operacional *Miller CMOS*. O histograma apresenta a distribuição desta tensão sobre 1000 amostras, em intervalos de 0,5 mV. O desvio-padrão calculado é 2,1 mV. A curva tracejada é a sua aproximação Gaussiana.

Este trabalho tem como objetivo o estudo teórico e experimental do descasamento entre transistores *MOS* (figura 1.4) e sua modelagem, visando contribuir de forma significativa no domínio desse efeito, por parte do projetista de circuitos integrados analógicos ou digitais.

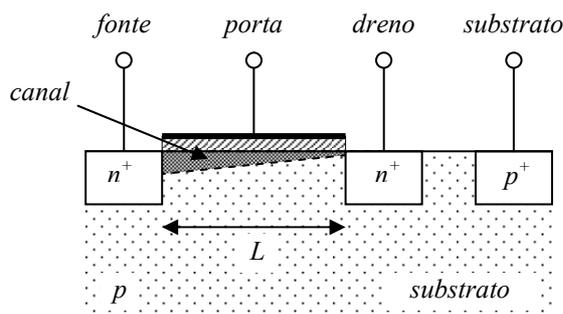


Figura 1.4: Diagrama esquemático da seção lateral de um transistor *MOS* do tipo canal N.

1.4 Causas e tipos de descasamento

O descasamento entre transistores *MOS* pode ser definido como o resultado de um conjunto de variações atemporais no seu comportamento elétrico, observado entre dois ou mais dispositivos identicamente desenhados e fabricados. Tais variações de comportamento são decorrência de variações físicas que ocorrem entre estes transistores nas várias etapas do processo de fabricação. Essas variações físicas podem ocorrer entre dispositivos de uma mesma pastilha (*intradie*), ou de pastilhas diferentes (*interdie*). No caso de pastilhas diferentes, podem ainda ocorrer entre dispositivos de uma mesma lâmina (*intrawafer*), de um mesmo lote de fabricação (*inrabatch*) ou ainda entre dispositivos fabricados em diferentes lotes (*interbatch*). Obviamente, quanto maior o domínio observado, maior a amplitude dessas variações.

Geralmente, o estudo do descasamento entre transistores *MOS* se restringe ao caso em que estes dispositivos estão em uma mesma pastilha (*intradie*). Isso decorre da estratégia de projeto adotada, onde os dispositivos em que se deseja casamento dentro de um circuito são dimensionados na forma de razões mais do que em termos absolutos.

As principais variações físicas observadas em lâminas fabricadas em tecnologia *MOS* são:

- flutuações na concentração e no perfil de dopantes no substrato e nas camadas eletricamente ativas;
- variações e flutuações na espessura e na qualidade (contaminação, solidificação não-uniforme, etc) das diversas camadas crescidas ou depositadas sobre o substrato (óxido, poli-silício e metal);
- formação de aglomerados de dopantes no poli-silício (segregação);

- variações nas dimensões laterais das formas geométricas das camadas depositadas ou decapadas;
- rugosidade de borda das formas geométricas das camadas depositadas ou decapadas;
- flutuações na qualidade do contato elétrico (resistividade, barreiras de potencial, etc) entre camadas condutoras diferentes (substrato, poli-silício e metal);
- flutuações no aprisionamento de cargas elétricas no interior do óxido ou na sua interface com o substrato;
- tensões mecânicas permanentes na superfície do substrato (p. ex. devido à forma de encapsulamento).

O modo como as variações físicas afetam o comportamento dos transistores em uma mesma lâmina é fundamentalmente dependente da relação entre as dimensões físicas (tamanho) dos dispositivos, e o *comprimento de onda espacial* das variações ao longo da mesma. Nesse aspecto, as variações físicas e seus efeitos sobre o descasamento são classificados em dois tipos: *globais* (aquelas relacionadas às variações físicas cujo comprimento de onda espacial é muito superior às dimensões físicas dos transistores) e *locais* (aquelas relacionadas às variações físicas cujo comprimento de onda espacial é muito inferior às dimensões físicas dos transistores).

Os fatores que provocam essas variações podem ser classificados em dois tipos, em função do modo como afetam os dispositivos fabricados:

- fatores *sistemáticos*: geralmente relacionados a variações ou deformações espaciais, que ocorrem de forma sistemática em componentes ou elementos do processo de fabricação. Decorrem de efeitos físicos e químicos incontroláveis, como a dilatação térmica nos equipamentos, as aberrações nas lentes e distorções nas máscaras de fotolitografia, as mudanças na concentração de substâncias (de ataque, deposição ou dopagem), as não-linearidades no equipamento de implante de íons e as flutuações térmicas em geral. São fatores que provocam gradientes nas propriedades físicas ao longo da pastilha, da lâmina ou do lote;
- fatores *estocásticos*: em geral relacionados à natureza discreta da matéria, provocando flutuações locais (microscópicas) nas propriedades físicas dos dispositivos fabricados.

De um modo geral, para dispositivos próximos entre si, pode-se dizer que os fatores de descasamento sistemático estão relacionados aos efeitos do tipo global, e os fatores estocásticos estão relacionados aos efeitos do tipo local.

1.5 Técnicas de redução dos efeitos globais

Considerando-se uma região da pastilha com dimensões suficientemente inferiores ao comprimento de onda espacial das variações globais, essas variações devem ocorrer dentro da região na forma de um gradiente. Usando esse conceito em mente, algumas estratégias de leiaute foram desenvolvidas [14], visando a redução destes efeitos no descasamento entre transistores *MOS*, e que são apresentadas a seguir.

- Em primeiro lugar, a geometria dos dispositivos a serem casados, bem como dos elementos usados na conexão destes ao circuito, deve ser idêntica e deve-se utilizar os mesmos tipos de camadas (metal-1, metal-2, poli-silício, contatos, vias, etc). Deve-se inclusive evitar que trilhas de outras camadas (p. ex. metal) passem de forma não-idêntica sobre os dispositivos (figura 1.5);
- deve-se agrupar os dispositivos que necessitam casamento da forma mais próxima possível, visando a exposição destes à menor variação total da grandeza que exhibe gradiente (figura 1.6);
- procura-se utilizar dispositivos de pequena dimensão, visando atender à regra anterior (isso se opõe à redução do descasamento local, como será visto a seguir);
- no caso da necessidade de dispositivos de grande dimensão, deve-se fracioná-los em um número de elementos menores idênticos e associá-los em paralelo, tomando o cuidado de intercalar (em uma ou em duas dimensões) estes elementos de modo que ambos os dispositivos a serem casados (ou todos) tenham seus elementos integrantes igualmente afastados (em média) de um ponto fictício central comum. Essa técnica é conhecida como centróide-comum e garante que todos os dispositivos casados sejam submetidos, em média, igualmente aos mesmos gradientes (essa técnica só tende a compensar os efeitos de primeira ordem, ou lineares, das variações globais). A figura 1.7 mostra como o transistor M_1 é mais afetado pelo gradiente em (a), sofrendo efeito médio equivalente a M_2 em (b);

- ao redor de um grupo de dispositivos que necessita casamento, deve-se posicionar outros dispositivos idênticos e sem funcionalidade no circuito (*dummy devices*), fazendo com que todos os dispositivos úteis sejam rodeados exatamente pelo mesmo tipo e quantidade de elementos (ou seja, “enxerguem” as mesmas condições de contorno sobre a pastilha, evitando que haja condições diferentes para as etapas de ataque e deposição dos elementos a serem casados – figura 1.8). Estes dispositivos *dummy* podem ser implementados apenas na camada de poli-silício, que é aquela a que mais o transistor *MOS* é sensível, ou através da reprodução idêntica de todo o dispositivo e conexões, levando a um maior fator de segurança no projeto;
- por fim, deve-se garantir que a direção e o sentido da corrente elétrica seja exatamente o mesmo em todos os dispositivos que necessitam de casamento (figura 1.9), visto que a mobilidade elétrica na superfície do substrato pode variar com a direção.

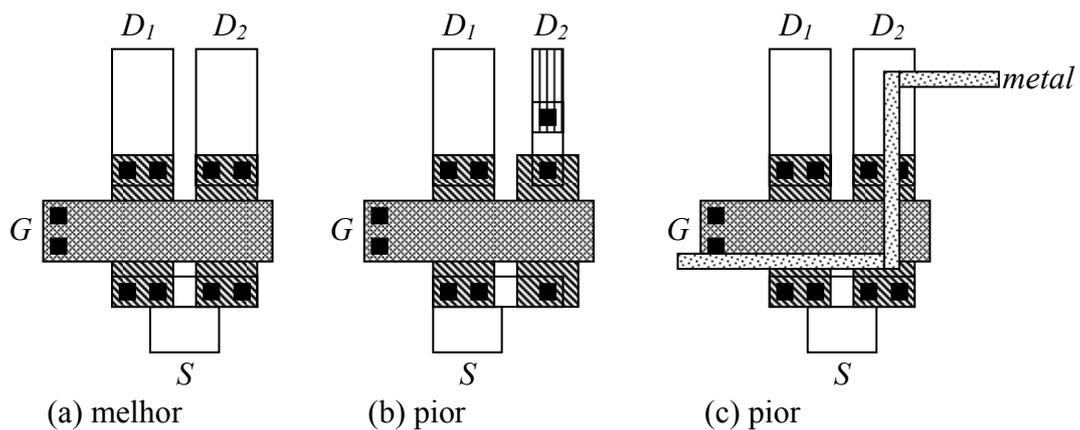


Figura 1.5: O casamento é melhor se os transistores forem idênticos.

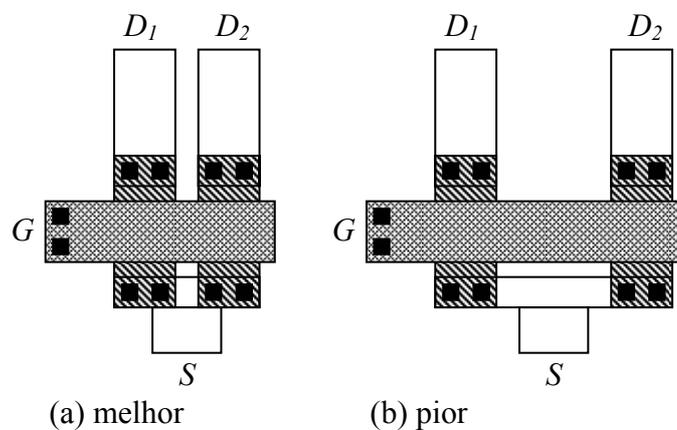


Figura 1.6: O casamento é melhor se os transistores estiverem próximos.

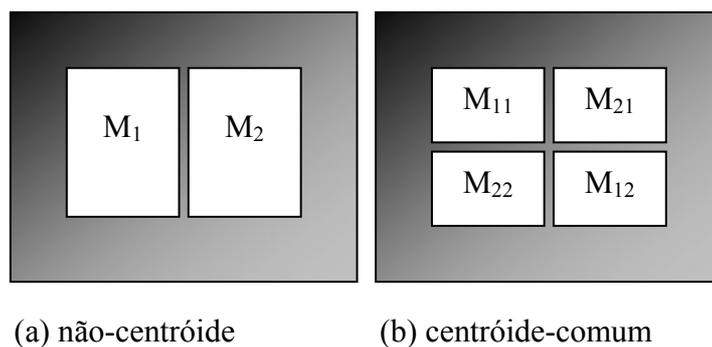


Figura 1.7: A distribuição centróide-comum (b) integra os efeitos de gradientes ao longo da área do circuito, fazendo com que, em média, os dois transistores sejam igualmente afetados.

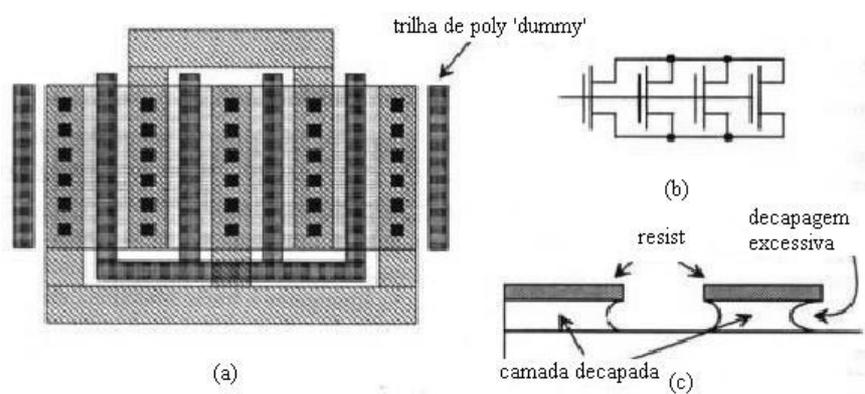


Figura 1.8: Deve-se implementar elementos ou dispositivos *dummy* ao redor dos dispositivos casados que se encontram na extremidade (a) do conjunto (b), fazendo com que estes tenham as mesmas condições de contorno durante a fabricação (c).

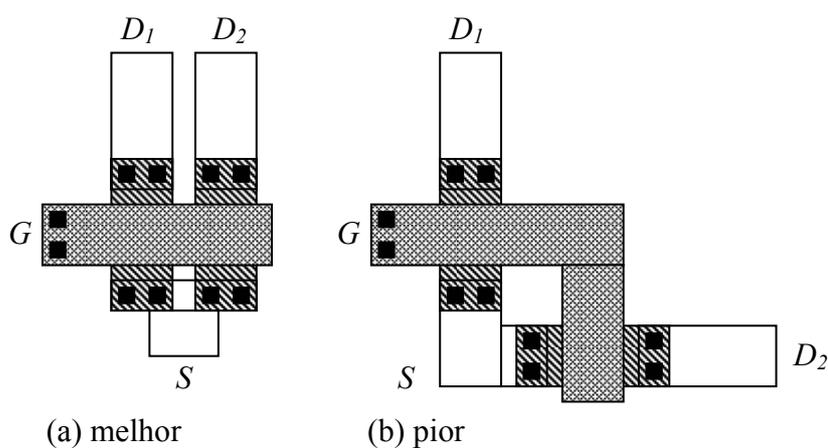


Figura 1.9: O casamento é melhor se a corrente nos transistores tiver a mesma direção.

1.6 Técnicas de redução dos efeitos locais

No caso das variações locais, considera-se que estas têm caráter estocástico, com média nula ao longo da pastilha, e que seu comprimento de onda espacial é muito inferior às dimensões dos transistores (sendo, pois, consideradas como flutuações microscópicas). Daí, pode-se deduzir que, quanto maiores as dimensões físicas de um dispositivo, maior o número de flutuações que serão integradas dentro deste, aproximando mais a variação total do seu comportamento elétrico de zero (média). Assim, percebe-se que o descasamento de uma característica elétrica de um dispositivo, devido aos efeitos locais, está relacionado ao inverso da dimensão física que se está analisando. Por exemplo, no caso do efeito de descasamento devido à rugosidade na lateral da área de poli-silício que forma o eletrodo de *porta*⁷ de um transistor *MOS*, o desvio-padrão deste efeito apresenta-se inversamente proporcional ao comprimento desta lateral (W , largura, ou L , comprimento, conforme a dimensão considerada). Mas, se for considerado o efeito de descasamento devido às flutuações na concentração de dopantes, distribuídos na região ativa do transistor (ou seja, na região do canal), o desvio-padrão passa a apresentar-se inversamente proporcional à raiz quadrada da área desta região (WL).

Dessa forma, conclui-se que o aumento nas dimensões físicas de um transistor *MOS* é um dos fatores que reduz o descasamento devido aos efeitos locais.

Uma maneira de se poder estimar o montante de descasamento entre transistores devido aos efeitos locais é se estabelecer uma relação física entre as grandezas que variam microscopicamente (concentração de dopantes, formação de aglomerados no poli-silício, rugosidade lateral e superficial, etc) e o comportamento elétrico do dispositivo. Essa relação é obtida através de um *modelo de comportamento elétrico do transistor MOS*. Entendendo-se os mecanismos por trás das principais causas de descasamento local e associando-se esses mecanismos a um modelo do transistor, pode-se deduzir um *modelo físico para o descasamento entre transistores MOS*. De posse desse modelo de descasamento, o projetista pode facilmente avaliar todos os graus de liberdade que o projeto permite no sentido de reduzir o descasamento e aumentar a repetibilidade e a exatidão do comportamento elétrico de um circuito integrado, reduzindo as perdas em um processo de fabricação (aumento de *yield*).

⁷ Um dos quatro terminais de um transistor *MOS*. Os outros são: *dreno*, *fonte* e *substrato*.

A figura 1.10 apresenta um diagrama de fluxo de projeto, que inclui a simulação estatística de desempenho de um circuito, com base em um modelo de descasamento e em parâmetros estatísticos desse processo. Como resultado pode-se avaliar o percentual de perdas na fabricação e melhorar o projeto.

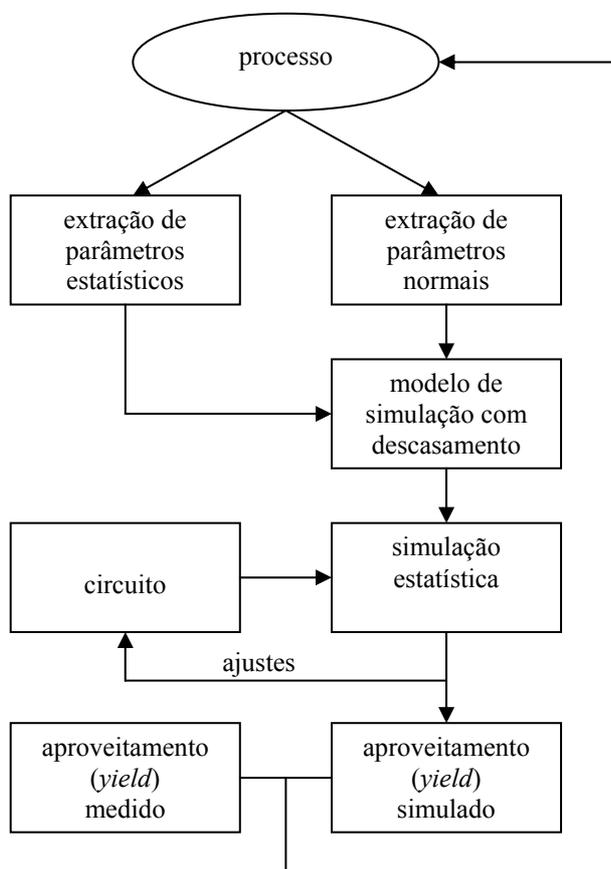


Figura 1.10: Fluxo de projeto, incluindo o modelo de descasamento (extraído de [10]).

1.7 Estrutura do texto

Este trabalho está dividido nas partes descritas a seguir.

O capítulo 1 apresenta uma introdução aos conceitos envolvidos no projeto de circuitos eletrônicos, em especial os integrados, e situa o leitor no assunto principal, que é o descasamento entre transistores *MOS* e na necessidade dos projetistas quanto ao modelamento deste efeito.

O capítulo 2 oferece uma revisão histórica e bibliográfica dos estudos relacionados ao descasamento entre dispositivos *MOS*, detalhando os avanços obtidos nesse campo e o modo como acompanharam a evolução tecnológica. Também apresenta detalhes da abordagem necessária nesse tipo de estudo, tendo em vista o desenvolvimento de um

modelo consistente e acurado, e que ofereça uma previsão teórica do descasamento sob uma ampla faixa de condições de polarização.

O capítulo 3 apresenta alguns conceitos fundamentais do modelo ACM e fundamentos da teoria estatística das distribuições aleatórias. O problema do reduzido número de átomos dopantes na região ativa dos transistores *MOS* é exposto, e o efeito das flutuações locais na carga de inversão ao longo do canal é analisado, além de suas implicações sobre a corrente de *dreno*. O modelo proposto nesta tese para o descasamento entre transistores *MOS* é exposto e detalhado, e o impacto de diversos efeitos secundários é analisado.

O capítulo 4 trata do projeto e da implementação de circuitos integrados voltados à caracterização elétrica do efeito do descasamento entre transistores *MOS*. Duas versões de circuitos de teste são apresentadas, implementadas em diversas tecnologias e caracterizadas experimentalmente, de modo a permitir que se obtenha o volume de medidas necessário para que o modelo aqui apresentado possa ser avaliado e utilizado. Um grande volume de resultados experimentais é incluído neste capítulo, além de sua interpretação.

O capítulo 5 apresenta o projeto e a implementação de um circuito conversor digital-analógico integrado, que utiliza uma rede de transistores *MOS* como elemento de ponderação binária. Sua metodologia de projeto é descrita, e o modelo de descasamento aqui apresentado é utilizado na previsão de sua inexatidão. Duas versões de 8 bits do conversor são apresentadas, implementadas e caracterizadas experimentalmente, demonstrando a utilização do modelo em uma aplicação prática.

No capítulo 6 são apresentadas e discutidas as conclusões deste trabalho.

Ao final são apresentados alguns anexos, que auxiliam o entendimento de partes do texto principal, a lista de publicações técnicas obtidas, e a lista das referências bibliográficas utilizadas.

2 MODELOS DE DESCASAMENTO

O presente capítulo apresenta uma revisão da bibliografia técnica relacionada ao descasamento entre transistores *MOS*, fornecendo suporte ao entendimento dos modelos já propostos para este efeito, bem como uma análise mais acurada do modelo de maior repercussão, e que é hoje adotado por grande parte das indústrias de circuitos integrados.

2.1 Revisão bibliográfica

O descasamento entre transistores *MOS* tem sido uma preocupação entre os pesquisadores e projetistas há mais de 3 décadas.

Em artigo publicado em 1975, Robert Keyes [2] (*IBM*) discute o limite na reprodutibilidade de transistores de efeito de campo *MOS* devido a um grau de imprevisibilidade na tensão de limiar (*threshold voltage* ou V_T) decorrente da flutuação na concentração de dopantes no substrato, problema já citado por Hoeneisen e Mead em 1972 [3] como uma possível limitação para a tecnologia *MOS*. Com o uso de uma discretização da região do canal (percolação), Keyes formaliza um modelo tentando prever a amplitude na variação de V_T para esse caso, mas que não leva em conta a condição de operação do transistor.

Com a expansão da tecnologia *MOS* na indústria de semicondutores, e o início do seu uso em circuitos analógicos, a preocupação com o casamento de dispositivos nessa tecnologia vem à tona. Em 1981, James McCreary [4] publica um artigo contendo uma análise estatístico-experimental dos erros de descasamento entre capacitores *MOS*, incluindo algumas técnicas de projeto. Tal descasamento é causa de queda de desempenho em conversores analógico-digitais, em filtros, em circuitos a capacitor chaveado e em amplificadores e atenuadores que dependam da relação entre capacitores.

O descasamento entre capacitores *MOS* é rediscutido por Shyu, Temes e Yao em artigo de 1982 [5], onde é feita uma abordagem mais formal e proposto um modelo para a dependência desse efeito nas rugosidades de borda, e nas variações de espessura e

permissividade do óxido. Neste trabalho os fatores de descasamento já são agrupados em sistêmicos e estocásticos. Em novo trabalho, publicado em 1984 [6], Shyu, Temes e Krummenacher apresentam um modelo mais completo para o descasamento de capacitores e transistores *MOS*, incluindo como fontes de erro as flutuações aleatórias nas dimensões físicas da região ativa (comprimento, largura e área) e nos parâmetros do processo de fabricação. O descasamento é analisado em função dos efeitos de rugosidade de borda, dos estados de superfície e dopantes implantados, do óxido e da mobilidade no canal, com uso das definições de variações globais e locais já sendo empregadas. Pela primeira vez a condição de operação (polarização) do dispositivo é levada em conta nesse tipo de análise através do modelo de primeira ordem do comportamento elétrico do transistor *MOS* (relação quadrática entre a corrente de *dreno* e a tensão de *overdrive*, representada por $V_{GS} - V_T$). São considerados apenas transistores de grandes dimensões e na região de saturação e *inversão forte*⁸. Na época, a condição de *inversão fraca*⁹ era praticamente desconsiderada como uma opção pelos projetistas. No artigo são apresentadas as expressões, depois consagradas, que relacionam o descasamento ao inverso da dimensão física do dispositivo a que está vinculado o fator de flutuação original (área, comprimento ou largura). Este trabalho usa de análise estatística para avaliar as variações estocásticas e é muito rico no embasamento experimental. Pode-se dizer que este é o marco inicial dos modelos de descasamento de transistores *MOS*.

Visto que nem todos os circuitos que exigem exatidão de desempenho podem ser feitos unicamente com capacitores, Lakshmikumar, Hadaway e Copeland publicam em 1986 [7] o primeiro artigo que analisa somente o descasamento entre transistores *MOS*, incluindo dispositivos de pequena dimensão na parte experimental. O descasamento é modelado com base no mesmo modelo de primeira ordem do comportamento elétrico do transistor *MOS* usado em [6], válido apenas para a situação de saturação e inversão forte. Os autores buscam um modelo de descasamento que necessite de poucos parâmetros experimentais e propõem como parâmetros naturais os desvios-padrão da tensão de limiar (σ_{VT}) e do fator de condutância (σ_K) do canal. É um artigo muito rico, onde é analisada a dependência de σ_{VT} e de σ_K com diversos tipos de flutuações: na concentração de dopantes, na mobilidade no canal, na espessura e permissividade do óxido, e nas dimensões da região

⁸ Situação onde a corrente de *dreno* é suficientemente superior a um limiar característico do processo chamado de *corrente específica* ($I_S = \frac{1}{2} \mu C'_{ox} n \phi_i^2 (W/L)$).

⁹ Situação oposta à inversão forte.

ativa. Os dados experimentais incluem transistores *NMOS* e *PMOS* escalados desde pequenas até grandes dimensões. A variedade de relações de aspecto W/L é pequena, dificultando a verificação da dependência do descasamento com os efeitos de borda. Ao final, um conversor digital-analógico (*DAC*) de 8 bits é projetado como forma de demonstrar a utilidade do modelo proposto.

Em 1989 Pelgrom, Duinmaijer e Welbers publicam um artigo onde são apresentadas as propriedades de casamento dos transistores *MOS* [8]. Ao invés de partir das suas causas, como nos artigos anteriores, esse trabalho parte de um tratamento matemático mais geral para o descasamento entre dois dispositivos que ocupam regiões diferentes de um substrato, onde um certo parâmetro P varia espacialmente. Tal análise leva naturalmente à dedução das dependências espaciais do descasamento. O conceito de dependência espacial do descasamento global é introduzido (efeito da distância entre os dispositivos), ao contrário de [6], onde esse descasamento era modelado por uma constante. São propostos novos parâmetros experimentais relacionados à variância de V_{TO} , K e β (tensão de limiar, fator de corpo e fator de corrente) em função de efeitos locais, ou $A_{V_{TO}}$, A_K e A_β , e em função de efeitos globais, ou $S_{V_{TO}}$, S_K e S_β , respectivamente. Uma vasta gama de dimensões e relações W/L diferentes é caracterizada em dois diferentes processos, permitindo a verificação das dependências do descasamento com a área e com os efeitos de borda dos transistores, e ainda com a distância entre eles. Ao final, o modelo proposto é utilizado na análise teórico-experimental de um espelho de corrente e de um circuito de referência tipo *band-gap*.

O modelo de Pelgrom conseguiu abranger de modo bastante formal e completo todos os pontos-de-vista até então levados em conta na análise de descasamento, incluindo os efeitos geométricos, de polarização e do processo de fabricação, vindo a tornar-se referência para a maioria dos projetistas e pesquisadores, inclusive na indústria. A simplicidade é evidente, visto que modela o descasamento local por apenas três parâmetros experimentais, relacionando-os às dimensões do transistor. Atualmente esse modelo ainda é utilizado como a base da análise de descasamento pela maioria das indústrias e projetistas, de forma que será tratado mais detalhadamente em uma seção posterior.

Muitas contribuições foram dadas a esse assunto desde então, como a adaptação desse modelo aos simuladores em uso [15], a verificação dos efeitos da topografia do circuito [16], o emprego do modelo na análise de produtividade, otimização [17] e

desempenho [13] de circuitos e a inclusão de alguns novos parâmetros de ajuste (*fitting*) [18].

Em 1992, Nishinohara, Shigyo e Wada [19] retomam a discussão iniciada por Keyes [2] sobre os efeitos causados na tensão de limiar pela flutuação microscópica na distribuição de dopantes na região ativa. A análise é feita através da modelagem do potencial de superfície. Em [20], [21] e [22], os autores demonstram experimentalmente que o comportamento da distribuição da tensão de limiar (V_T) é dado por uma função gaussiana e que essa distribuição é principalmente resultante da flutuação na carga líquida da zona de depleção decorrente da flutuação no número de dopantes nessa região. O efeito da flutuação na concentração de dopantes torna-se muito significativo na análise do descasamento local [23], [24], [25], [26], principalmente devido à progressiva redução nas dimensões dos dispositivos, já em escala sub-micrométrica, onde os efeitos de canal curto e estreito apresentam grande influência no comportamento dos mesmos [27], [28]. Os efeitos do posicionamento dos dopantes na região ativa, sobre o descasamento entre transistores *MOS*, são discutidos em [29], [30], [31], [32].

Outros efeitos relacionados ao processo de dopagem são estudados em outras publicações, como o impacto do implante *halo* (*pocket*) [33], [34], [35], e o efeito da formação de aglomerados (segregação de dopantes) no poli-silício (e conseqüente surgimento de uma zona de depleção não uniforme neste eletrodo) [36], [37]. O descasamento em transistores *SOI* (*silicon-on-insulator*) também é discutido [30]. O efeito do perfil vertical de dopantes da região ativa sobre o descasamento [38], [39], [40], e a proposta de usar esse grau de liberdade como forma de reduzi-lo [41] também é apresentada em alguns artigos.

Em 1994, Forti e Wright [42] expandem o problema de descasamento à região de inversão fraca da operação do *MOSFET*, mas apenas apresentando uma análise experimental, visto que os próprios autores consideram a carência de um modelo de comportamento elétrico do transistor que seja adequado a uma análise teórica nessa região. Os resultados experimentais, obtidos de uma grande gama de dimensões de transistores em quatro processos diferente, mostram claramente uma diferença de comportamento do descasamento nessa região, em relação às análises feitas em publicações anteriores, todas em regime de inversão forte. O descasamento relacionado à inversão fraca é também modelado em [43].

A dependência do descasamento com a tensão de substrato (*bulk*) é apresentada e modelada em [44] e [45] para a condição de inversão fraca, mas apenas para transistores saturados. O trabalho mostra que a polarização fracamente direta da junção *substrato-fonte* pode reduzir o efeito de descasamento entre transistores *MOS* (reduz a zona de depleção do canal). Em [46] o efeito da polarização de substrato foi analisado apenas através da alteração provocada na variação da tensão de limiar.

Outro modelo de destaque foi introduzido em 1999 por Drennan e McAndrew [47], [48]. Nesse modelo o descasamento na corrente de *dreno* foi analisado através da sua sensibilidade a cada um dos parâmetros do processo (derivadas parciais) resultando numa matriz de parâmetros de descasamento. Essa matriz, quando multiplicada pelo vetor composto pelo desvio-padrão de cada parâmetro do processo, resulta na variação total da corrente em função dos mesmos. O procedimento resulta em boas aproximações do descasamento, mas só tem utilidade como ferramenta computacional, visto que a complexidade dos cálculos o torna inviável para estimativas manuais do projetista. Além disso, a necessidade de um grande número de parâmetros torna complexa a caracterização do processo e extração destes parâmetros. Esse modelo também está restrito à região de inversão forte, divergindo na medida que a tensão de *overdrive* se aproxima de zero.

Em 2000, Lan e Geiger [49], [50] analisam o desempenho do modelo de Pelgrom para geometrias de transistores *MOS* não-tradicionais, concluindo que este modelo carece de consistência para associações equivalentes tipo série. Tal consistência é de suma importância para a implementação do modelo em simuladores de comportamento elétrico.

Em 2002, Lannacconeand e Amirante [51] consideram o impacto dos efeitos quânticos no descasamento de transistores de canal curto, tema que volta a ser estudado em [52], [53].

Croon *et alli* publicam em 2002 [54] um modelo para descasamento baseado no modelo de Pelgrom, mas que consegue avaliar de forma contínua as regiões de inversão fraca a forte, e ainda modelando fisicamente os efeitos da polarização de substrato. Tal modelo sofre da mesma inconsistência do modelo de Pelgrom e necessita um número maior de parâmetros de ajuste.

Em 2003, Yang *et alli* [55] observam uma incoerência no modelo proposto por Pelgrom (mais exatamente na “regra da área”, reflexo de sua inconsistência) e propõem uma forma de contorná-la, dividindo o transistor *MOS* em duas partes em série, sendo uma

representando a região não-saturada e outra a região saturada, e modelando-as através das equações de Pelgrom. Tal modelo mostra-se mais acurado, embora restrito à região de inversão forte, pois diverge para tensões de *overdrive* próximas de zero.

A partir de 2003, o impacto da rugosidade presente nas bordas (*line edge roughness* – LER) do eletrodo de *porta*, sobre o descasamento, passa a ser um fator preocupante em dispositivos com dimensões deca-nanométricas [56], [57], [58], [59], [60].

Em virtude da progressiva redução geométrica dos dispositivos, fator que intensifica o descasamento entre dispositivos, o entendimento dos limites de casamento impostos pela natureza discreta da matéria tem se tornado fundamental para a indústria de semicondutores, o que pode ser observado pelo grande volume de publicações recentes relacionadas ao assunto [61], [62], [63].

Diversos outros artigos foram estudados, relacionados a formas de medida de descasamento e estruturas ou circuitos de teste, sendo que esses artigos serão citados no capítulo correspondente ao circuito de teste.

2.2 O modelo de Pelgrom

O modelo para descasamento entre transistores *MOS* mais conhecido e divulgado nos dias de hoje, inclusive na indústria, ainda é o publicado em 1989 por Pelgrom, Duinmaijer e Welbers [8], também conhecido como *modelo de Pelgrom*, em função do nome do seu principal autor.

Nesse modelo, os autores partem de um tratamento matemático inicial que cobre todas as causas de descasamento entre dispositivos que são dependentes de área. Feito isso, as constantes de descasamento são derivadas do confronto dessa teoria com os dados experimentais.

Optou-se por reproduzir o texto do artigo de Pelgrom, traduzido e adaptado para o português, mas mantendo as expressões e considerações originais do autor, de modo a levar ao leitor as idéias contidas nesse artigo da forma mais fiel possível e sem novas interpretações.

Inicialmente, é considerado o valor de um parâmetro P qualquer, como sendo composto de uma parte fixa e outra que varia aleatoriamente, e que faz com que $P(x,y)$

tenha diferentes valores em posições (x,y) diferentes. O descasamento do parâmetro $P(x,y)$ entre dois dispositivos dependentes de área pode então ser expressa como

$$\Delta P(x_{12}, y_{12}) = \frac{1}{\text{área}} \left[\iint_{\text{área}(x_1, y_1)} P(x', y') dx' dy' - \iint_{\text{área}(x_2, y_2)} P(x', y') dx' dy' \right], \quad (2.2.1)$$

onde x e y são as coordenadas na superfície considerada.

Aplicando-se a transformada de Fourier bidimensional à integral, sua parte geométrica pode ser separada da *fonte* de descasamento

$$\tilde{\Delta P}(\omega_x, \omega_y) = \tilde{G}(\omega_x, \omega_y) \tilde{P}(\omega_x, \omega_y), \quad (2.2.2)$$

onde o til representa as variáveis transformadas para o domínio frequência (ω) .

Nesse domínio, o processo gerador do descasamento em P e a geometria do dispositivo G podem ser interpretados como uma *fonte* espacial de frequências \tilde{P} e uma função filtro \tilde{G} , respectivamente, de forma que possam ser analisados separadamente.

A função filtro de um par de geometrias retangulares e idênticas, com área WL , afastadas entre si no eixo x de uma distância D_x (o afastamento foi considerado apenas no eixo x por simplicidade), é obtida pela análise de Fourier resultando

$$\tilde{G}(\omega_x, \omega_y) = \frac{\sin(\omega_x L / 2)}{\omega_x L / 2} \frac{\sin(\omega_y W / 2)}{\omega_y W / 2} [2 \sin(\omega_x D_x / 2)]. \quad (2.2.3)$$

Nesse ponto, são consideradas duas classes distintas de causas para o descasamento, que abrangem causas físicas de comportamento semelhante. A primeira classe compreende os processos estocásticos, causadores de descasamento em P com variações de curta distância, cuja distância de correlação seja muito menor que as dimensões físicas do dispositivo (W e L , respectivamente largura e comprimento do dispositivo). No domínio de Fourier, esses processos representam um “ruído branco” de alta frequência espacial e cuja média é zero. Encontram-se nessa classe as flutuações locais de parâmetros físicos, como a distribuição de dopantes no substrato, flutuações locais de mobilidade, granularidade do óxido (n.t.: supõe-se que o autor quis dizer rugosidade do óxido), flutuações locais na permissividade do óxido, etc. Esses processos são considerados pelo autor como tendo efeitos muito pequenos e não tendo correlação entre si, de modo que suas contribuições possam ser simplesmente somadas.

Na segunda classe encontram-se as variações de grande distância de correlação, bem acima das dimensões dos dispositivos, sendo representada no domínio de Fourier por uma frequência espacial muito baixa. Nessa classe são encontrados os parâmetros físicos com variações globais, como gradientes térmicos, aberrações nas lentes e máscaras de fotolitografia, distorções mecânicas na lâmina ou nos equipamentos, variações de concentração de substâncias de ataque, deposição ou implante, etc.

A variância no descasamento do parâmetro P , pode então ser calculada através da análise da densidade de potência no domínio de Fourier de $\tilde{G} \cdot \tilde{P}$, ou seja

$$\sigma^2(\Delta P) = \frac{1}{4\pi^2} \int_{\omega_y \rightarrow -\infty}^{\omega_y \rightarrow \infty} \int_{\omega_x \rightarrow -\infty}^{\omega_x \rightarrow \infty} \left| \tilde{G}(\omega_x, \omega_y) \right|^2 \cdot \left| \tilde{P}(\omega_x, \omega_y) \right|^2 d\omega_x d\omega_y. \quad (2.2.4)$$

Considerando-se a expressão (2.2.3) e as definições apresentadas das classes de variações de curta e de longa distância, em termos de frequências espaciais, pode-se calcular a variância do parâmetro P , através de (2.2.4), resultando

$$\sigma^2(\Delta P) = \frac{A_P^2}{WL} + S_P^2 D_x^2, \quad (2.2.5)$$

onde A_P é o fator de proporcionalidade do parâmetro P com a área (WL) dos dispositivos e S_P representa a variação de P com a distância (D_x) entre os mesmos.

Com o uso de um modelo de comportamento elétrico do transistor *MOS*, o resultado obtido em (2.2.5) pode ser aplicado ao descasamento entre transistores. O modelo empregado pelos autores foi o de primeira ordem, restrito aos transistores de canal longo e na região linear:

$$I_D = \beta \left[\frac{(V_{GS} - V_T - V_{DS}/2)V_{DS}}{1 + \theta(V_{GS} - V_T)} \right], \quad (2.2.6)$$

onde $\beta = \mu C'_{ox} W/L$ é o fator de corrente, sendo μ a mobilidade de *portadores* no canal e C'_{ox} a capacitância de óxido de *porta* por unidade de área, e θ representa o fator de redução de mobilidade para campos transversais (nesse parâmetro foram incluídos também os efeitos das resistências série). I_D representa a corrente de *dreno*, V_{GS} a tensão entre *porta* e *fonte*, V_{DS} a tensão entre *dreno* e *fonte* e V_T a tensão de limiar, que pode ser decomposta na soma de uma parcela constante V_{T0} com outra cujo valor depende da relação entre o fator de corpo K e a tensão entre *fonte* e substrato V_{SB} , ou $V_T = V_{T0} + K(\sqrt{|V_{SB}| + 2\phi_F} - \sqrt{2\phi_F})$, onde

ϕ_F representa o nível de Fermi dos portadores majoritários. O fator de corpo K , por sua vez, é um parâmetro do processo dado por $K = \sqrt{2q\epsilon_{Si}N_A} / C'_{ox}$, onde q representa a carga do elétron, ϵ_{Si} é a permissividade do silício, N_A é a concentração de dopantes no substrato, e C'_{ox} é a capacitância do óxido por unidade de área.

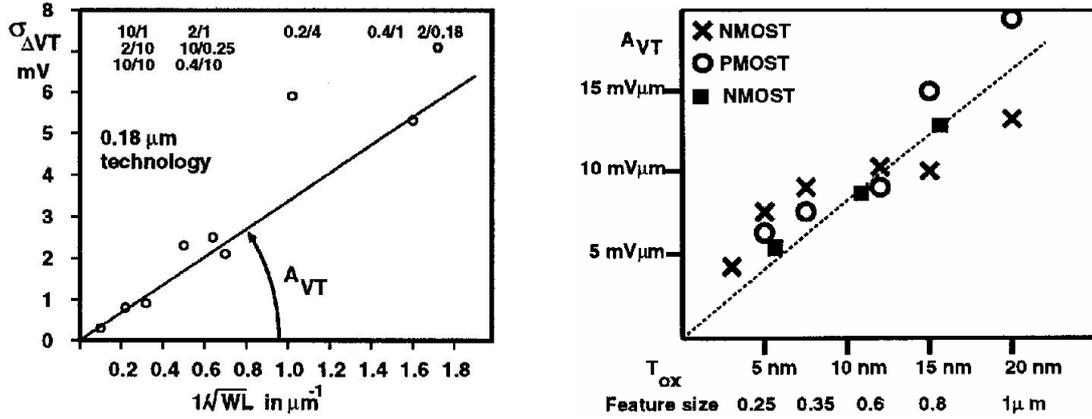


Figura 2.1: Medida do desvio-padrão da tensão de limiar ($\sigma_{\Delta V_T}$), em função da área do transistor para transistores *NMOS*, e o parâmetro A_{V_T} de diversas gerações de processos (extraído de [10]).

Os parâmetros escolhidos pelos autores para representar o descasamento são V_{T0} , K e β . As causas de variações para os dois primeiros se enquadram nos quesitos da análise matemática anterior (sem correlação entre si, sendo mutuamente independentes), de forma que seus desvios-padrão podem ser expressos diretamente por (2.2.5). O parâmetro β depende em parte de outros fatores (W , L , μ e C'_{ox}) do modelo, tendo sido expresso por (2.2.5) como uma aproximação. Assim, tem-se

$$\sigma^2(V_{T0}) = \frac{A_{V_{T0}}^2}{WL} + S_{V_{T0}}^2 D^2, \quad (2.2.7a)$$

$$\sigma^2(K) = \frac{A_K^2}{WL} + S_K^2 D^2, \quad (2.2.7b)$$

$$\sigma^2(\beta) = \frac{A_\beta^2}{WL} + S_\beta^2 D^2, \quad (2.2.7c)$$

onde $A_{V_{T0}}$, A_K e A_β são os fatores de proporcionalidade de V_{T0} , K e β com a área dos dispositivos e $S_{V_{T0}}$, S_K e S_β representam as variações de V_{T0} , K e β com a distância D entre eles. A figura 2.1 apresenta os valores do desvio-padrão da tensão de limiar ($\sigma_{\Delta V_T}$), extraídos de transistores *NMOS* com diversas dimensões em uma tecnologia com $0,18\mu\text{m}$

de comprimento mínimo de canal (esquerda), e os valores do parâmetro A_{VT} para diversas gerações de processos (direita).

Desconsiderando-se os efeitos da polarização de substrato no descasamento, o desvio-padrão normalizado da corrente de *dreno* pode ser expresso, a partir do modelo utilizado, por

$$\frac{\sigma^2(I_D)}{I_D^2} = \frac{4\sigma^2(V_{T0})}{(V_{GS} - V_{T0})^2} + \frac{\sigma^2(\beta)}{\beta^2} \quad (2.2.8)$$

sendo esta a expressão mais usual do modelo de Pelgrom, onde os valores dos desvios-padrão de V_{T0} e β são dados pelas expressões (2.2.7a) e (2.2.7c). Normalmente, o fabricante da tecnologia em que se tem interesse se limita (quando muito) a informar os parâmetros de descasamento local A_{VT0} e A_β . Os efeitos do descasamento global devem ser atenuados mediante as técnicas de leiaute já mencionadas. O nível de inversão do transistor entra no modelo através do valor da tensão V_{GS} .

Outra forma para a expressão (2.2.8) e que é mais adequada à representação do estado de polarização do transistor através de seu nível de inversão é

$$\frac{\sigma^2(I_D)}{I_D^2} = \left(\frac{g_m}{I_D}\right)^2 \sigma^2(V_{T0}) + \frac{\sigma^2(\beta)}{\beta^2}, \quad (2.2.9)$$

onde g_m representa a transcondutância vista pelo terminal de *porta* e é facilmente representável em função dos níveis de inversão direta e reversa do dispositivo.

2.3 Considerações sobre modelagem

A modelagem do descasamento entre transistores *MOS* tem sido feita, por vários autores, de duas formas: através da análise do comportamento físico do transistor, necessitando portanto de um modelo físico para o comportamento elétrico deste dispositivo, ou então expressando-se a sensibilidade dos parâmetros de um modelo de ajuste experimental do dispositivo às variações nos diversos fatores físicos a que estão relacionados. A primeira opção leva a modelos de descasamento com sentido físico e equacionamento mais compacto, apresentando menor número de parâmetros [8], [18], [26], [28], [51], [55]. A segunda geralmente leva à representação do descasamento através de conjuntos (matrizes) de fatores de sensibilidade, e em maior número pois muitos são inter-

relacionados, resultando em um modelo com um maior número de equações, sendo a maioria sem sentido físico explícito. Os modelos obtidos desta forma geralmente têm aplicação apenas computacional [47], [48], devido à sua complexidade.

Considerando-se o primeiro caso (modelos físicos), o estudo dos diversos modelos e tentativas de modelagem, publicados ao longo dos anos, revela a presença de dois grandes pontos-fracos nesse processo, a saber, a indisponibilidade de um bom modelo de comportamento elétrico do transistor *MOS* e a confusão entre descasamento de parâmetros e descasamento de comportamento.

Para que o modelo de descasamento seja acurado, abrangente e consistente, é necessário que seja baseado em um modelo de comportamento elétrico do transistor *MOS* que também seja acurado, abrangente e consistente. Não é suficiente que a análise das causas do descasamento seja feita de forma extremamente criteriosa, tentando prever todos os possíveis fatores locais e globais de descasamento, se o modelo elétrico do transistor utilizado no desenvolvimento for pobre e limitado a uma ou a poucas regiões de operação. Em geral, os modelos de descasamento apresentados nas publicações pesquisadas utilizam-se de modelos elétricos de primeira ordem, ou limitados apenas a uma condição de operação do transistor. Ou ainda utilizam equações distintas para representar condições de operação distintas, tornando o modelo de descasamento de uso mais difícil e confuso. Assim, deduz-se que é necessário o uso de um modelo de comportamento elétrico do transistor *MOS* que tenha origem física e consiga representar sua operação nas mais variadas condições (região linear e saturação, e inversões forte, moderada e fraca), através de uma única e compacta equação e com o uso de poucos parâmetros.

Por outro lado, visto que o comportamento elétrico do transistor *MOS* não é linear, deve-se ter o cuidado de concatenar os efeitos do descasamento na ordem correta, sob pena de se chegar a um modelo inconsistente. O *Anexo A – Consistência de um modelo de descasamento* apresenta um estudo sobre a consistência de modelos de descasamento para associações paralelo e série de transistores.

3 O MODELO DE DESCASAMENTO PROPOSTO

O presente capítulo expõe em detalhe o modelo para descasamento local entre transistores *MOS* proposto nesta tese, além de outros tópicos periféricos. Inicialmente, são apresentados alguns conceitos fundamentais do modelo *ACM*, que foi utilizado como espinha dorsal deste trabalho. Alguns fundamentos relacionados à teoria estatística das distribuições aleatórias são também apresentados, e o problema, cada vez mais grave, do reduzido número de átomos dopantes na região ativa dos transistores *MOS* é exposto. O efeito das flutuações locais na carga de inversão ao longo do canal, e suas implicações sobre a corrente é analisado. O modelo proposto para o descasamento local é detalhadamente apresentado, no mesmo formato em que foi publicado em [64]. Finalmente, outros efeitos físicos que afetam o descasamento são expostos e analisados.

3.1 Considerações gerais

Sob o ponto-de-vista externo, o transistor *MOS* se comporta como um dispositivo de quatro terminais, cuja corrente que circula entre estes depende (de forma não-linear e interdependente) da tensão simultaneamente aplicada aos mesmos. Idealmente, quando dois transistores idênticos (com a mesma geometria e fabricados no mesmo processo) são submetidos ao mesmo conjunto de tensões em seus terminais, as correntes estabelecidas nestes adquirem exatamente os mesmos valores.

Entretanto, no mundo real, estes mesmos dois transistores apresentam pequenas diferenças nas correntes resultantes, fruto do grau de incontabilidade (variações e flutuações) nos parâmetros físicos dos materiais e processos utilizados na sua fabricação, e a que chamamos de descasamento (*mismatch*). Ao projetista de circuitos integrados, que é o usuário final desse dispositivo, interessa a possibilidade de avaliar o descasamento entre as grandezas elétricas externas (correntes), em função dos seus graus de liberdade de projeto (geometria e polarização), de modo que possa agir no sentido de reduzi-lo a valores toleráveis pelas especificações do circuito que está desenvolvendo. Assim, vê-se que um

modelo que não expresse o descasamento entre transistores *MOS* na forma de um desvio na corrente de *dreno*, em função das dimensões (W e L) e da condição de operação (polarização), dificilmente terá aplicação prática.

Em um transistor *MOS* ideal, a interação dos campos elétricos, resultantes dos potenciais aplicados aos seus quatro terminais, com as cargas elétricas que compõem os materiais que o formam, estabelece uma corrente elétrica que se distribui homoganeamente ao longo da região ativa (canal) e que flui tipicamente entre *dreno* e *fonte* (desconsiderando aqui as correntes tuneladas através do óxido da *porta* e as correntes de fuga através do substrato).

Em transistores *MOS* reais, as flutuações microscópicas naturalmente presentes nos parâmetros físicos dos materiais que os compõem provocam flutuações locais na densidade da corrente elétrica ao longo da região ativa, variando-a ao redor de um valor médio. Como estas flutuações de corrente ocorrem em número finito e de forma aleatória, sua integração por sobre a região ativa resulta num valor de desvio líquido da corrente de *dreno* em relação ao valor médio (descasamento).

Assim, a maneira correta de se modelar o descasamento é através da integração dos seus efeitos na corrente de *dreno*, em função das flutuações nos parâmetros que os provocam, utilizando-se um modelo de comportamento elétrico que seja suficientemente abrangente e acurado.

A maioria dos modelos publicados (inclusive o de Pelgrom) faz o procedimento inverso, integrando primeiro as flutuações nos parâmetros físicos do processo em parâmetros de modelo (p. ex. A_{VT0} , A_K e A_β) para depois refletirem essas flutuações como uma variação da corrente de *dreno* através do modelo elétrico do transistor.

Ora, como o comportamento elétrico do transistor *MOS* (e do seu modelo) não é linear, o resultado deste segundo procedimento diverge do primeiro, fazendo com que o modelo de descasamento resultante apresente estimativas incorretas e seja inconsistente para associações do tipo série de transistores.

A análise dos efeitos globais sobre o descasamento não foi incluída neste trabalho, pois os resultados experimentais apresentados na vasta literatura já publicada sobre este assunto indicam que estes efeitos têm um impacto muito inferior ao dos efeitos locais, para circuitos projetados utilizando-se as técnicas de leiaute descritas no item 1.5.

3.2 O modelo ACM

O modelo *ACM* (*advanced compact model*) [65], [66] é um modelo consistente para o comportamento elétrico do transistor *MOS*, que parte da aproximação de dependência linear entre a densidade de carga de inversão Q'_I e o potencial de superfície ϕ_S , a qual é válida para as condições de inversão fraca, moderada e forte

$$dQ'_I = (C'_b + C'_{ox})d\phi_S = nC'_{ox}d\phi_S. \quad (3.2.1)$$

Em (3.2.1) n é o fator de rampa, levemente dependente do potencial de *porta*, e C'_b , C'_{ox} são as capacitâncias de depleção e óxido por unidade de área.

A corrente de *dreno*, considerando-se o transistor de canal longo, é calculada através de (3.2.1) e da aproximação de folha de carga [67]. Assim, esta corrente é representada como o somatório das componentes de deriva (resultante da ação de um campo elétrico sobre os portadores) e de difusão (resultante da ação da agitação térmica sobre um gradiente de concentração de portadores), sendo dada por

$$I_D = \frac{\mu W}{nC'_{ox}} \left(-Q'_I + nC'_{ox}\phi_t \right) \frac{dQ'_I}{dx}, \quad (3.2.2)$$

onde μ representa a mobilidade efetiva de portadores, W é a largura do canal, ϕ_t é o potencial térmico e x representa a posição ao longo do canal, em relação à *fonte*.

A outra especificidade do modelo *ACM* é o uso do modelo unificado de controle de carga (*unified charge control model* ou *UCCM*) [68], que relaciona a densidade de carga dos portadores com o potencial no canal, ou

$$V_P - V_x = \phi_t \left[\frac{Q'_I}{Q'_{IP}} - 1 + \ln \left(\frac{Q'_I}{Q'_{IP}} \right) \right], \quad (3.2.3)$$

onde $Q'_{IP} = -nC'_{ox}\phi_t$ é a densidade de carga de inversão na situação de estrangulamento (*pinch-off* - definida segundo o modelo *ACM* como a condição onde as duas componentes da corrente, difusão e deriva, se igualam, ou seja, é o ponto de transição da condição de inversão fraca para a forte). O potencial $V_P = (V_{GB} - V_T)/n$ é a tensão de estrangulamento (*pinch-off voltage*), e V_x é o potencial do canal na posição x .

O uso de (3.2.2) em conjunto com (3.2.3) resulta em

$$I_D = -\mu W Q'_I(x) \frac{dV_x}{dx}. \quad (3.2.4)$$

Conseqüentemente, o modelo *ACM* é integralmente consistente com a formulação de potencial de quasi-Fermi.

Integrando-se (3.2.2) da *fonte* ao *dreno*, resulta a expressão da corrente que circula entre esses terminais

$$I_D = I_F - I_R, \quad (3.2.5)$$

sendo as correntes direta e reversa (I_F e I_R) dadas por

$$I_{F(R)} = \mu n C'_{ox} \frac{W}{L} \frac{\phi_t^2}{2} \left[\left(\frac{Q'_{IS(D)}}{n C'_{ox} \phi_t} \right)^2 - 2 \frac{Q'_{IS(D)}}{n C'_{ox} \phi_t} \right] = I_{SQ} \frac{W}{L} i_{f(r)}, \quad (3.2.6)$$

onde L é o comprimento do canal, $Q'_{IS(D)}$ é a densidade de carga de inversão na *fonte* (*dreno*), I_{SQ} é a corrente específica por quadrado e $i_{f(r)}$ é o nível de inversão direto (reverso).

As seguintes relações definem a corrente específica de folha e o nível de inversão, em termos da densidade de carga de inversão

$$I_{SQ} = \frac{1}{2} \mu n C'_{ox} \phi_t^2, \quad (3.2.7)$$

$$-\frac{Q'_{IS(D)}}{n C'_{ox} \phi_t} = \sqrt{1 + i_{f(r)}} - 1. \quad (3.2.8)$$

3.3 Distribuição de Poisson

A teoria matemática da probabilidade [69] pode ser aplicada a algumas classes específicas de sistemas físicos, fornecendo uma previsão numérica do comportamento aproximado de uma grandeza que varia aleatoriamente entre situações semelhantes. Isso é feito através de dados obtidos pela análise de um conjunto limitado de situações representativas do todo (*amostra*). A maneira como cada valor da grandeza se repete (*freqüência*) nesse conjunto é denominado de *distribuição*.

Uma distribuição é chamada de *binária* [69], quando o resultado da grandeza em questão se enquadra em apenas duas possibilidades mutuamente excludentes, gerando

resultados do tipo *sim* e *não* (*sucesso* e *insucesso*). A face de uma moeda (*cara* ou *coroa*), a ocorrência de um *seis* em um dado (*seis* ou *não-seis*), ou a existência de um átomo estranho em uma posição de uma rede cristalina (*existe* ou *não-existe*), são exemplos de distribuições binárias. Em uma distribuição deste tipo, na medida em que o número de situações semelhantes consideradas (tamanho da amostra) aumenta, a distribuição dos valores resultantes se aproxima de uma função matemática contínua, conhecida como função *normal* ou *gaussiana* [69], e que pode ser determinada através de apenas duas quantidades: *média* (μ) e *desvio-padrão* (σ). A média, representa o ponto “central” no entorno do qual a distribuição se situa, e o desvio-padrão representa a maneira como essa distribuição está concentrada ou expandida simetricamente ao redor da média, sendo que ambas podem ser extraídas experimentalmente de uma amostra. O grau de confiabilidade da média e do desvio-padrão, como representativos de um comportamento geral, depende do tamanho da amostra (quantidade de situações representativas do sistema, analisadas experimentalmente).

Um caso interessante de distribuição binária é a *distribuição de Poisson* [69], que se caracteriza por apresentar um número muito pequeno de sucessos, frente ao número de insucessos. Neste tipo de distribuição, na medida em que aumentamos o tamanho da amostra, a média e o desvio-padrão resultantes passam a se correlacionar ($\sigma \cong \sqrt{\mu}$), de modo que o comportamento da distribuição pode ser aproximado através da determinação experimental uma única quantidade, ou seja, da média.

O processo de dopagem de um substrato semiconductor ocorre inserindo-se átomos de um material estranho (*dopante*), de forma que estes ocupem posições de átomos originais do semiconductor na rede cristalina. Este processo se caracteriza por uma distribuição aleatória dos átomos dopantes no volume de substrato considerado, de forma que a média de concentração volumétrica de dopantes seja a mesma em todas as regiões do volume considerado.

A rede cristalina do silício, substrato geralmente utilizado em dispositivos eletrônicos, apresenta uma densidade volumétrica de aproximadamente 5×10^{22} átomos/cm³ [67]. A concentração de dopantes que é utilizada no substrato, na região do canal dos transistores *MOS* nas tecnologias atuais, se situa geralmente entre 10^{15} e 10^{18} átomos/cm³ [67]. Assim, tem-se uma relação maior que 1:10.000 entre o número médio de sucessos (existir um átomo dopante numa posição possível da rede) e de insucesso (não existir), de

modo que esta situação enquadra-se perfeitamente no modelo de uma distribuição de Poisson. Como resultado, a curva de distribuição estatística de átomos dopantes em um substrato semiconductor pode ser quantificada apenas através do conhecimento da sua concentração média.

3.4 Flutuação na concentração de dopantes

Com os avanços tecnológicos levando a uma contínua redução nas dimensões dos transistores *MOS*, o número de átomos dopantes na região ativa (canal) deste tem se reduzido no mesmo ritmo. Em uma tecnologia com comprimento efetivo de canal de 250 nm, um transistor de dimensões mínimas apresenta algo como 1100 átomos dopantes na região de depleção, sob o canal, número esse que cai para apenas cerca de 200 em tecnologias de 100 nm (há muito em uso na indústria de semicondutores) [70].

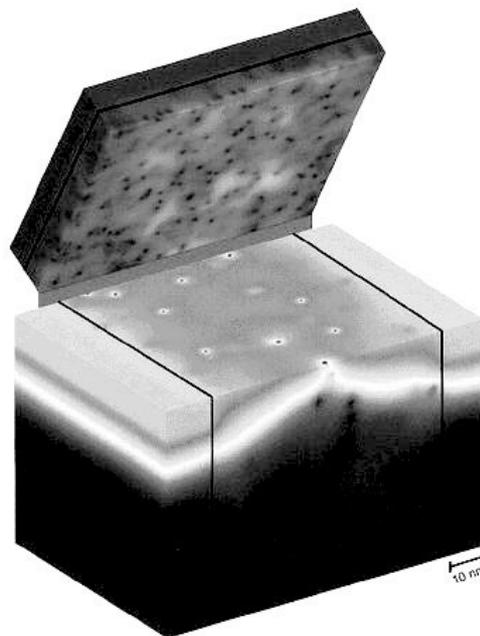


Figura 3.1: Simulação da distribuição de dopantes e potenciais no substrato e no poli-silício de um MOSFET com 50nm de comprimento de canal.

A figura 3.1 foi extraída de [36] e apresenta uma simulação da distribuição atomística de dopantes no substrato e no poli-silício, em um transistor *MOS* com 50 nm de comprimento de canal. O óxido foi retirado e o poli-silício foi aberto como “um livro” de modo que se consiga vislumbrar essas cargas e o comportamento da distribuição de potenciais. Os pontos escuros representam a localização das impurezas implantadas no substrato e no poli-silício.

Observando-se essa figura, vê-se claramente que em tecnologias sub-micrométricas, a flutuação aleatória dos dopantes na região ativa passa a exercer grande impacto na determinação das características elétricas do transistor, provocando variações locais na corrente do canal, e se tornando a principal causa do descasamento entre transistores *MOS*, ponto de partida no desenvolvimento do modelo aqui proposto.

A flutuação dos dopantes na região ativa tem sido analisada como um processo estocástico, principalmente sob o aspecto da flutuação do *número de impurezas* e seus efeitos no comportamento do transistor, geralmente centrado o estudo nas implicações dessa flutuação sobre a variação da tensão de limiar (V_T) [11], [19], [20], [21]. Outro aspecto, menos analisado, está relacionado à flutuação da *posição das impurezas* dentro do volume da região ativa do dispositivo [29], [30], [31]. De fato, a flutuação de dopantes se dá como uma combinação dos dois efeitos, sendo mais adequadamente representada como o somatório de variações microscópicas de concentração, localizadas tridimensionalmente dentro do volume da região ativa [29], [71].

No modelo aqui proposto, integra-se o efeito da flutuação do número de impurezas sobre a corrente de *dreno* de um *MOSFET*, mas associado ao seu efeito posicional ao longo do canal (em uma dimensão apenas; eixo x). O efeito das flutuações de concentração de dopantes no sentido da profundidade (eixo y) do volume da região de depleção é analisado separadamente, incluindo as deformações sofridas por essa região com os potenciais aplicados aos terminais do transistor.

3.5 Variações de condutância no canal e seus efeitos

A corrente de *dreno* que flui pelo canal de um transistor *MOS* pode ser estabelecida através da relação entre a diferença de potencial existente entre os extremos de um segmento do mesmo e sua condutância média. Considerando-se a região do canal apresentada na figura 3.2, pode-se afirmar que a corrente de *dreno* I_D , resultante num segmento transversal Δx qualquer, é dada por

$$I_D = G(x)\Delta V_C = \frac{W}{\Delta x} \sigma(x)\Delta V_C, \quad (3.5.1)$$

onde $G(x)$ representa a condutância média do segmento na posição x , $\sigma(x)$ representa a condutividade superficial da região delimitada por Δx e W , na coordenada x , e ΔV_C é a queda de tensão no segmento.

Lembrando que, para um determinado potencial de *porta* V_G constante, considerando-se a mobilidade de portadores μ como uma grandeza independente, a condutividade superficial na coordenada x depende apenas do potencial do canal naquele ponto, $V_C(x)$, ou seja, $\sigma(x) = \sigma(V_C(x))$. Assim, pode-se integrar (3.5.1) entre *dreno* e *fonte*, resultando em

$$I_D = \frac{W}{L} \int_{V_S}^{V_D} \sigma(V_C) dV_C, \quad (3.5.2)$$

onde V_D e V_S representam os potenciais aplicados aos terminais de *dreno* e *fonte*, respectivamente.

Supondo agora que em uma região delimitada pelas coordenadas x_0 e $x_0 + \Delta x_0$, a condutividade superficial sofra um acréscimo de $\delta\sigma$, lembrando que os potenciais nos terminais são fixos externamente, tem-se um acréscimo resultante na corrente de δI_D , ou

$$I_D + \delta I_D = \frac{W}{L} \left[\int_{V_S}^{V_C(x_0)} \sigma(V_C) dV_C + \int_{V_C(x_0)}^{V_C(x_0 + \Delta x_0)} [\sigma(V_C) + \delta\sigma] dV_C + \int_{V_C(x_0 + \Delta x_0)}^{V_D} \sigma(V_C) dV_C \right], \quad (3.5.3)$$

que com o auxílio de (3.5.2), resulta

$$\delta I_D = \frac{W}{L} \delta\sigma \int_{V_C(x_0)}^{V_C(x_0 + \Delta x_0)} dV_C = \frac{W}{L} \delta\sigma \Delta V_{C0}. \quad (3.5.4)$$

A eq. (3.5.1) pode então ser reescrita para a região delimitada na condição atual, ou

$$I_D + \delta I_D \cong \frac{W}{\Delta x_0} (\sigma(x_0) + \delta\sigma) \Delta V_{C0}, \quad (3.5.5)$$

onde, isolando-se ΔV_{C0} e aplicando-se a (3.5.4), resulta

$$\delta I_D = \frac{\Delta x_0}{L} \frac{\delta\sigma}{\sigma(x_0) + \delta\sigma} (I_D + \delta I_D) = \frac{\Delta x_0 \delta\sigma}{L(\sigma(x_0) + \delta\sigma) - \Delta x_0 \delta\sigma} I_D. \quad (3.5.6)$$

Considerando-se que $L\sigma(x_0) \gg \Delta x_0 \delta\sigma$, pode-se simplificar (3.5.6), tendo-se

$$\delta I_D \cong \frac{\Delta x_0}{L} \frac{\delta\sigma}{\sigma(x_0) + \delta\sigma} I_D \cong \frac{\Delta x_0}{L} \frac{\delta\sigma}{\sigma(x_0)} I_D. \quad (3.5.7)$$

A última simplificação de (3.5.7) é válida para os casos em que $\sigma(x_0) \gg \delta\sigma$, situação que geralmente ocorre, considerando-se que a flutuação local na carga de inversão do canal

é normalmente muito menor que seu valor médio. Este resultado é similar ao apresentado em [72], [73].

Lembrando que a relação entre a condutividade superficial do canal $\sigma(x)$ e a densidade de carga de inversão $Q'_I(x)$ é dada por

$$\sigma(x) = \mu Q'_I(x), \quad (3.5.8)$$

pode-se reescrever a expressão (3.5.7) como

$$\frac{\delta I_D}{I_D} \cong \frac{\Delta x}{L} \frac{\delta Q'_I}{Q'_I(x)}. \quad (3.5.9)$$

Desta expressão observa-se que uma pequena flutuação na densidade de carga de inversão ($\delta Q'_I$), que ocorra em um pequeno segmento do canal, provoca uma variação proporcional na corrente de *dreno* (δI_D). Considerando-se as flutuações ao longo do canal como aleatórias, e lembrando que de fato ocorrem em número finito e afetando segmentos não infinitesimais do mesmo, sua integração média quadrática entre *dreno* e *fonte* resulta num desvio total da corrente do transistor. Assim, pode-se afirmar que dois transistores geometricamente idênticos, quando submetidos a potenciais de polarização idênticos, poderão apresentar correntes de *dreno* ligeiramente diferentes, em decorrência das possíveis flutuações aleatórias na densidade de carga de inversão, que ocorram ao longo do canal.

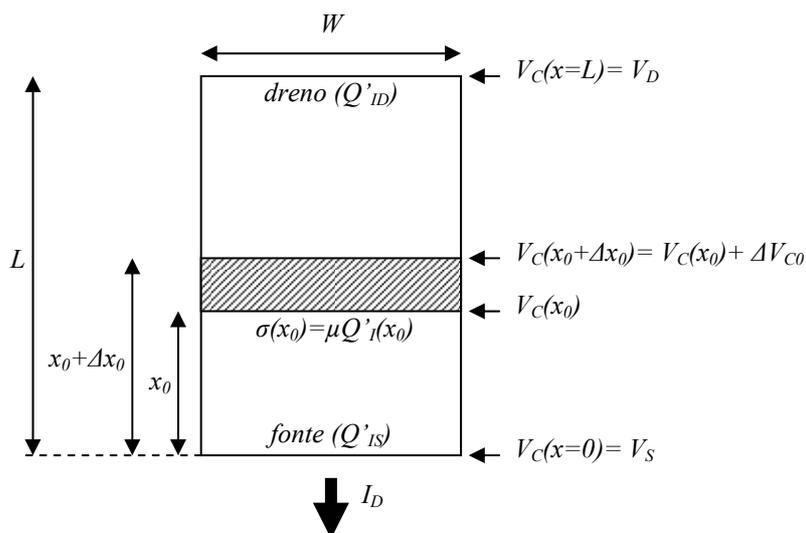


Figura 3.2: Representação da região ativa de um transistor MOS.

3.6 Um modelo compacto para o descasamento entre MOSFETs

O modelo aqui proposto parte da consideração de que o descasamento na corrente de *dreno* entre transistores *MOS* é decorrente do somatório dos efeitos das flutuações que ocorrem na corrente em cada segmento da área do canal.

A figura 3.2 apresenta a vista superior do canal do transistor *MOS*, com dimensões W e L , onde se destaca uma pequena seção de comprimento Δx_0 na posição x_0 (em relação à *fonte*).

Supondo que o transistor está polarizado, uma corrente de dreno (I_D) flui entre *dreno* e *fonte*. Imaginando-se que naquela pequena região do canal, a concentração média de dopantes é ligeiramente diferente do valor médio nas outras regiões do transistor, tem-se como consequência um desvio na corrente. A integração do efeito desse desvio em I_D devido a flutuações semelhantes em todas as pequenas regiões ao longo do canal, da *fonte* ao *dreno*, resulta num fator médio de incerteza da corrente, aqui denominado *descasamento*.

Para que essa integração seja feita, modela-se o dispositivo apresentado como sendo resultado da composição série de três outros dispositivos: um transistor com comprimento de canal $L-x$ (próximo ao *dreno*), um segundo transistor com comprimento de canal x (próximo à *fonte*) e um resistor caracterizado pela condutância do segmento de área $W\Delta x_0$.

Considerando-se que uma flutuação na concentração de dopantes, que estão dentro da região de depleção nesse segmento, provoca uma flutuação correspondente na densidade de carga de inversão $Q'_I(x)$, e que a corrente que circula entre *dreno* e *fonte* é função da integração de $Q'_I(x)$ ao longo do canal, o resultado final é um desvio médio na corrente que flui pelo dispositivo.

Através do modelo *ACM*, pode-se calcular as respectivas transcondutâncias dos dois transistores do modelo (de comprimentos $L-x$ e x), percebidas por essa flutuação de corrente e, considerando-se a condutância do segmento resistivo, calcular o efeito dessa flutuação na corrente total I_D .

Dessa forma, ao se integrar todas as contribuições das flutuações ao longo do canal, e supondo que estas não são correlacionadas entre si, tem-se a dispersão média total na

corrente I_D resultante (desvio-padrão de I_D , ou σ_{ID}), que vem a ser o descasamento médio esperado da corrente.

Um tratamento completo para o modelo aqui descrito foi publicado no *IEEE Journal of Solid-State Circuits* [64], e é apresentado de forma adaptada a seguir.

3.6.1 Introdução

É de amplo reconhecimento que o desempenho de grande parte dos circuitos analógicos, e até dos digitais, está limitado pelo descasamento de seus transistores *MOS* [6], [8], [42], [44]. Nos circuitos analógicos, os desvios das características *CC* dos dispositivos, devido à natureza estocástica da dopagem na região do canal, produz um funcionamento impreciso ou até anormal [70]. Nos circuitos digitais, quando em tecnologias sub-micrométricas, estes desvios causam variações nos atrasos que podem chegar ao equivalente ao atraso de várias portas lógicas [70]. Como exposto pelo professor J. Meindl, no artigo de R. Wilson [74], “*variations will set the ultimate limits on scaling of MOSFETs*”.

A progressiva redução de escala e diminuição na tensão de alimentação dos circuitos integrados tem tornado as limitações impostas pelo descasamento ainda mais importantes, de modo que este tem sido um grande objeto de estudo em anos recentes [48], [49], [50], [75]. Os modelos de descasamento existentes atualmente ou são muito simples, ficando limitados a uma condição de operação específica [6], [7], [8], [44], [49], [50], ou fazem uso de expressões muito complexas [48] como as do modelo BSIM (*Department of Electrical Engineering and Computer Sciences – EECS – at the University of California, Berkeley*).

Em geral, o uso de modelos *CC* na análise do descasamento não é contestado, e é amplamente aceito que o descasamento pode ser modelado através das variações aleatórias que ocorrem nos parâmetros geométricos, do processo e/ou do dispositivo, e o efeito dessas variações aleatórias sobre a corrente de *dreno* pode ser quantizado usando tal modelo. Mas o uso que tem sido feito desses modelos apresenta uma falha fundamental, como apontado por [49] e [50], e mais recentemente [55], a qual implica em resultados inconsistentes. De fato, esses modelos implicitamente assumem que os valores de alguns parâmetros globais do modelo *CC*, podem ser obtidos pela sua integração posicional sobre a área da região do canal, como por exemplo, para a tensão de limiar V_T

$$V_T = \frac{1}{WL} \iint_{\text{area-do-canal}} V_T(x, y) dx dy, \quad (3.6.1)$$

onde W e L são a largura e o comprimento do transistor.

Como demonstrado em [49], [50] e [55], o uso de (3.6.1) na associação série ou paralela de transistores acarreta modelos de descasamento inconsistentes, devido à natureza não-linear do *MOSFET*. Conseqüentemente, a simplória consideração das flutuações aleatórias de um parâmetro global do modelo CC não é apropriada ao desenvolvimento de modelos de descasamento, sendo que novas expressões para este devem ser derivadas do comportamento físico do dispositivo.

O impacto das flutuações locais das impurezas na tensão de limiar dos *MOSFETs*, inicialmente reconhecida em 1975 [2], é uma das principais fontes de descasamento nos atuais dispositivos *MOS* [2], [21], [25], [37], [55]. Na medida que os *MOSFETs* são reduzidos em escala, até as tecnologias *deep-submicron*¹⁰, o número de dopantes na região de depleção reduz, sendo da ordem de apenas algumas centenas para dispositivos de tamanho mínimo [76]. Por exemplo, um transistor mínimo em um processo 0,25 μm , apresenta cerca de 1100 átomos dopantes no volume da região de depleção, enquanto que em um processo 0,1 μm este número cai para cerca de 200 somente [70]. A flutuação no número de átomos dopantes na região de depleção causa uma incerteza na tensão de limiar, que aumenta a cada nova geração tecnológica [76]. Ainda que os *MOSFETs* de *porta* dupla e levemente dopados [76], [77] evitem dopantes, e conseqüentemente os efeitos da flutuação no número de dopantes, os dispositivos dopados de *porta* simples são os predominantes, e ainda o serão pelos próximos anos. Assim, a previsão dos efeitos do número aleatório de dopantes no descasamento de *MOSFETs* é de suma importância.

O modelo proposto avalia o efeito das variações aleatórias no número de portadores devido às flutuações espaciais na concentração de impurezas. A abordagem convencional leva em conta a flutuação de dopantes sobre toda a região ativa, enquanto nesta proposta o efeito local destas flutuações é explicitamente considerado. A contribuição das flutuações locais é integrada ao longo do canal considerando-se as principais não-linearidades do *MOSFET*. Felizmente o formalismo necessário para se incluir as flutuações locais é

¹⁰ Designação empregada para tecnologias cujo comprimento de canal se situa abaixo dos 200 nm.

disponível na modelagem de ruído *flicker* ou $1/f$, sendo conhecido como *carrier number fluctuation theory* [78].

O descasamento e o ruído interno são fatores limitantes do funcionamento de circuitos eletrônicos. A analogia entre estes efeitos já havia sido previamente citada [13] através da avaliação dos limites por eles estabelecidos para o mínimo consumo de potência, para se atingir uma determinada velocidade e exatidão em um circuito eletrônico. Descasamento (flutuação espacial) e ruído (flutuação temporal) são fenômenos similares, sendo ambos função do processo, das dimensões do dispositivo e de sua polarização. De forma grosseira, o descasamento poderia ser entendido como um “ruído CC”.

Aqui será mostrado que a mesma teoria empregada para se derivar o ruído $1/f$ em transistores *MOS* pode ser também aplicada na modelagem do seu descasamento na corrente.

Para se obter resultados gerais, válidos para todas as regiões de operação do transistor, o modelo *ACM* foi utilizado, sendo este um modelo com bases físicas que cobre todas as regiões de operação através de uma única equação [65], [66]. O modelo aqui apresentado não inclui os efeitos da degradação na mobilidade, saturação de velocidade, ou resistências série. A inclusão desses efeitos é essencial à implementação computacional do modelo, mas leva a expressões complexas cuja discussão não será feita neste momento.

3.6.2 Um modelo consistente para as flutuações na corrente de *dreno*

O desvio na corrente de *dreno*, no entorno do seu valor nominal, resulta da soma de contribuições das flutuações locais ao longo do canal, sejam quais forem suas origens. Para se calcular o efeito dessas flutuações, podemos dividir o transistor em 3 elementos em série, como mostrado na fig. 3.3(a): um transistor superior, um transistor inferior, e um pequeno elemento de canal, com comprimento Δx e área $\Delta A = W\Delta x$. Na fig. 3.3(a), x representa a distância entre o elemento de canal e a *fonte* do transistor.

A flutuação local na corrente ($i_{\Delta A}$) é considerada aqui como sendo um processo estacionário de média zero em relação à variável x . A análise pelo método das perturbações (simpliciter chamado aqui como análise para pequenos-sinais), permite que se calcule o efeito de $i_{\Delta A}$ no desvio da corrente de *dreno* (ΔI_d), como mostrado na fig. 3.3(b).

Observando-se que [65] $g_u = -\mu \frac{W}{L-x} Q'_{Ix}$ e $g_l = -\mu \frac{W}{x} Q'_{Ix}$, a divisão de corrente entre o

elemento de canal e a condutância equivalente do resto do canal, resulta em $\Delta I_d = (\Delta x/L)i_{\Delta A}$. Este simples resultado para a divisão de corrente, proporcional à relação geométrica do comprimento do canal, é uma consequência da formulação por potencial de quasi-Fermi para a corrente de *dreno*, ou seja, a condutância do elemento de canal e as transcondutâncias dos transistores de cima e de baixo são proporcionais à densidade de carga de inversão local [65]-[67]. Assim, considerando-se que as flutuações de corrente ao longo do canal não são correlacionadas, tem-se que o quadrado da flutuação de corrente é

$$\overline{\Delta I_D^2} = \sum_{\text{channel-length}} (\Delta I_d)^2 = \lim_{\Delta x \rightarrow 0} \sum \left(\frac{\Delta x}{L} i_{\Delta A} \right)^2 = \frac{1}{L^2} \int_0^L \Delta x (i_{\Delta A})^2 dx. \quad (3.6.2)$$

Essas flutuações locais na corrente resultam principalmente de três origens físicas independentes: flutuações na concentração de dopantes na região do canal, nos estados de superfície e na espessura do óxido de *porta* [21]. Lembrando que, como $i_{\Delta A}$ está relacionado à flutuação local na área $W\Delta x$, sua variância deve ser proporcional a $1/(W\Delta x)$. Assim como em [21], assumiu-se que a flutuação na dopagem do canal é o principal fator que determina as flutuações locais na corrente.

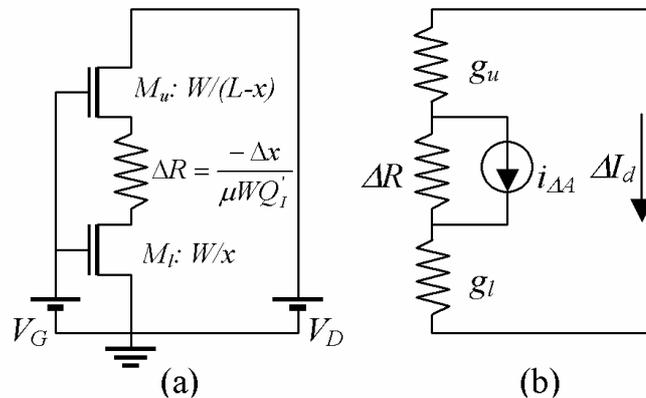


Figura 3.3: Divisão de um transistor em três elementos em série: (a) circuito equivalente com transistores, e (b) equivalente para pequenas perturbações (ou pequenos-sinais).

3.6.3 Um modelo para a flutuação na concentração de dopantes

A flutuação local na corrente de *dreno*, que resulta de uma flutuação local na densidade da carga de inversão, pode ser calculada através de (3.2.4), como em [72], [73], [79], ou por (3.5.9), considerando-se que, sob o ponto-de-vista do elemento do canal, $\Delta x=L$ nesta expressão, resultando

$$i_{\Delta A} = I_D \frac{\Delta Q'_I}{Q'_I}, \quad (3.6.3)$$

onde $\Delta Q'_I$ é a flutuação na densidade de carga de inversão no elemento de canal de área ΔA . Buscando a simplicidade no modelo, considerou-se apenas as flutuações no número de portadores, mas a análise poderia ser estendida para incluir flutuações na mobilidade, por exemplo, como feito em [79] para o ruído $1/f$.

Como em [21], assume-se que as flutuações no número de impurezas são a única causa de flutuações no número de portadores. Para derivar a flutuação na densidade de carga de inversão, o modelo capacitivo do transistor *MOS* pode ser utilizado. O princípio de conservação de cargas no modelo da fig. 3.4 requer que

$$\Delta Q'_B(x) + \Delta Q'_G(x) + \Delta Q'_I(x) = 0, \quad (3.6.4)$$

onde $\Delta Q'_G = -C'_{ox} \Delta \phi_S$ é a flutuação na carga de *porta*, e $\Delta Q'_I = -C'_i \Delta \phi_S$. A variação na carga de depleção $\Delta Q'_B$ resulta da soma de duas componentes, sendo a primeira igual a $-C'_b \Delta \phi_S$ e associada à flutuação no potencial de superfície, e a segunda designada $\Delta Q'_{IMP}$ e associada com a flutuação no número de impurezas ionizadas na região de interesse. Assim, a variação na carga de depleção é dada por $\Delta Q'_B = -C'_b \Delta \phi_S + \Delta Q'_{IMP}$.

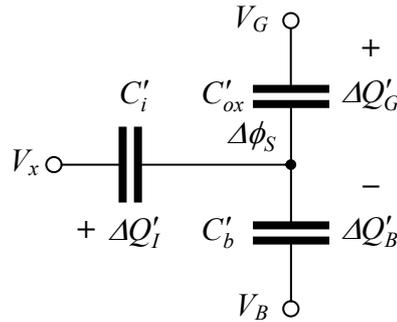


Figura 3.4: Modelo capacitivo do canal do *MOSFET* para a análise de descasamento. A tensão nos terminais é considerada constante.

A definição das capacitâncias, em conjunto com (3.6.4), resulta em

$$\Delta Q'_I = -\frac{C'_i}{C'_i + C'_b + C'_{ox}} \Delta Q'_{IMP}. \quad (3.6.5)$$

Como em [65] e [80], considera-se que $C'_i \cong -Q'_i / \phi_t$ e $C'_b = (n-1)C'_{ox}$ sob qualquer condição de polarização. Assim a expressão (3.6.5) pode ser reescrita como

$$\Delta Q'_i = -\frac{Q'_i}{Q'_i - nC'_{ox}\phi_t} \Delta Q'_{IMP}. \quad (3.6.6)$$

A fig. 3.5 apresenta, de forma esquematizada e exagerada, as flutuações nas cargas de inversão e depleção, causadas pelas flutuações na concentração de dopantes ao longo do canal do *MOSFET*.

A flutuação local do número de impurezas em uma fina fatia elementar horizontal da região de depleção, localizada na posição y em relação à interface óxido-substrato, é calculada assumindo-se que esta é uma variável aleatória com distribuição de Poisson [7], [21]. O número médio de dopantes (\bar{n}_a) nesse volume elementar da região de depleção com comprimento Δx [21] e profundidade Δy is

$$\bar{n}_a = N_a \Delta y W \Delta x, \quad (3.6.7)$$

onde N_a é a concentração total de dopantes (aceitadores e doadores) nesse volume elementar.

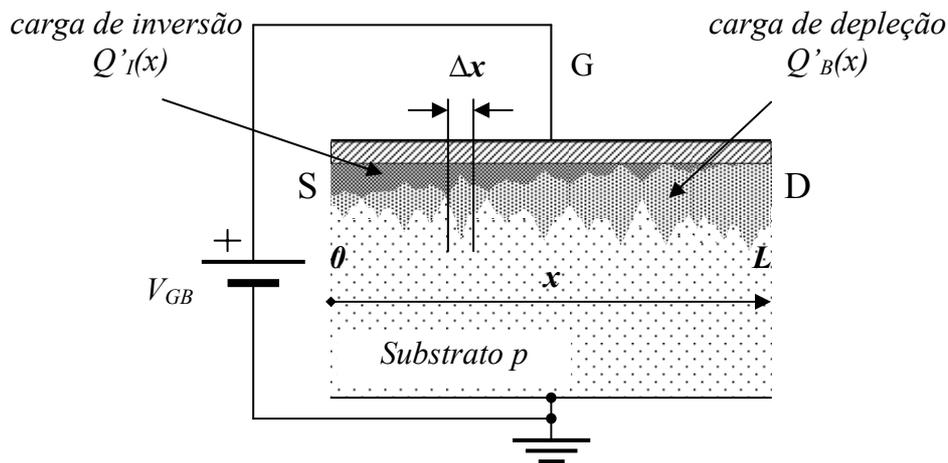


Figura 3.5: Corte longitudinal da região ativa do *MOSFET*, onde as flutuações nas cargas de inversão e depleção, causadas pelas flutuações na concentração de dopantes, estão representadas de forma exagerada.

O desvio-padrão de uma variável com distribuição do tipo Poisson é igual à raiz-quadrada da sua média, assim

$$\sigma^2(n_a) = N_a \Delta y W \Delta x. \quad (3.6.8)$$

Agora, para calcular o desvio-padrão de $\Delta Q'_{IMP}$, pode-se proceder como em [25] e [41], onde assume-se que a contribuição individual das variações locais no número de dopantes, que compõem $\Delta Q'_{IMP}$, não é correlacionada, resultando (vide o *anexo B*)

$$\sigma^2(\Delta Q'_{IMP}) = \frac{q^2}{W \Delta x} \int_0^{y_d} N_a \left(1 - \frac{y}{y_d}\right)^2 dy, \quad (3.6.9)$$

onde q é a carga do elétron, y é a distância para dentro do substrato a partir da interface óxido-semicondutor e y_d é a profundidade da região de depleção.

Usando-se (3.6.3), (3.6.6) e (3.6.9), $(i_{\Delta A})^2$ pode ser calculado e, inserindo-se o resultado em (3.6.2), obtém-se uma expressão para $\overline{\Delta I_D^2}$. Com o auxílio de (3.2.2), a integração sobre o comprimento do canal pode ser alterada para uma integração sobre a densidade de carga de inversão do canal, chegando-se a

$$\sigma_{I_D}^2 = \overline{\Delta I_D^2} = \frac{q^2 \mu I_D}{nC_{ox} L^2} \int_{Q'_{IS}}^{Q'_{ID}} \frac{\int_0^{y_d} N_a \left(1 - \frac{y}{y_d}\right)^2 dy}{nC'_{ox} \phi_t - Q'_I} dQ'_I, \quad (3.6.10)$$

onde Q'_{IS} e Q'_{ID} são as densidades superficiais de carga de inversão nas regiões de *fonte* e *dreno*, respectivamente. A expressão (3.6.10) permite que se avalie o descasamento da corrente de *dreno* em termos da concentração de dopantes na região de depleção, e da polarização, aqui representada pelas densidades de carga de inversão no *dreno* e *fonte*. A principal dificuldade em se calcular a integral de (3.6.10) advém da não-uniformidade do perfil de dopagem do substrato e da variação da profundidade da região de depleção ao longo do canal. Para um perfil constante de dopagem, estimou-se que a introdução de uma profundidade variável da zona de depleção, ao longo do canal, mostra-se geralmente irrelevante, com exceção da operação sob níveis de inversão muito altos (sob saturação, o impacto no descasamento estimado, de uma região de depleção com profundidade variável, é menor que 6% para um nível de inversão de 1000, na tecnologia *TSMC 0.35*).

Com o objetivo de se obter uma expressão simples para o descasamento, pode-se assumir que a integral em (3.6.9) seja constante, o que será considerado daqui por diante, através da seguinte notação

$$N_{oi} = \int_0^{y_d} N_a \left(1 - \frac{y}{y_d}\right)^2 dy. \quad (3.6.11)$$

onde o parâmetro N_{oi} reflete a influência do perfil vertical de dopantes nas flutuações na carga de depleção. Cita-se aqui uma importante conclusão relativa a (3.6.11), feita pelos autores de [41]: “Como a função de ponderação $(1 - y/y_d)^2$ em (3.6.11) apresenta seu valor máximo para $y=0$, a ausência de impurezas nas proximidades da interface se mostra efetiva na redução dos desvios”. Assim, ajustando-se o perfil de distribuição de dopantes, pode-se reduzir o efeito de descasamento, mantendo-se a tensão de limiar (V_T) de interesse.

Finalmente, utilizando-se (3.6.11) e integrando-se (3.6.10) da *fonte* ao *dreno*, chega-se a

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{q^2 N_{oi} \mu}{L^2 n C'_{ox} I_D} \ln \left(\frac{n C'_{ox} \phi_t - Q'_{IS}}{n C'_{ox} \phi_t - Q'_{ID}} \right). \quad (3.6.12)$$

O resultado em (3.6.12) é essencialmente o mesmo que foi derivado para o ruído *flicker* de transistores em [81]. Isso ocorre porque o descasamento se comporta como se fosse um “ruído CC”, sendo que a origem física de ambos efeitos (ruído e descasamento) está relacionada a flutuações ao longo do canal, seja de cargas fixas (descasamento), ou de estados de interface (ruído).

3.6.4 O modelo de descasamento em termos de nível de inversão

Uma expressão alternativa para (3.6.12), muito útil para os projetistas de circuitos, pode ser obtida se as cargas de inversão de *dreno* e *fonte* forem reescritas em termos dos fatores de corrente direta e reversa de um transistor [65], [66], [82]. No modelo *ACM* [65], [66], a corrente de *dreno* é expressa como a diferença entre suas componentes direta (I_F) e reversa (I_R), ou

$$I_D = I_F - I_R = I(V_G, V_S) - I(V_G, V_D) = I_S (i_f - i_r), \quad (3.6.13)$$

onde $I_S = \frac{1}{2} \mu C'_{ox} n \phi_t^2 (W/L)$ é a corrente específica de normalização, proporcional à razão de aspecto W/L do transistor. V_G , V_S , e V_D são os potenciais de *porta*, *fonte* e *dreno*, respectivamente, com referência ao substrato. Os parâmetros i_f e i_r são as correntes normalizadas direta e reversa, ou níveis de inversão na *fonte* e no *dreno*, respectivamente. Observa-se que, na condição de saturação, a corrente é praticamente independente de V_D ,

ou seja, $i_f \gg i_r$ e $I_D \cong I_F$. Por outro lado, caso V_{DS} seja pequeno (região linear), tem-se $i_f \cong i_r$. Usando-se a relação entre a densidade de carga de inversão e a corrente, [65], [66], ou $-Q'_{IS(D)}/nC'_{ox}\phi_t = \sqrt{1+i_{f(r)}} - 1$, e (3.6.13), a expressão (3.6.12) pode ser reescrita como

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{oi}}{WLN^{*2}} \frac{1}{i_f - i_r} \ln\left(\frac{1+i_f}{1+i_r}\right), \quad (3.6.14)$$

onde o parâmetro N^* foi definido como em [79], [81]

$$N^* = \frac{-Q'_{IP}}{q} = \frac{nC'_{ox}\phi_t}{q}, \quad (3.6.15)$$

sendo Q'_{IP} a densidade de carga de inversão na condição de estrangulamento (*pinch-off*).

A expressão (3.6.14) indica que o quadrado do descasamento normalizado é inversamente proporcional à área, relação esta que é amplamente conhecida. Em inversão forte, caso consideremos $n=1$, a expressão (3.6.14) se reduz à expressão (12) de [55]. Ainda, o quadrado do descasamento normalizado é proporcional a t_{ox}^2 e a $N_a^{1/2}$ (pois, para uma concentração homogênea, (3.6.11) se reduz a $N_{oi} = N_a y_d/3$ e $y_d \propto 1/\sqrt{N_a}$). Finalmente, o potencial de substrato também afeta o parâmetro N_{oi} através da modulação da profundidade da região de depleção. Para valores fixos de i_f e i_r , o aumento da polarização reversa entre substrato e *fonte* aumenta a profundidade da região de depleção, conseqüentemente aumentando N_{oi} . Como resultado, pode-se concluir que há uma piora no casamento entre transistores, resultante do aumento da polarização reversa entre substrato e *fonte*.

A expressão (3.6.14) pode ser simplificada sob condições específicas. Em regime linear, da inversão fraca à forte, tem-se $i_f \cong i_r$ e (3.6.14) é reduzida a

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{oi}}{WLN^{*2}} \frac{1}{1+i_f}. \quad (3.6.16a)$$

Esta última expressão pode também ser escrita como

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{oi}}{WLN^{*2}} \left(n\phi_t \frac{g_m}{I_D} \right)^2. \quad (3.6.16b)$$

Sob inversão fraca, tem-se $i_f \ll I$, de modo que uma expansão em uma série de primeira ordem de (3.6.14) resulta em

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{oi}}{WLN^{*2}}, \quad (3.6.17)$$

válida tanto para condição linear quanto para saturação.

Sob saturação, tem-se $i_r \rightarrow 0$, de modo que a expressão (3.6.14) pode ser escrita como

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{oi}}{WLN^{*2}} \frac{\ln(1+i_f)}{i_f}. \quad (3.6.18a)$$

Esta última expressão pode também ser escrita como

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{oi}}{WLN^{*2}} \frac{\ln(1+I_D/I_S)}{I_D/I_S}. \quad (3.6.18b)$$

Visando tornar-se o modelo mais completo, pode-se ainda incluir nele os erros aleatórios relacionados à corrente específica de folha $I_{SQ} = \frac{1}{2}\mu C'_{ox} n \phi_t^2$, como feito em [6], o que resulta na seguinte modificação da expressão (3.6.14)

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{1}{WL} \left[\frac{N_{oi}}{N^{*2}} \frac{1}{i_f - i_r} \ln\left(\frac{1+i_f}{1+i_r}\right) + B_{ISQ}^2 \right]. \quad (3.6.19)$$

Nesta expressão, B_{ISQ} representa um parâmetro de descasamento que, em uma primeira aproximação, é um fator constante que engloba variações de mobilidade, espessura de óxido de *porta*, e do fator de rampa. Um modelo mais elaborado poderia ainda considerar B_{ISQ} como um termo dependente da polarização.

3.6.5 Resultados experimentais

O descasamento intra-pastilha (*intradie*) da corrente foi medido utilizando-se um circuito-teste fabricado na tecnologia *TSMC 0.35μm 3.3V CMOS n-well*, contendo um conjunto de transistores NMOS e PMOS. Esse processo apresenta espessura de óxido de *porta* de 78 angstroms. No circuito-teste, os transistores são distribuídos em arranjos de 20 idênticos dispositivos funcionais, com a mesma orientação, cuja periferia é terminada por dispositivos *dummy*, com o objetivo de garantir condições de contorno uniformes. As

dimensões dos transistores ($W \times L$) de cada arranjo são: $12\mu\text{m} \times 8\mu\text{m}$ (*grande*), $3\mu\text{m} \times 2\mu\text{m}$ (*médio*), $0,75\mu\text{m} \times 8\mu\text{m}$ (*estreito* – largura mínima), $12\mu\text{m} \times 0,5\mu\text{m}$ (*curto* – comprimento mínimo) e $0,75\mu\text{m} \times 0,5\mu\text{m}$ (*pequeno* – tamanho mínimo). No leiaute foram empregadas trilhas de conexão largas e múltiplas janelas de contato, de forma a se reduzir as quedas ôhmicas das conexões [14]. Todos os dez circuitos caracterizados, de um lote de quarenta, apresentaram um comportamento similar de descasamento.

O circuito apresentado na fig. 3.6 foi utilizado nos ensaios de descasamento. V_D , $V'_D (= V_D)$, e V_B são fontes de tensão externas, e I_B é uma fonte de corrente, empregada na polarização do transistor de referência M_{REF} . As unidades de estímulo e medida (*source/monitor unit* - SMU) do analisador de parâmetros de semicondutores HP4145B foram empregadas no ensaio. O mesmo transistor M_{REF} foi utilizado para todas as medidas de um mesmo arranjo, enquanto os 19 transistores restantes foram medidos em pares adjacentes, M_i e M_{i+1} ($i=1, \dots, 18$). A técnica de medida diferencial foi empregada [83] para cada par de dispositivos adjacentes, pois assim é possível se ressaltar o descasamento local, com menor influência de efeitos globais.

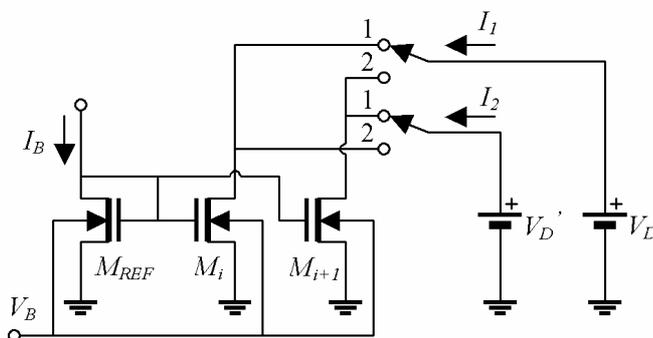


Figura 3.6: Circuito para ensaio experimental - M_{REF} é um transistor de referência, enquanto M_i e M_{i+1} são os transistores sob teste. I_B (V_B , $V'_D=V_D$) é uma fonte de corrente (tensão) para polarização.

Cada par de transistores M_i e M_{i+1} , $i=1, \dots, 18$, foi seqüencialmente caracterizado, com as correntes de ambos (I_1 e I_2) sendo medidas simultaneamente, para cada condição de polarização, com as chaves em cada uma das posições 1 e 2. A corrente cc que passava por cada dispositivo, $I_{D(i)}$ e $I_{D(i+1)}$, foi armazenada, após feita a média das duas situações de medida para cada transistor (chaves em 1 ou 2). Esse procedimento reduz os erros que possam resultar do descasamento entre as duas SMUs empregadas para se medir os dois transistores. No total, são necessárias de três a quatro SMUs durante o ensaio: duas para

excitar e medir os drenos dos transistores sob teste, uma terceira para polarização em corrente do transistor de referência, e uma quarta (opcional) para a polarização em tensão do terminal de substrato dos transistores (se desejado).

O descasamento normalizado para cada arranjo foi calculado utilizando-se a seguinte expressão

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{1}{2(N-1)I_D^2} \sum_{i=1}^N (I_{D(i)} - I_{D(i+1)})^2, \quad (3.6.20)$$

onde N representa o número total de pares adjacentes em cada grupo de transistores idênticos ($N = 18$ para as estruturas de teste ensaiadas). O fator 2 no denominador de (3.6.20) é necessário para se converter a variância medida a partir de um par, na variância de um único dispositivo [83], [84].

As figs. 3.7 a 3.10 apresentam o descasamento normalizado para uma tensão *dreno-fonte* variando de $+(-)10\text{mV}$ (região linear) até $+(-)2\text{V}$ (saturação) para os transistores *grandes* e *médios* NMOS (PMOS). O descasamento foi medido sob seis níveis de inversão diferentes, cobrindo assim uma faixa de cinco décadas completas (0,01; 0,1; 1; 10; 100; e 1000). A tensão *substrato-fonte* foi mantida em zero volts. As curvas simuladas através deste modelo resultam de (3.6.19), com i_r calculado através do modelo *ACM* para canal longo [65], [66].

Da inversão moderada ($i_f = 1$ e 10) à forte ($i_f = 100$), as curvas simuladas e medidas apresentam comportamento semelhante, aumentando a partir da região linear, até a saturação, onde seu valor estabiliza. Algumas diferenças entre as curvas simuladas e medidas, mais aparentes nos dispositivos *médios* do que nos *grandes*, podem ser associadas à não-uniformidade espacial da concentração de átomos dopantes [22].

O parâmetro N_{oi} foi estimado a partir das medidas sob inversão fraca, utilizando-se (3.6.17). O parâmetro N^* foi calculado com base nos parâmetros do processo utilizado, fornecidos pela *MOSIS* (www.mosis.org). A largura e o comprimento efetivos do canal (W_{eff} e L_{eff}) foram calculados [27] através dos parâmetros do modelo BSIM3v3 *WINT* e *LINT* ($0,065\mu\text{m}$ e $0,075\mu\text{m}$, respectivamente) [85], também fornecido pela *MOSIS*. A partir dos arranjos de transistores *grande* e *médio* foi obtido o mesmo valor de N_{oi} para os dispositivos NMOS, $1,8 \times 10^{12} \text{ cm}^{-2}$, e $7 \times 10^{12} \text{ cm}^{-2}$ para os PMOS.

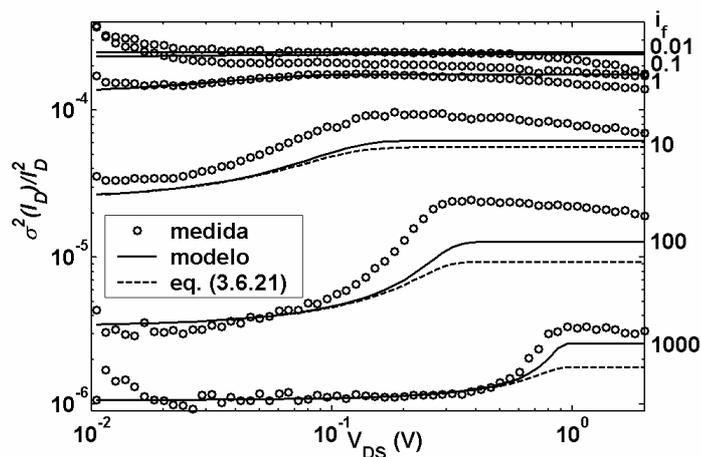


Figura 3.7: Descasamento normalizado da corrente para o arranjo *grande* NMOS. A junção substrato-*fonte* foi mantida a zero volt. As curvas do modelo resultam de (3.6.19).

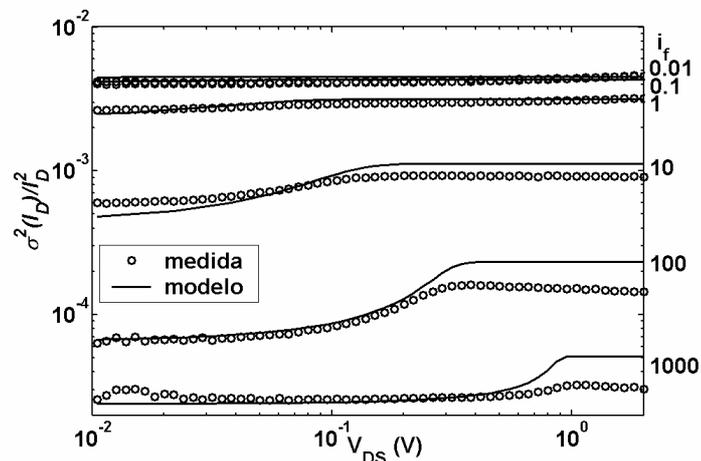


Figura 3.8: Descasamento normalizado da corrente para o arranjo *médio* NMOS. A junção substrato-*fonte* foi mantida a zero volt. As curvas do modelo resultam de (3.6.19).

Sob inversão fraca ($i_f = 0,01$ e $0,1$), o descasamento é quase constante da região linear até a saturação, como previsto por (3.6.17). As curvas simuladas e medidas sob inversão fraca são quase coincidentes, sendo de difícil diferenciação.

Deve-se observar que N_{oi} inclui tanto as impurezas doadoras, quanto as aceitadoras [84]. Como consequência, N_{oi} é geralmente maior que o valor dado pelo produto da concentração líquida de dopantes pela profundidade da região de depleção.

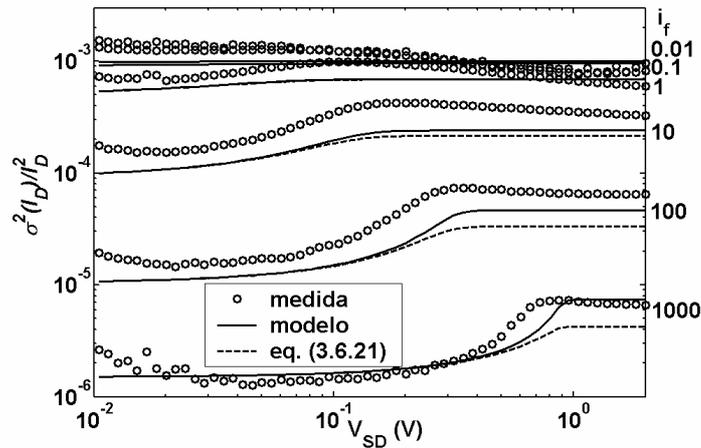


Figura 3.9: Descasamento normalizado da corrente para o arranjo *grande* PMOS. A junção substrato-*fonte* foi mantida a zero volts. As curvas do modelo resultam de (3.6.19).

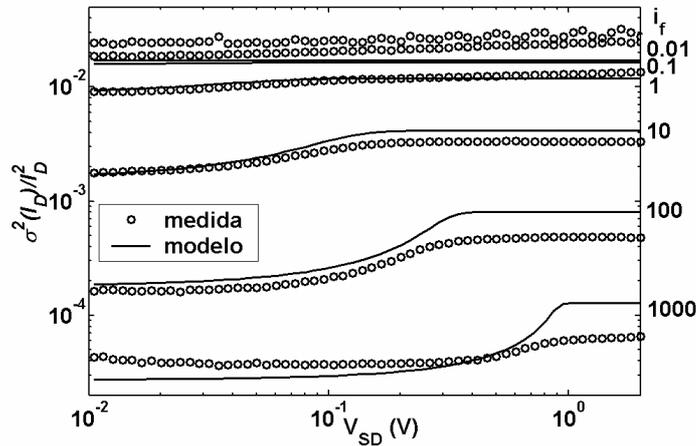


Figura 3.10: Descasamento normalizado da corrente para o arranjo *médio* PMOS. A junção substrato-*fonte* foi mantida a zero volts. As curvas do modelo resultam de (3.6.19).

Considerando-se as concentrações líquidas de dopantes obtidas através dos parâmetros do modelo fornecido pela *MOSIS* (calculadas através de $\gamma_{N(P)}C'_{ox} = \sqrt{2\epsilon_{Si}qN_{A(D)}}$ e da aproximação do fator de corpo $\gamma \approx KI$, onde KI é um dos parâmetros do modelo BSIM, resultando $N_A = 2,1 \times 10^{17} \text{ cm}^{-3}$ e $N_D = 1,1 \times 10^{17} \text{ cm}^{-3}$), e aproximando para o caso de concentração homogênea de impurezas, onde $N_{oi} = N_a y_d/3$, pode-se estimar a profundidade da região de depleção como $y_{dN} = 0,25 \text{ }\mu\text{m}$ e $y_{dP} = 1,9 \text{ }\mu\text{m}$. Essa estimativa é totalmente irreal, apresentando ao menos três erros importantes, que podem colaborar para o absurdo dos valores obtidos, principalmente no caso do PMOS: (1) as concentrações de dopantes na região ativa são dadas em valor *líquido*, enquanto N_a corresponde à concentração *total* de impurezas; (2) os valores utilizados nos parâmetros do

modelo BSIM representam mais elementos de ajuste de suas curvas, que quantidades físicas do dispositivo; e principalmente (3) em dispositivos sub-micrométricos a concentração de dopantes não é homogênea na sua região ativa.

Sob alto nível de inversão, o componente de descasamento de (3.6.19) associado com N_{oi} passa a ser da mesma ordem ou menor que a contribuição associada a B_{ISQ} . Assim, para altos níveis de inversão, o descasamento permanece constante em um mínimo determinado por B_{ISQ} , como pode ser observado, por exemplo, na fig. 3.8. Experimentalmente foi observado que a razão entre o descasamento mínimo sob inversão forte e o máximo descasamento medido sob inversão fraca, é a mesma para os arranjos *grande e médio*.

O parâmetro B_{ISQ} foi estimado a partir das medidas sob inversão forte, na região linear, utilizando-se (3.6.19). B_{ISQ} da ordem de 0,89 %- μm e 0,71 %- μm foi extraído para os dispositivos NMOS e PMOS, respectivamente, tanto para os *grandes*, como para os *médios*. As curvas simuladas apresentadas nas figs. 3.7 a 3.10 são baseadas nos valores extraídos de N_{oi} e B_{ISQ} , para os transistores NMOS e PMOS.

Alguns autores [48] sugerem que os dispositivos PMOS apresentam melhor casamento que os NMOS. Outros citam o contrário [42]. Os resultados obtidos neste trabalho indicam que o descasamento depende de vários detalhes intrínsecos de cada processo, como os perfis de dopagem vertical e horizontal (implante *halo*, disponibilidade de duplo-poço, implantes superficiais de ajuste, implante retrógrado, etc). Como pode ser visto nestas medidas, para dispositivos com mesma geometria, nível de inversão e tensão de *dreno*, os PMOS (em poço de compensação N) apresentam descasamento superior ao dos NMOS. Outros autores obtiveram resultados experimentais onde os PMOS apresentam descasamento superior aos NMOS [43], [86], enquanto outros ainda obtiveram o oposto [55]. Conclui-se assim que não há uma simples “regra geral” que informe qual tipo de transistor *MOS* apresenta melhor casamento, contrariando o senso comum de muitos autores.

As figs. 3.11 e 3.12 apresentam a dependência do descasamento no nível de inversão, para as regiões linear e de saturação, para três tamanhos diferentes de transistores NMOS, sob duas polarizações diferentes de substrato-*fonte* (V_{BS}). Através dessas figuras, pode-se observar que os transistores maiores seguem a “regra da área”, como apresentado no modelo. Para uma polarização específica de substrato, foi utilizado o mesmo N_{oi} para a

simulação dos transistores *grande* e *médio*, nas regiões linear e de saturação. Também foi utilizado o mesmo valor de B_{ISQ} para a simulação do descasamento dos transistores *grande* e *médio* em ambas as polarizações de substrato.

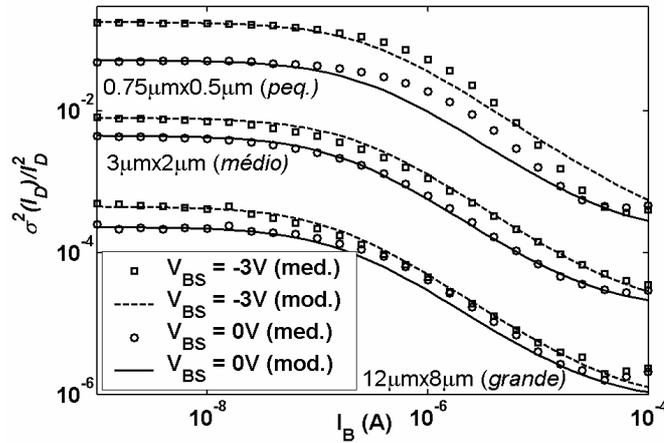


Figura 3.11: Descasamento normalizado da corrente medido (méd.) na região linear ($V_{DS}=20\text{mV}$), em função do nível de inversão, para os arranjos *grande*, *médio* e *pequeno* NMOS, sob duas tensões de substrato-fonte. As curvas do modelo (mod.) resultam de (3.6.19).

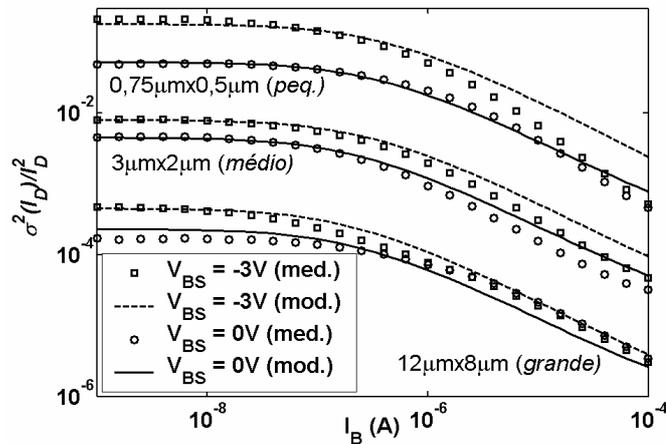


Figura 3.12: Descasamento normalizado da corrente medido (méd.) na saturação ($V_{DS}=2V$), em função do nível de inversão, para os arranjos *grande*, *médio* e *pequeno* NMOS, sob duas tensões de substrato-fonte. As curvas do modelo (mod.) resultam de (3.6.19).

Os transistores *pequenos* (mínimos) não seguem esta regra, apresentando um descasamento 55% menor que o previsto pelo modelo (sob polarização de substrato igual a zero) utilizando o mesmo N_{oi} . Com $V_{BS} = -3V$, o descasamento medido para os dispositivos

pequenos apresenta boa concordância com a estimativa do modelo. Entretanto, utilizou-se para estes dispositivos valores de N_{oi} diferentes daqueles medidos para os dispositivos *grandes*, de modo a se obter um melhor ajuste às curvas experimentais. Para as pastilhas caracterizadas, os dispositivos *pequenos* apresentaram um N_{oi} imprevisível, como já observado nas referências [10] e [23].

De fato, as características elétricas de dispositivos de canal curto são muito sensíveis a flutuações devido à grande dependência dos efeitos de borda (deformações progressivas em campos e concentrações, devido aos efeitos de fronteira que ocorrem nos limites da região do canal, ou à dopagem com concentração não-homogênea, utilizada para compensar alguns efeitos de fronteira). Esta alta sensibilidade dos dispositivos de canal curto é uma das principais razões para as dificuldades encontradas na modelagem do descasamento, principalmente nas complexas tecnologias sub-micrométricas atuais. Ainda, nos dispositivos de canal mínimo, as regiões dopadas de *dreno* e *fonte* são muito próximas entre si, afetando fortemente o perfil da região de depleção sob o canal. Para os dispositivos *pequenos*, B_{ISQ} foi calculado usando-se a mesma proporção do N_{oi} utilizado nestes transistores em relação ao utilizado para os *grandes/médios*. Como demonstrado experimentalmente, este modelo apresenta uma boa aproximação até para os dispositivos de dimensões mínimas, ainda que seja necessário o ajuste dos parâmetros N_{oi} e B_{ISQ} . Uma boa estratégia para a modelagem do descasamento de dispositivos de canal curto, pode ser a definição de uma faixa de valores “mínimo-máximo” para N_{oi} [9]. Em um projeto mais conservador, o máximo valor de N_{oi} seria utilizado, de modo a se prever o pior caso do descasamento.

Das figs. 3.11 a 3.15, pode-se observar que o descasamento da corrente aumenta com a polarização reversa de substrato. O motivo desse aumento advém do acréscimo na profundidade da região de depleção resultante de uma maior polarização reversa do substrato, resultando em um maior N_{oi} [44], [45], [46], [87]. A dificuldade em se desenvolver um modelo para a modulação de N_{oi} com a polarização de substrato fez com que se preferisse, em um primeiro momento, escolher valores de N_{oi} que se ajustassem às medidas para cada condição de polarização de substrato.

A fig. 3.13 apresenta o descasamento da corrente para os arranjos de transistores NMOS *estreito e curto*. Observa-se que este modelo pode ser aplicado a transistores com

comprimento ou largura mínima, apresentando uma boa aproximação com os valores medidos.

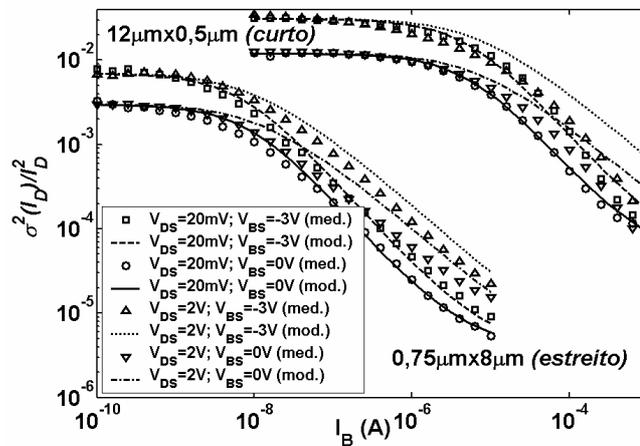


Figura 3.13: Descasamento normalizado da corrente medido (méd.) nas regiões linear e de saturação, em função do nível de inversão, para os arranjos *estrito* e *curto* NMOS, sob duas tensões de substrato-fonte. As curvas do modelo (mod.) resultam de (3.6.19).

As figs. 3.14 e 3.15 apresentam os resultados obtidos dos arranjos PMOS. Para estes, a polarização de substrato parece ter menor impacto que para os NMOS. Os resultados medidos apresentam-se também de acordo com a “regra da área”, com exceção novamente dos dispositivos mínimos. O valor de N_{oi} utilizado para a simulação das curvas dos dispositivos *pequenos* apresenta-se 80% maior que o estimado para os grandes PMOS. A partir dos resultados obtidos dos dispositivos fabricados nesta tecnologia, tem-se que, para uma determinada geometria e polarização equivalentes, os dispositivos PMOS apresentaram um maior descasamento que os NMOS.

Além da flutuação na concentração de dopantes no canal, flutuações na dopagem de *porta* e variações geométricas também são fatores relevantes de descasamento [8], [25], [37], [46], [88]. Vários autores têm demonstrado experimentalmente que o primeiro fator é o dominante no descasamento da tensão de limiar (V_T – resultando em descasamento de corrente), sendo o segundo também importante para processos sub-micrométricos, e o terceiro o menos relevante, em geral. Como pode ser visto, outras fontes de descasamento poderiam ser incluídas neste modelo, mas neste momento ainda se preferiu mantê-lo o mais simples possível.

Resultados experimentais obtidos de outra tecnologia CMOS, e que foram publicados em [89] e [90], também corroboram o modelo aqui apresentado.

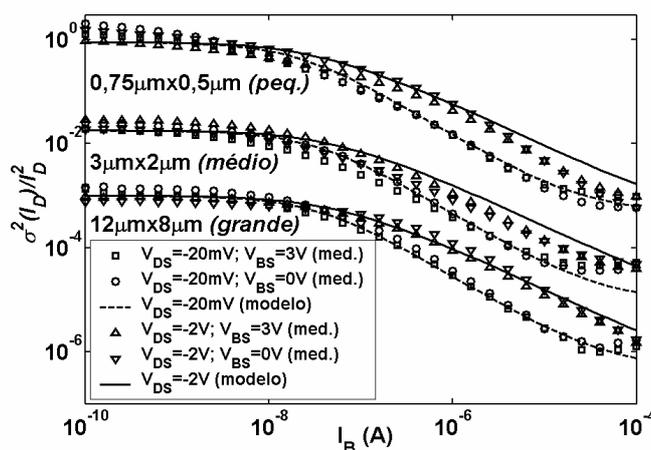


Figura 3.14: Descasamento normalizado da corrente medido (méd.) nas regiões linear e de saturação, em função do nível de inversão, para os arranjos *grande*, *médio* e *pequeno* PMOS, sob duas tensões de substrato-*fonte*. As curvas do modelo resultam de (3.6.19).

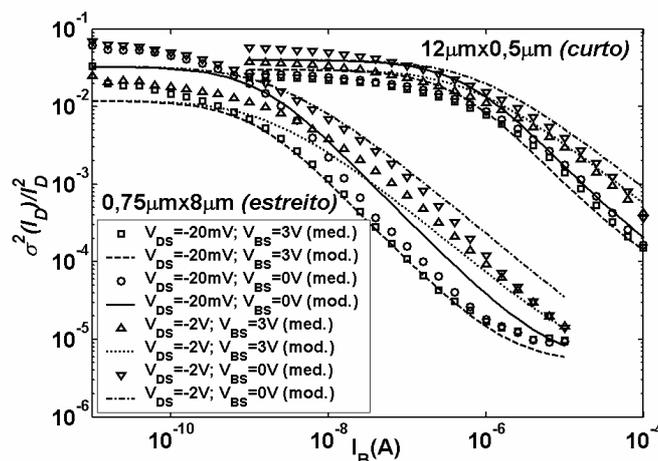


Figura 3.15: Descasamento normalizado da corrente medido (méd.) nas regiões linear e de saturação, em função do nível de inversão, para os arranjos *estreito* e *curto* PMOS, sob duas tensões de substrato-*fonte*. As curvas do modelo (mod.) resultam de (3.6.19).

3.6.6 Conclusões

Aqui foi descrito um modelo que foi desenvolvido para o descasamento entre transistores *MOS*, e que é contínuo sob qualquer condição de operação, e consistente para associações série. O enfoque proposto na modelagem é baseado na integração do efeito da variação aleatória do número de portadores ao longo do canal. Essa estratégia, em conjunto com o uso do modelo *CC ACM*, resultou em uma expressão compacta para o descasamento, de fácil uso e que cobre todas as condições de operação. Os resultados aqui obtidos estão intimamente relacionados àqueles já derivados em [81] para o ruído $1/f$, o

que se deve ao fato de que os mecanismos na origem dos dois fenômenos são similares. A partir deste trabalho, conclui-se ainda que o uso da variação de parâmetros globais, como a tensão de limiar, não é apropriado para a descrição do descasamento, em decorrência da distribuição não-linear do efeito dos portadores ao longo do canal do transistor. Foi demonstrado como se pode incluir no modelo as flutuações aleatórias tanto dos átomos dopantes, quanto da corrente específica. Um conjunto de grupos de transistores idênticos foi fabricado em uma tecnologia CMOS 0,35 μm para que se pudesse avaliar a influência da geometria e da polarização no descasamento. Os resultados experimentais confirmam a exatidão do modelo sob uma ampla faixa de geometrias e polarizações. Para a tecnologia sob análise, pode-se concluir que o fator determinante no descasamento é o N_{oi} , que está relacionado ao número médio de dopantes por unidade de área na região de depleção sob o canal. Espera-se que este trabalho possa trazer uma nova luz à modelagem do descasamento de *MOSFETs*, ajudando os projetistas de circuitos a prever o descasamento através de apenas um par de parâmetros (N_{oi} e B_{ISQ}).

3.6.7 Apêndice

Através do modelo de descasamento de Pelgrom [8], em conjunto com o modelo *ACM* [65], [66], tem-se derivada a seguinte expressão para o descasamento normalizado

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{A_{VT}^2}{WL} \left(\frac{g_m}{I_D} \right)^2 = \frac{N_{oi}}{WLN^{*2}} \left(\frac{2}{\sqrt{1+i_f} + \sqrt{1+i_r}} \right)^2, \quad (3.6.21)$$

onde sua dependência em i_f é a mesma de $(g_m/I_D)^2$ em i_f . Ambas expressões, (3.6.14) e (3.6.21), tendem a (3.6.17) em inversão fraca ($i_f \ll 1$), tornando-se ambas quase insensíveis à corrente. Mas, sob inversão forte e saturação ($i_f \gg 1$, $i_r \rightarrow 0$) elas divergem. Por exemplo, no caso de $i_f=1000$ e $i_r=0$, a expressão (3.6.14) prediz um valor 80% superior que (3.6.21). A explicação para essa diferença pode ser atribuída à natureza do *MOSFET*. Enquanto o modelo de Pelgrom assume um V_T único e global para o *MOSFET*, o modelo aqui apresentado supõe um V_T distribuído ao longo do canal. Como consequência, sob inversão forte e saturação, a parte do canal mais próxima ao *dreno* passa a ter um papel menos importante na flutuação da carga ao longo do canal que a parte mais próxima à *fonte*. As figs. 3.7 e 3.9 apresentam também a equação (3.6.21) para comparação.

Também, ainda que o modelo de Pelgrom, quando usado em conjunto com a expressão para g_m/I_D obtida do modelo *ACM*, forneça uma boa estimativa para o descasamento, sob níveis de inversão não muito altos, este é inconsistente para associações de transistores em série, conforme já citado por [49], [50] e [55].

3.7 Efeitos secundários que afetam o descasamento

No item anterior, foi desenvolvido um modelo estatístico preditivo para o descasamento entre transistores *MOS*, que leva em conta o comportamento não-linear das relações entre cargas e potenciais. Entretanto, esse modelo foi baseado na aproximação de canal longo do modelo *ACM*, de modo que muitos efeitos chamados “secundários” foram negligenciados.

A seguir, alguns desses efeitos são analisados e modelados de forma a se prever seu impacto no modelo de descasamento.

3.8 Efeito da saturação na velocidade dos portadores

Na modelagem do deslocamento de cargas elétricas em materiais semicondutores, o parâmetro *mobilidade elétrica de portadores* (μ_0) é utilizado para representar a relação entre a velocidade em que estas cargas trafegam (v_d) e a intensidade de campo elétrico (E) que provoca este movimento, de modo que, em uma primeira aproximação,

$$\mu_0 = v_d / E . \quad (3.8.1)$$

Em função da interação das cargas em movimento com a estrutura atômica do semicondutor por onde elas trafegam, a velocidade de deslocamento apresenta uma limitação superior, ou *saturação*, representada pelo parâmetro v_{sat} . Esta limitação também pode ser representada através do valor de um *campo elétrico crítico* (E_C), a partir do qual v_d não é mais proporcional a E .

Na modelagem de transistores *MOS*, o modelo mais simples de mobilidade elétrica, e que incorpora este efeito, estabelece que

$$\mu_{sat} = \frac{\mu_0}{1 + E_x / E_C} , \quad (3.8.2)$$

onde E_x corresponde ao valor do campo elétrico existente na direção x do canal (longitudinal), junto à superfície, que é onde se encontra a maior concentração de cargas em movimento, constituindo a corrente de *dreno*.

A expressão (3.8.2) estabelece uma boa aproximação para a dependência da mobilidade com o campo elétrico longitudinal nas regiões de assíntotas (E_x muito abaixo ou muito acima de E_C). Entretanto, na região de transição de comportamento (E_x da mesma ordem de E_C), o valor de mobilidade resultante de (3.8.2) é de 30 a 50% inferior ao medido experimentalmente. Assim, uma expressão que melhor aproxima o comportamento da mobilidade, inclusive na região de transição, é dada por [91]

$$\mu_{sat} = \frac{\mu_0}{\sqrt{1 + (E_x/E_C)^2}}. \quad (3.8.3)$$

Geralmente, utiliza-se em modelagem compacta a expressão (3.8.2), mais pela sua simplicidade que pela sua exatidão, implicando em expressões finais mais simples e de mais fácil integração ao longo do canal, estratégia que será adotada aqui.

3.8.1 Inclusão na corrente de *dreno*

A expressão de Pao-Sah [92] para a corrente de *dreno* é

$$I_D = -\mu W Q'_I(x) \frac{dV_x}{dx}, \quad (3.8.4)$$

onde μ representa a mobilidade local de portadores, W é a largura da região considerada do canal, $Q'_I(x)$ é a densidade superficial da carga de inversão (portadores) na coordenada x (direção do comprimento do transistor) e V_x é o potencial do canal na mesma posição.

Através do modelo UCCM [68], tem-se a relação entre a carga de inversão e o potencial do canal em uma determinada coordenada x , ou

$$\frac{V_P - V_x}{\phi_t} = \frac{Q'_I}{Q'_{IP}} + \ln\left(\frac{Q'_I}{Q'_{IP}}\right) - 1, \quad (3.8.5)$$

que, se reescrita na forma diferencial, resulta

$$dV_x = \left(\frac{1}{nC'_{ox}} - \frac{\phi_t}{Q'_I} \right) dQ'_I = -\phi_t \left(1 + \frac{1}{q'_I} \right) dq'_I. \quad (3.8.6)$$

O termo mais à direita de (3.8.6) está apresentado sob a forma normalizada, onde $q'_I = Q'_I/Q'_{IP}$. No modelo ACM, Q'_{IP} representa a densidade de carga correspondente ao produto da capacitância efetiva do canal, $C'_{ox} + C'_b = nC'_{ox}$, pelo potencial térmico, ou $Q'_{IP} = -nC'_{ox}\phi_t$, sendo muito utilizado como carga de normalização. É chamado de *densidade de carga de inversão na condição de estrangulamento (pinch-off)* por fatores históricos.

Em alguns desenvolvimentos analíticos, a forma normalizada de uma expressão facilita seu manuseio, pois reduz a quantidade de parâmetros. Neste texto, algumas grandezas poderão ser apresentadas na forma normalizada, quando representadas através da variável principal em caractere minúsculo. As condições de normalização para corrente, tensão, densidade superficial de carga e distância do terminal de *fonte* ao longo do comprimento do transistor (x), são apresentadas a seguir.

$$i = I/I_S, \text{ (onde } I_S = (W/2L)\mu nC'_{ox}\phi_t^2 \text{)} \quad (3.8.7a)$$

$$v = V/\phi_t \quad (3.8.7b)$$

$$q' = Q'/Q'_{IP} \quad (3.8.7c)$$

$$\xi = x/L \quad (3.8.7d)$$

Lembrando que, no modelo de folha de carga [67], o campo elétrico longitudinal é representado através do potencial de superfície ($E_x = -d\phi_S/dx$), e utilizando-se a relação linear entre este potencial e a carga de inversão ($d\phi_S = dQ'_I/(nC'_{ox})$), junto com (3.8.6), pode-se reescrever (3.8.4) como

$$I_D = \frac{-\mu_0 W Q'_I}{1 + \frac{1}{nC'_{ox} E_C} \frac{dQ'_I}{dx}} \left(\frac{\phi_t}{Q'_I} - \frac{1}{nC'_{ox}} \right) \frac{dQ'_I}{dx}, \quad (3.8.8a)$$

ou na forma normalizada

$$i_D = -2 \frac{1 + q'_I}{1 - \zeta \frac{dq'_I}{d\xi}} \frac{dq'_I}{d\xi}, \quad (3.8.8b)$$

onde $\zeta = \frac{\phi_t}{L.E_C} = \frac{\mu_0\phi_t}{L.v_{sat}}$. O fator ζ pode ser entendido como a relação entre a velocidade de difusão ($\mu_0\phi_t/L$) e a velocidade de saturação (v_{sat}).

Integrando-se (3.8.8) ao longo do canal, da *fonte* ao *dreno*, obtém-se

$$I_D = \frac{\overbrace{\mu_0 n C'_{ox} \phi_t^2}^{I_s} \overbrace{W}^{sat. velocidade}}{2 L} \left(\frac{1}{1 + \frac{\phi_t}{L.E_C} \left(\frac{Q'_{IS} - Q'_{ID}}{Q'_{IP}} \right)} \right) \left(\overbrace{\frac{Q'^2_{IS} - Q'^2_{ID}}{Q'^2_{IP}}}^{deriva} + 2 \overbrace{\frac{Q'_{IS} - Q'_{ID}}{Q'_{IP}}}^{difusão} \right), \quad (3.8.9a)$$

ou, na forma normalizada

$$\begin{aligned} i_D &= \frac{q'^2_{IS} - q'^2_{ID} + 2(q'_{IS} - q'_{ID})}{1 + \zeta(q'_{IS} - q'_{ID})} = \frac{(q'_{IS} + 1)^2 - (q'_{ID} + 1)^2}{1 + \zeta(q'_{IS} - q'_{ID})} = \dots \\ &\dots = \frac{(q'_{IS} + q'_{ID} + 2)(q'_{IS} - q'_{ID})}{1 + \zeta(q'_{IS} - q'_{ID})}. \end{aligned} \quad (3.8.9b)$$

Na expressão (3.8.9a) pode-se identificar a corrente específica de normalização, o fator devido à saturação na velocidade dos portadores, e as componentes de deriva e de difusão da corrente de *dreno*. Observa-se nessa expressão que a inclusão do efeito de saturação na velocidade de portadores aparece na forma final da corrente como um fator de atenuação multiplicativo. A expressão normalizada (3.8.9b) foi apresentada em três formas equivalentes, que serão convenientes para deduções posteriores.

Considerando-se em (3.8.9b) a relação entre as densidades de carga de inversão nas regiões de *dreno* e *fonte*, com os níveis de inversão nessas regiões, ou $(q'_{IS(D)} + 1)^2 = 1 + i_{f(r)}$, tem-se que

$$i_D = \frac{i_f - i_r}{1 + \zeta(\sqrt{1 + i_f} - \sqrt{1 + i_r})}, \quad (3.8.9c)$$

que é a forma equivalente da expressão clássica do modelo *ACM* para a corrente, em termos dos níveis de inversão no transistor canal-longo, mas com o efeito da saturação na velocidade de portadores incluído.

A expressão (3.8.9) representa uma forma compacta da corrente de *dreno* em termos das densidades de carga de inversão (ou dos níveis de inversão) nas regiões de

dreno e *fonte*, incluindo as correntes de deriva e de difusão, o que faz com que seja válida continuamente desde a condição de inversão fraca até a forte.

3.8.2 Desvio na corrente de *dreno* devido a flutuações locais

O desvio na corrente de *dreno*, causado por uma flutuação local que ocorra em uma área infinitesimal do canal do transistor, pode ser estimado através da representação do canal como uma associação série de dois transistores, unidos por esta região. A perturbação é representada através duas fontes de corrente ($i_{\Delta A}$), conforme a fig. 3.16. Este método é conhecido como “*impedance field method*” [93].

Inicialmente, desdobra-se o canal do *MOSFET* em dois transistores em série (M_1 e M_2), que podem ser representados pelos seus modelos para pequenas variações, de modo que se possa calcular o reflexo que uma perturbação local ($i_{\Delta A}$) causa na corrente de *dreno* (ΔI_d), conforme a fig. 3.17. Nesta figura, a posição em relação ao comprimento do canal está indicada de forma absoluta (x) e normalizada (ζ).

Através do modelo da fig. 3.17b, pode-se estabelecer a relação entre as duas correntes como

$$\Delta I_d = \frac{g_2}{g_1 + g_2} i_{\Delta A}, \quad (3.8.10)$$

onde g_1 e g_2 são as condutâncias dos dois segmentos de canal, correspondentes aos transistores M_1 e M_2 , vistos através do nó localizado na posição x .

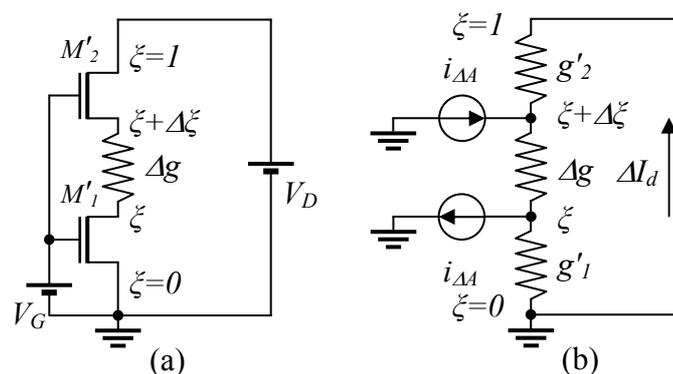


Figura 3.16: Representação do transistor *MOS* através de dois transistores, em série com a região onde ocorre a flutuação (a), e seu equivalente para pequenas variações (b).

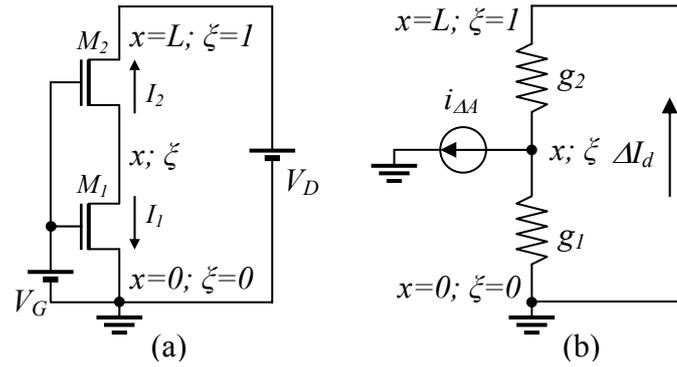


Figura 3.17: Representação do transistor *MOS* através de dois transistores em série (a), e seu equivalente para pequenas variações (b).

Aplicando-se a superposição de efeitos ao modelo da fig. 3.16b, em conjunto com a expressão (3.8.10), pode-se estabelecer a relação entre a flutuação da corrente de *dreno* (ΔI_d), e a flutuação local no segmento de canal determinado por $\Delta \zeta$ ($i_{\Delta A}$). Isso é feito considerando-se que, em cada situação da superposição, o nó de interesse de (3.8.10) se localiza em ζ e $\zeta + \Delta \zeta$, respectivamente, ou

$$\Delta I_d = \left(\frac{g_2}{g_1 + g_2} \Big|_{\zeta + \Delta \zeta} - \frac{g_2}{g_1 + g_2} \Big|_{\zeta} \right) i_{\Delta A}. \quad (3.8.11)$$

A condutância de cada um dos segmentos do canal, vistos pelo nó intermediário, é dada por

$$g_{1(2)} = \frac{dI_{1(2)}}{dV_x} = I_S \frac{di_{1(2)}}{dV_x} = g_0 \frac{di_{1(2)}}{dv_x}, \quad (3.8.12a)$$

onde $I_{1(2)}$ é a corrente cc que percorre o canal do segmento $M_{1(2)}$, $i_{1(2)}$ é esta mesma corrente normalizada, V_x é a tensão no nó intermediário de posição x , v_x é esta mesma tensão normalizada e $g_0 = I_S / \phi_t$ é a condutância de normalização.

Como a relação entre a corrente e a carga de inversão é determinada por (3.8.9b), e a relação da a carga de inversão com o potencial no canal é dada por (3.8.6) (modelo UCCM), a expressão (3.8.12a) pode ser reescrita como

$$g_{1(2)} = g_0 \frac{di_{1(2)}}{dq'_I} \frac{dq'_I}{dV_x / \phi_t} = -g_0 \frac{di_{1(2)}}{dq'_I} \frac{q'_I}{1 + q'_I}, \quad (3.8.12b)$$

onde q'_I representa a densidade superficial de carga de inversão normalizada, na posição x .

Se (3.8.8b) for integrada da *fonte* até a posição normalizada ζ , resulta na expressão a seguir, que é uma adequação de (3.8.9b) para o transistor M_1

$$i_1 = \frac{(q'_{IS} + 1)^2 - (q'_I + 1)^2}{\xi + \zeta(q'_{IS} - q'_I)}, \quad (3.8.13)$$

e cuja derivada é

$$\frac{di_1}{dq'_I} = \frac{-2(q'_I + 1) + \zeta i_1}{\xi + \zeta(q'_{IS} - q'_I)}. \quad (3.8.14)$$

O mesmo pode ser feito em relação a M_2 , integrando-se (3.8.8b) da posição normalizada ξ até o *dreno*, o que resulta na expressão a seguir

$$i_2 = -\frac{(q'_I + 1)^2 - (q'_{ID} + 1)^2}{(1 - \xi) + \zeta(q'_I - q'_{ID})}, \quad (3.8.15)$$

e cuja derivada é

$$\frac{di_2}{dq'_I} = \frac{-2(q'_I + 1) - \zeta i_2}{(1 - \xi) + \zeta(q'_I - q'_{ID})}. \quad (3.8.16)$$

Utilizando-se a relação obtida em (3.8.10), junto com (3.8.14) e (3.8.16), e observando que $i_1 = -i_2$ (forma normalizada de $I_1 = -I_2$), pode-se obter a razão do divisor de corrente do modelo da fig. 3.17b

$$\frac{g_2}{g_1 + g_2} = \frac{\xi + \zeta(q'_{IS} - q'_I)}{1 + \zeta(q'_{IS} - q'_{ID})}. \quad (3.8.17)$$

Este resultado é uma generalização do princípio de divisão de correntes, baseada na suposição que a mobilidade dos portadores depende do campo elétrico longitudinal conforme dado por (3.8.2). Em um transistor de canal longo ($\zeta \rightarrow 0$), a equação (3.8.17) é simplificada, estabelecendo uma proporcionalidade entre o efeito da flutuação local na corrente de *dreno*, e a posição em que essa flutuação ocorre ao longo do canal, conforme já observado anteriormente.

Aplicando-se (3.8.17) a (3.8.11), conforme o modelo da fig. 3.16b, chega-se à expressão que representa a relação entre a flutuação da corrente de *dreno* provocada por uma flutuação local, no segmento infinitesimal do canal $\Delta\xi$, com a carga de inversão, incluindo o efeito da saturação na velocidade dos portadores, ou

$$\Delta I_d = \frac{\Delta \xi - \zeta \Delta q'_I}{1 + \zeta (q'_{IS} - q'_{ID})} i_{AA} = \Delta \xi \frac{1 - \zeta \frac{dq'_I}{d\xi}}{1 + \zeta (q'_{IS} - q'_{ID})} i_{AA}. \quad (3.8.18)$$

3.8.3 Inclusão no modelo de descasamento

Conforme visto na expressão (3.6.2), a variância da corrente de *dreno* é dada pelo somatório quadrático de todas as parcelas de desvio resultantes das flutuações locais, ao longo do canal. Aplicando-se (3.8.18) a (3.6.2), tem-se

$$\sigma_{i_D}^2 = \overline{\Delta I_D^2} = \sum_{channel-length} (\Delta I_d)^2 = \sum \left(\frac{\Delta \xi - \zeta \Delta q'_I}{1 + \zeta (q'_{IS} - q'_{ID})} i_{AA} \right)^2. \quad (3.8.19)$$

Com o uso de (3.6.3), estabelece-se a relação entre a perturbação local na corrente e a flutuação de densidade de carga na região infinitesimal, fazendo com que a expressão anterior possa ser reescrita, sob a forma normalizada, como

$$\sigma_{i_D}^2 = \sum \left(\frac{\Delta \xi - \zeta \Delta q'_I}{1 + \zeta (q'_{IS} - q'_{ID})} \frac{\Delta q'_I}{q'_I} i_D \right)^2, \quad (3.8.20)$$

onde i_D é a corrente de *dreno* normalizada.

O uso de (3.6.6), (3.6.9) e (3.6.11) permite que se inclua na expressão anterior, a dependência da flutuação da densidade de carga de inversão com a flutuação na concentração local de impurezas, aqui representada pelo parâmetro N_{oi} , ou

$$\sigma_{i_D}^2 = \left(\frac{q}{Q'_{IP}} \right)^2 \frac{N_{oi}}{WL} i_D \sum \left[\left(\frac{\Delta \xi - \zeta \Delta q'_I}{1 + \zeta (q'_{IS} - q'_{ID})} \frac{1}{q'_I + 1} \right)^2 \frac{i_D}{\Delta \xi} \right], \quad (3.8.21)$$

que pode ser reescrita como a seguinte integral, quando consideramos o segmento do canal $\Delta \xi$ tendendo a zero e utilizamos a relação (3.8.8b)

$$\sigma_{i_D}^2 = -2 \frac{i_D}{WL} \frac{N_{oi}}{N^{*2}} \frac{1}{[1 + \zeta (q'_{IS} - q'_{ID})]^2} \int_{q'_{IS}}^{q'_{ID}} \left(1 - \zeta \frac{dq'_I}{d\xi} \right) \frac{dq'_I}{q'_I + 1}. \quad (3.8.22)$$

Normalizando-se essa expressão, e com o uso de (3.8.8b), chega-se a

$$\frac{\sigma_{i_D}^2}{i_D^2} = -2 \frac{1}{WL} \frac{N_{oi}}{N^{*2}} \frac{1}{i_D} \frac{1}{[1 + \zeta (q'_{IS} - q'_{ID})]^2} \int_{q'_{IS}}^{q'_{ID}} \left(1 - \frac{-i_D \zeta / 2}{q'_I + 1 - i_D \zeta / 2} \right) \frac{dq'_I}{q'_I + 1}, \quad (3.8.23)$$

que, se integrada entre *dreno* e *fonte*, e substituindo-se as densidades de carga de inversão pelos níveis de inversão, e com o uso de (3.8.9c), resulta

$$\frac{\sigma_{i_D}^2}{i_D^2} = \frac{2}{WL} \frac{N_{oi}}{N^{*2}} \frac{1}{i_f - i_r} \frac{1}{1 + \zeta (\sqrt{1+i_f} - \sqrt{1+i_r})} \ln \left(\frac{\sqrt{1+i_f} - \frac{\zeta}{2} i_D}{\sqrt{1+i_r} - \frac{\zeta}{2} i_D} \right), \quad (3.8.24a)$$

onde a corrente de dreno normalizada é dada por

$$i_D = \frac{i_f - i_r}{1 + \zeta (\sqrt{1+i_f} - \sqrt{1+i_r})}. \quad (3.8.24b)$$

Observando-se essa expressão, pode-se verificar que ela corresponde à mesma expressão para o descasamento desenvolvida no item 3.6, (3.6.14), mas acrescida de termos relacionados ao efeito da saturação na velocidade dos portadores da corrente média de *dreno*. Em um transistor de canal longo ($\zeta \rightarrow 0$), esta expressão é simplificada para o mesmo resultado obtido naquele item.

No cálculo de (3.8.24) deve-se também observar que existe uma densidade mínima da carga de inversão de *dreno* (q'_{IDsat}), que ocorre quando os portadores atingem a velocidade de saturação [94], e que é dada por

$$q'_{IDsat} = q'_{IS} + \frac{\zeta + 1}{\zeta} \left[1 - \sqrt{1 + 2 \frac{\zeta}{(\zeta + 1)^2} q'_{IS}} \right], \quad (3.8.25a)$$

resultando em um nível mínimo de inversão no *dreno*, dado por

$$i_{r\min} = (q'_{IDsat} + 1)^2 - 1. \quad (3.8.25b)$$

Assim, para transistores de canal curto, onde se leva em conta o efeito da saturação na velocidade dos portadores no modelo de descasamento através de (3.8.24), o cálculo do nível de inversão no *dreno* (i_r) deve ser feito observando-se que este não pode ser inferior ao valor mínimo dado por (3.8.25).

A fig. 3.18 apresenta o efeito da saturação na velocidade dos portadores (ESVP), sobre o cálculo do descasamento, com parâmetros da tecnologia *TSMC 0.35* e $v_{sat} = 10^7$

cm/s [67]. O ESVP é obtido através da razão entre a diferença das expressões do descasamento para transistores de canal longo (CL – eq. 3.6.14) e canal curto (CC – eq. 3.8.24), com a expressão para transistores de canal curto, ou

$$ESVP = \frac{\frac{\sigma_{I_D}(CL) - \sigma_{I_D}(CC)}{I_D}}{\frac{\sigma_{I_D}(CC)}{I_D}}. \quad (3.8.26)$$

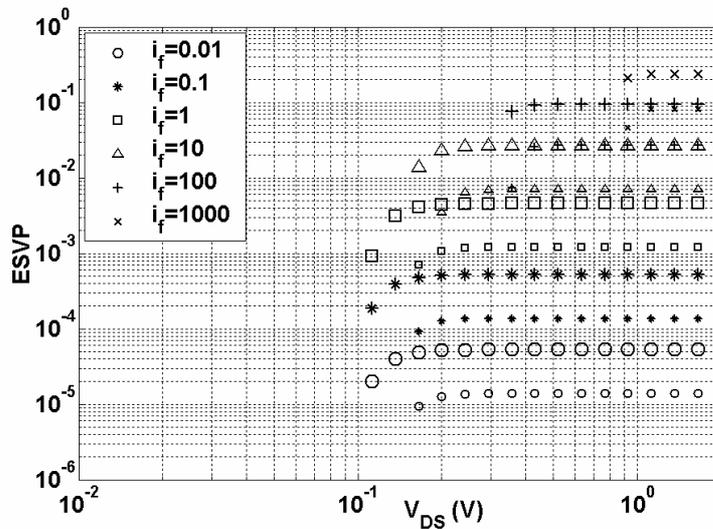


Figura 3.18: Efeito da saturação na velocidade dos portadores (ESVP) sobre o descasamento, em função do nível de inversão (i_f) e da condição de saturação (V_{DS} variando entre 20 mV e 2 V), para um transistor NMOS na tecnologia TSMC 0.35. Os símbolos maiores representam um transistor com $L = 2\lambda$ (canal mínimo) e os menores, um transistor com $L = 8\lambda$.

Observa-se nesta figura que em um transistor de canal mínimo sob saturação, o erro resultante de se negligenciar o efeito da saturação na velocidade dos portadores no cálculo do descasamento é da ordem de 2,5% para $i_f = 10$, 10% para $i_f = 100$ e 23% para $i_f = 1000$. Para um transistor de canal médio (4 vezes o mínimo), esse erro resulta em 0,7% para $i_f = 10$, 2,5% para $i_f = 100$ e 9% para $i_f = 1000$, também sob saturação. Fora da condição de saturação, esse erro é desprezível.

3.9 Depleção no substrato

O modelo desenvolvido no item 3.6 pressupõe o número de impurezas ionizadas, existente dentro da região de depleção, como sendo a principal causa do descasamento entre transistores *MOS* identicamente desenhados e fabricados. Para tanto, considerou-se a região de depleção, limitada pelas dimensões W , L e y_d , como um volume constante e independente da condição de polarização. Sabe-se que, de fato, a profundidade da região de depleção sob a área ativa do transistor depende tanto da polarização *porta-substrato* (V_{GB}), quanto da polarização *dreno-fonte* (V_{DS}) e *fonte-substrato* (V_{SB}). Esse efeito de modulação do volume da região de depleção pode ser entendido como uma dependência do fator de descasamento N_{oi} com a condição de polarização, de modo que pode ser incluído através de um termo de correção no modelo.

3.9.1 Efeito da polarização *porta-substrato*

O aumento da polarização entre os terminais de *porta* e *substrato* provoca um aumento na profundidade da região de depleção, pois são necessárias mais cargas ionizadas para gerar o campo elétrico que compense o resultante do aumento de potencial aplicado.

A equação de potenciais do capacitor *MOS*, pode ser escrita como

$$V_{GB} = \phi_S + \phi_{ox} + V_{FB} = \phi_S - \frac{Q'_I + Q'_B}{C'_{ox}} + V_{FB}, \quad (3.9.1)$$

onde ϕ_S representa a queda de potencial entre a região não depletada do substrato e a interface substrato-óxido (potencial de superfície), ϕ_{ox} representa a queda de potencial através da espessura do óxido, V_{FB} é o potencial de banda-plana, que resulta dos potenciais de contato entre o terminal de *porta* e o de substrato, acrescido dos efeitos das cargas aprisionadas no óxido, Q'_I é a densidade superficial de carga de inversão, Q'_B é a densidade de carga de depleção e C'_{ox} é a capacitância do óxido por unidade de área.

Considerando-se a concentração líquida de dopantes (N_A) como sendo uniforme no volume do substrato, na região do canal, pode-se expressar a relação entre ϕ_S e Q'_B como

$$\phi_S = \frac{Q'^2_B}{2q\epsilon_S N_A}. \quad (3.9.2)$$

Deseja-se obter a relação entre uma variação aplicada ao potencial V_{GB} e seu efeito sobre a carga de depleção. Assim, pode-se substituir (3.9.2) em (3.9.1) e derivar-se esta expressão, resultando

$$dV_{GB} = \left(\frac{Q'_B}{q\varepsilon_S N_A} - \frac{1}{C'_{ox}} \right) dQ'_B - \frac{dQ'_I}{C'_{ox}}. \quad (3.9.3a)$$

Considerando-se o modelo UCCM na forma diferencial de (3.8.6), aplicado à condição de *pinch-off* ($V_x = V_P$), e lembrando que $dV_{GB} \approx ndV_P$ [94], pode-se reescrever (3.9.3a) como

$$\frac{dV_{GB}}{1 + \frac{Q'_I}{Q'_{IP}}} = \left(\frac{Q'_B}{q\varepsilon_S N_A} - \frac{1}{C'_{ox}} \right) dQ'_B. \quad (3.9.3b)$$

Através de (3.9.3b) pode-se avaliar o efeito das variações do potencial *porta-substrato*, sobre a carga de depleção, para as condições de inversão forte e fraca.

Sob inversão forte ($|Q'_I| \gg |Q'_{IP}|$), (3.9.3b) indica que as variações em V_{GB} resultam em variações tanto menores em Q'_B , quanto mais intensa for a condição de inversão. Isso pode ser compreendido lembrando-se que nessa condição a “folha” onde se concentra a carga de inversão, na interface óxido-substrato, forma uma espécie de “blindagem” condutiva, reduzindo os efeitos das variações de campo elétrico que ocorram no óxido, devido a variações no potencial da *porta*, sobre o substrato. Assim, sob inversão forte, o potencial V_{GB} afeta pouco o parâmetro de descasamento N_{oi} , que está relacionado às flutuações espaciais dos dopantes que formam a carga de depleção.

Sob inversão fraca ($|Q'_I| \ll |Q'_{IP}|$), (3.9.3b) pode ser simplificada, resultando

$$dV_{GB} = \left(\frac{Q'_B}{q\varepsilon_S N_A} - \frac{1}{C'_{ox}} \right) dQ'_B. \quad (3.9.3c)$$

Considerando-se uma concentração de átomos dopantes homoganeamente distribuída no volume da região de depleção, pode-se estimar a densidade de carga de depleção por

$$Q'_B = -qN_A y_d, \quad (3.9.4)$$

onde N_A é a concentração líquida de dopantes e y_d é a profundidade da região de depleção, a partir da interface óxido-semicondutor.

Sob a mesma condição, de concentração homogênea de dopantes, pode-se reduzir (3.6.11) a

$$N_{oi} = N_a y_d / 3, \quad (3.9.5)$$

onde N_a é a concentração total de dopantes, incluindo doadores e aceitadores.

Considerando-se que no volume da região ativa, para que o substrato tenha característica predominante N ou P, a concentração líquida de dopantes tem valor próximo da total ($N_A \approx N_a$), (3.9.4) e (3.9.5) resultam em

$$Q'_B \approx -3qN_{oi}. \quad (3.9.6)$$

Através de (3.9.6), pode-se reescrever (3.9.3c) como

$$\frac{dN_{oi}}{dV_{GB}} = \frac{1}{3q \left(\frac{3N_{oi}}{\epsilon_S N_A} + \frac{1}{C'_{ox}} \right)}. \quad (3.9.7)$$

Assim, sob inversão fraca e considerando-se os parâmetros relativos à tecnologia TSMC 0.35 obtidos através da MOSIS ($N_A = 2,2 \times 10^{17} \text{ cm}^{-3}$, $N_D = 8,5 \times 10^{16} \text{ cm}^{-3}$, $C'_{ox} = 445 \text{ nF/cm}^2$) e os extraídos experimentalmente ($N_{oiN} = 1,9 \times 10^{12} \text{ cm}^{-2}$, $N_{oiP} = 6,6 \times 10^{12} \text{ cm}^{-2}$), estima-se que a dependência percentual de N_{oi} em V_{GB} , sob as condições de polarização aqui descritas, seja de aproximadamente 4%/V para os transistores NMOS e 0,15%/V para os PMOS. Essa dependência tende a diminuir na medida em que o dispositivo passa a operar em níveis de inversão mais altos.

3.9.2 Efeito da polarização *dreno-fonte*

A existência de polarização não nula entre os terminais de *dreno* e *fonte* faz com que ocorra uma variação no potencial ao longo do canal de um transistor MOS que esteja em condução. Esta variação de potencial implica numa variação na densidade superficial de carga de inversão (Q'_I) ao longo do canal, o que resulta também numa variação (contrária) na densidade de carga de depleção (Q'_B) ao longo do canal. Tal variação em Q'_B , representa uma deformação na profundidade da região de depleção (y_d), tornando-a

tanto maior quanto mais próximo do terminal que apresenta a menor diferença de potencial com a *porta*.

O modelo desenvolvido no item 3.6 pressupõe uma profundidade constante da região de depleção ao longo do canal. Aqui o impacto da modulação da profundidade da região de depleção, ao longo do canal, é analisado.

Considerando-se um transistor *MOS* com potenciais fixos aplicados a seus terminais, pode-se extrair a seguinte relação a partir de (3.9.1)

$$\frac{dQ'_B}{d\phi_S} + \frac{dQ'_I}{d\phi_S} = C'_{ox}, \quad (3.9.8)$$

que, com o uso da aproximação (3.2.1), pode ser reescrita como

$$dQ'_B = -\frac{n-1}{n}dQ'_I. \quad (3.9.9)$$

Pode-se representar a densidade de carga de inversão na coordenada x , como sendo seu valor na coordenada $x=0$ (Q'_{IS}) acrescido da integração de dQ'_I entre a *fonte* ($x=0$) e a coordenada x , ou

$$Q'_I(x) = Q'_{IS} + \int_S^x dQ'_I. \quad (3.9.10)$$

Uma relação análoga pode ser desenvolvida para a densidade de carga de depleção na coordenada x

$$Q'_B(x) = Q'_{BS} + \int_S^x dQ'_B, \quad (3.9.11)$$

que, através de (3.9.9), pode ser reescrita como

$$Q'_B(x) = Q'_{BS} - \frac{n-1}{n} \int_S^x dQ'_I, \quad (3.9.12)$$

e que, com o auxílio de (3.9.10), pode ser representada por

$$Q'_B(x) = Q'_{BS} + \frac{n-1}{n} (Q'_{IS} - Q'_I(x)). \quad (3.9.13)$$

Com o uso de (3.6.11) e (3.9.6), pode-se reescrever (3.6.10) como

$$\sigma_{I_D}^2 = \frac{q \mu I_D}{3n C'_{ox} L^2} \int_{Q'_{IS}}^{Q'_{ID}} \frac{-Q'_B}{n C'_{ox} \phi_t - Q'_I} dQ'_I, \quad (3.9.14)$$

que, com o uso de (3.9.13), e depois de integrada entre *fonte* e *dreno*, resulta em

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{q\mu}{3L^2 n C'_{ox} I_D} \left\{ \frac{n-1}{n} (Q'_{IS} - Q'_{ID}) - \left[\frac{n-1}{n} (Q'_{IS} - n C'_{ox} \phi_t) + Q'_{BS} \right] \ln \left(\frac{n C'_{ox} \phi_t - Q'_{IS}}{n C'_{ox} \phi_t - Q'_{ID}} \right) \right\}. \quad (3.9.15)$$

Através da aproximação (3.9.6), pode-se obter uma relação entre Q'_{BS} e N_{oi} , que fica referenciado à *fonte*. Utilizando-se então a relação entre densidade de carga de inversão e nível de inversão, pode-se reescrever (3.9.15) como a soma de duas parcelas. A primeira, é a própria expressão (3.6.14), aqui copiada por conveniência

$$\frac{\sigma_{I_D}^2}{I_D^2} (1) = \frac{N_{oi}}{WLN^{*2}} \frac{1}{i_f - i_r} \ln \left(\frac{1+i_f}{1+i_r} \right). \quad (3.9.16)$$

A segunda parcela representa o efeito da modulação da região de depleção sobre o descasamento, sendo

$$\frac{\sigma_{I_D}^2}{I_D^2} (2) = \frac{n-1}{n} \frac{1}{3WLN^*} \frac{1}{i_f - i_r} \left[\sqrt{1+i_f} \ln \left(\frac{1+i_f}{1+i_r} \right) - 2(\sqrt{1+i_f} - \sqrt{1+i_r}) \right]. \quad (3.9.17)$$

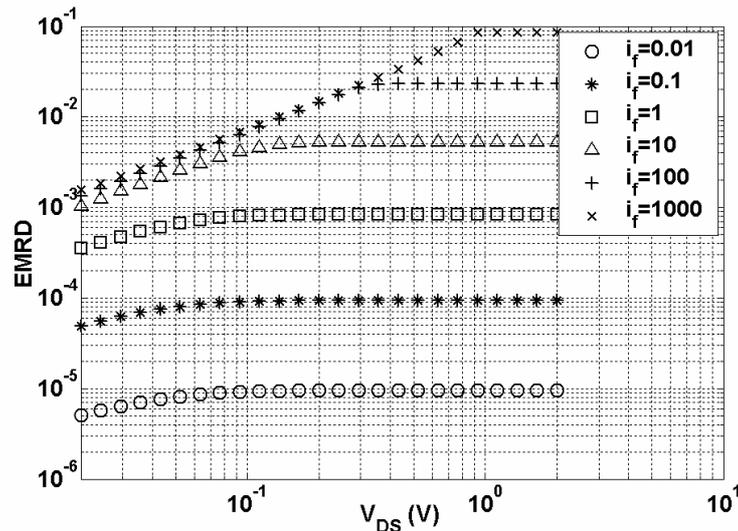


Figura 3.19: Efeito da modulação da região de depleção (EMRD), em função do nível de inversão (i_f) e da condição de saturação (V_{DS} variando entre 20 mV e 2 V), de um transistor NMOS. EMRD resulta da razão entre a eq. (3.9.17) e a eq. (3.9.16). Nesta curva utilizou-se parâmetros extraídos da tecnologia TSMC 0.35.

A fig. 3.19 apresenta a razão entre o módulo de (3.9.17) e (3.9.16), aqui denominada EMRD (efeito da modulação da região de depleção), cobrindo desde níveis de inversão muito fracos até muito fortes, e desde a região linear até a saturação. Observa-se que o maior impacto deste efeito no descasamento de um *MOSFET* ocorre apenas sob inversão muito forte e saturação, que é a condição onde ocorre a maior deformação na profundidade da região de depleção, entre *fonte* e *dreno*. Utilizou-se parâmetros extraídos da tecnologia *TSMC 0.35* para transistores NMOS ($N_{oi} = 1,9 \times 10^{12} \text{ cm}^{-2}$, $n \approx 1,3$, $C'_{ox} = 444 \text{ nF/cm}^2$). Para os transistores PMOS, o EMRD é menos de 1/3 do observado na figura, visto que o N_{oi} extraído é mais de 3 vezes o do NMOS.

3.10 Depleção na porta

A tecnologia empregada na fabricação de transistores *MOS* nos últimos 30 anos, tem empregado silício dopado, sob a forma policristalina (ou também amorfa), na implementação da região de *porta*. Isso decorre da vantagem do auto-alinhamento deste sobre a região ativa.

Mas, o silício dopado não pode ser considerado equipotencial como um metal. Em função do campo elétrico presente entre *porta* e substrato, surge na interface *porta*-óxido uma região depletada, que apresenta queda de potencial [36], [37].

Essa depleção pode ser analisada sob dois aspectos: seu efeito médio sobre as cargas de inversão e depleção ao longo da região ativa; e o efeito das flutuações que ocorrem nessa depleção, em função da natureza discreta dos íons implantados na região da *porta*. Por simplicidade, o impacto de cada um destes efeitos sobre o descasamento entre transistores idênticos é tratado separadamente neste item.

3.10.1 Efeito médio na depleção de *porta*

Considerando-se que a dopagem do poli-silício de *porta* está homogênea e continuamente distribuída, e aplicando-se a lei de Gauss a um pequeno segmento longitudinal do transistor, pode-se reescrever a equação de potenciais do capacitor *MOS* como

$$V_{GB} = \phi_S + \phi_{ox} + \phi_P + V_{FB} = \phi_S - \frac{Q'_C}{C'_{ox}} + \phi_P + V_{FB}, \quad (3.10.1)$$

onde ϕ_S representa a queda de potencial entre a região não depletada do substrato e a interface substrato-óxido (potencial de superfície), ϕ_{ox} representa a queda de potencial através da espessura do óxido, ϕ_P representa a queda de potencial entre a região não depletada da *porta* e a interface *porta*-óxido, V_{FB} é o potencial de banda plana, que resulta da diferença entre os potenciais de contato do terminal de *porta* e do *substrato*, acrescido dos efeitos das cargas aprisionadas no óxido, Q'_C corresponde à densidade superficial de carga total abaixo da interface óxido-substrato, resultando da soma de Q'_I (a densidade superficial de carga de inversão) com Q'_B (densidade de carga de depleção), e C'_{ox} é a capacitância do óxido por unidade de área.

Diferenciando-se (3.10.1) em relação a Q'_C , chega-se ao modelo do capacitor *MOS*, adaptado à situação de depleção de *porta*, como apresentado na fig. 3.20, e que pode ser resumido como

$$\frac{1}{C'_{gb}} = \frac{1}{C'_p} + \frac{1}{C'_{ox}} + \frac{1}{C'_i + C'_b} = \frac{1}{C'_{oxp}} + \frac{1}{C'_i + C'_b}, \quad (3.10.2)$$

onde C'_{oxp} representa a associação equivalente das capacitâncias de óxido e da depleção do poli-silício.

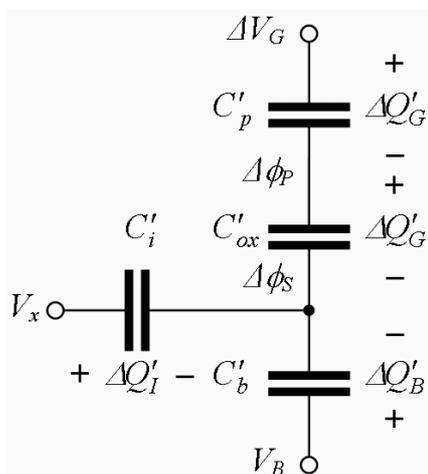


Figura 3.20: Modelo do capacitor *MOS*, alterado para incorporar o efeito de depleção que ocorre no poli-silício da *porta*.

Neste modelo, a flutuação na carga de *porta* é dada por $\Delta Q'_G = C'_{ox}(\Delta \phi_P - \Delta \phi_S) = C'_p(\Delta V_G - \Delta \phi_P) = C'_{oxp}(\Delta V_G - \Delta \phi_S) = C'_{gb}\Delta V_G$, $\Delta Q'_I = -C'_i \Delta \phi_S$ é a flutuação de carga de inversão e $\Delta Q'_B = -C'_b \Delta \phi_S$ é a flutuação de carga de depleção.

V_B e V_x são respectivamente os potenciais de substrato e do canal, que são considerados constantes, e ΔV_G é a variação aplicada ao terminal de *porta*.

Da relação entre a carga uniformemente distribuída numa região de depleção e a queda de potencial através dessa região, resulta que

$$Q'_B = -\sqrt{2q\varepsilon_S N_A} \sqrt{\phi_S} = -C'_{ox} \gamma \sqrt{\phi_S}, \quad (3.10.3)$$

$$Q'_G = \sqrt{2q\varepsilon_S N_P} \sqrt{\phi_P} = C'_{ox} \gamma_p \sqrt{\phi_P}, \quad (3.10.4)$$

onde N_A e N_P são as concentrações líquidas médias de dopantes no substrato e no polissilício da *porta*, e γ e γ_p são definidos como os *coeficientes de efeito de corpo* do substrato e da *porta*, respectivamente.

Através de (3.10.4), pode-se então estabelecer que

$$C'_p = \frac{\sqrt{2q\varepsilon_S N_P}}{2\sqrt{\phi_P}} = \frac{q\varepsilon_S N_P}{Q'_P}. \quad (3.10.5)$$

Como o modelo da figura 3.20 deve respeitar a conservação de cargas, ou seja, $\Delta Q'_C + \Delta Q'_G = \Delta Q'_I + \Delta Q'_B + \Delta Q'_G = 0$, pode-se estabelecer através de (3.10.3) e (3.10.4) que

$$\frac{\phi_P}{\phi_S} = \left(1 + \frac{Q'_I}{Q'_B}\right)^2 \frac{N_A}{N_P} = \left(\frac{Q'_G}{Q'_B}\right)^2 \frac{N_A}{N_P}. \quad (3.10.6)$$

Através de (3.10.1) e (3.10.3), pode-se explicitar a carga de inversão como

$$Q'_I = -C'_{ox} (V_G - V_{FB} - \phi_S - \gamma \sqrt{\phi_S} - \phi_P). \quad (3.10.7)$$

A partir de (3.10.6) e (3.10.7), pode-se estabelecer o *fator de rampa* (n_p) e o incremento na tensão de limiar (ΔV_{T0}), devido à depleção na *porta*, resultando

$$n_p = 1 + \frac{\gamma}{2\sqrt{\phi_{Sa}}} + \frac{N_A}{N_P}, \quad (3.10.8)$$

$$\Delta V_{T0} = 2\phi_F \frac{N_A}{N_P}, \quad (3.10.9)$$

onde ϕ_{Sa} representa o potencial de superfície sob condição de inversão fraca e ϕ_F representa o nível de Fermi no substrato. Deve-se lembrar que estes dois parâmetros são estimados sob inversão fraca ($Q'_I \approx 0$).

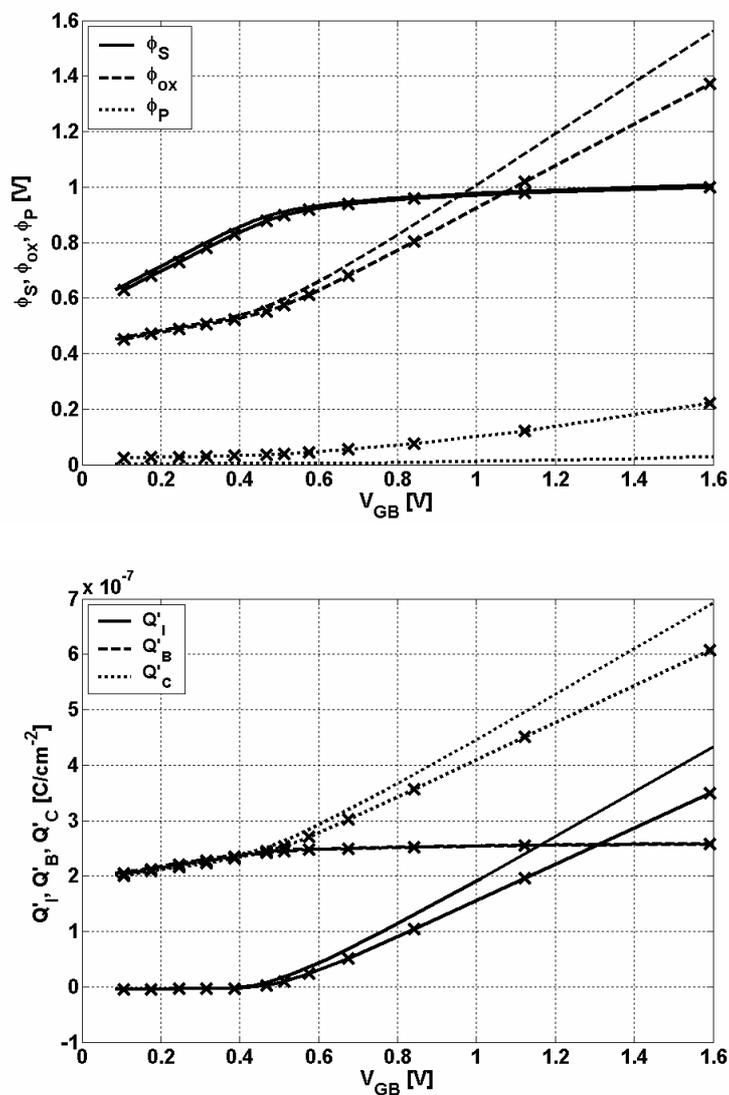


Figura 3.21: Simulação de (3.10.1), para: $V_{FB} = -1$ V, $C'_{ox} = 445$ nF/cm², $N_a = 2 \times 10^{17}$ cm⁻³, para dois valores de N_p : 5×10^{18} cm⁻³ (marca x) e 5×10^{19} cm⁻³ (sem marca).

Pode-se então afirmar que o efeito da depleção no poli-silício da *porta* pode ser incluído no modelo de um *MOSFET* através de correções no modelo de capacitância de *porta*, no fator de rampa e na tensão de limiar. Em uma abordagem mais simplista, considera-se que esse efeito pode ser incorporado à capacitância de óxido, através de um aumento equivalente de espessura do óxido (t_{ox}), que é calculado de modo a causar um efeito semelhante à depleção do poli-silício no capacitor *MOS*.

Supondo que dois transistores idênticos estejam sob as mesmas condições de polarização e que o nível médio de concentração de dopantes no poli-silício de *porta* seja semelhante para os dois, teremos ambos operando sob os mesmos níveis de inversão de *dreno* e *fonte*, resultando que a depleção média que ocorre na *porta* não deve afetar significativamente o descasamento entre eles. Assim, o efeito da depleção média sobre o descasamento se reduz ao seu efeito sobre os parâmetros do modelo de descasamento que dependam de n ou de C'_{ox} , ou seja, N^* e B_{ISQ} .

A figura 3.21 mostra o efeito da depleção no poli-silício de *porta* sobre os potenciais de superfície e de óxido de um capacitor *MOS*, e suas densidades superficiais de carga correspondentes, para duas concentrações diferentes de dopantes no poli-silício.

3.10.2 Flutuações na depleção de *porta*

Microscopicamente, a densidade de carga de depleção superficial do poli-silício de *porta* pode variar ao longo da área desta, em função das flutuações que ocorrem na concentração dos dopantes implantados no material [36], [37]. Tais flutuações provêm de duas causas: em função da natureza discreta dos átomos implantados (como ocorre no substrato), e em função do processo de aglomeração destes átomos, que ocorre devido ao silício que forma este terminal não ser um mono-cristal, mas sim composto por pequenos fragmentos cristalinos (poli-cristal). Assume-se aqui que estes dois fatores são aleatórios e com média zero sobre toda a área da região de *porta*.

Como no item 3.6, pode-se analisar o efeito das flutuações aleatórias que ocorrem na região depletada da *porta*, dividindo-se a região ativa do transistor como na fig. 3.2, e aplicando-se o princípio de conservação de cargas ao segmento considerado, através do modelo da fig. 3.20, de modo que

$$\Delta Q'_B(x) + \Delta Q'_G(x) + \Delta Q'_I(x) = 0, \quad (3.10.10)$$

onde $\Delta Q'_B = -C'_b \Delta \phi_S$ é a flutuação na carga de depleção de *substrato*, e $\Delta Q'_I = -C'_i \Delta \phi_S$ é a flutuação na carga de inversão. A variação na carga de depleção de *porta*, $\Delta Q'_G$, resulta da soma de duas componentes, sendo a primeira igual a $-C'_{oxp} \Delta \phi_S$ e associada à flutuação no potencial de superfície, e a segunda designada $\Delta Q'_{IMPP}$ e associada com a flutuação no número de impurezas ionizadas na região de interesse do poli-silício. Assim, a variação na

carga de *porta* é dada por $\Delta Q'_G = -C'_{oxp} \Delta \phi_S + \Delta Q'_{IMPP}$, onde C'_{oxp} resulta da associação equivalente entre C'_{ox} e C'_p .

Observando-se (3.10.10), pode-se concluir que uma flutuação no número de impurezas presentes na região depletada do poli-silício tem efeito equivalente a uma flutuação semelhante nas impurezas do substrato, efeito já analisado no item 3.6. Pode-se assim deduzir que o desenvolvimento teórico feito neste item, e conseqüentemente os resultados obtidos, também se aplicam ao caso das flutuações aleatórias de concentração de dopantes no poli-silício de *porta*, desde que se utilize neste desenvolvimento os parâmetros C'_{oxp} e n_p .

Assim, pode-se concluir que a inclusão no modelo de descasamento, das flutuações no número de átomos dopantes do poli-silício, resulta, em uma primeira aproximação, no aumento do parâmetro N_{oi} .

Ainda, pode-se incluir em N_{oi} todas as causas de descasamento que tenham mecanismos semelhantes, produzindo flutuações de carga em (3.10.10), que sejam espacialmente aleatórias, mas estáticas no tempo. Assim, N_{oi} pode também englobar os efeitos das *cargas fixas no óxido*.

Caso se confirme a atual tendência de se resgatar o uso de portas metálicas nas tecnologias *MOS* mais avançadas, os efeitos de depleção no poli-silício deverão desaparecer.

3.11 Efeitos de rugosidade de borda

Alguns dos primeiros estudos sobre o descasamento entre dispositivos *MOS* identicamente fabricados [6], [7], supunham que um dos fatores causadores deste seria a rugosidade existente nas bordas da região de *porta* (*LER* ou *line edge roughness*). Essa rugosidade pode ser atribuída ao processo de foto-litografia e ao processo de ataque químico (*etching*), que ocorrem durante a definição da região onde o poli-silício de *porta* será depositado [56]. Outros estudos da época demonstraram que este fator apresentava mínimo impacto no descasamento, sendo a flutuação na concentração de dopantes no volume da região ativa o principal fator [8], [21], [25], [46], [88], de modo que as pesquisas sobre o *LER* não avançaram.

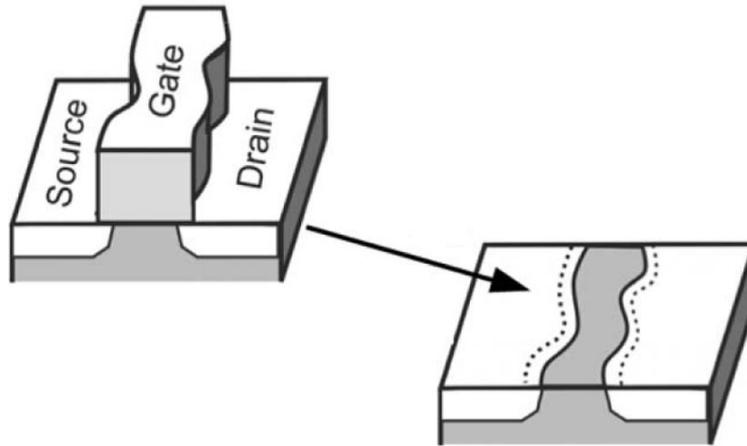


Figura 3.22: Efeito da rugosidade das bordas do poli-silício da *porta*, sobre a definição da região do canal (extraído de [58]).

A tecnologia atual utiliza o processo de auto-alinhamento para a definição da região do canal dos transistores *MOS*. Neste processo, a próprio poli-silício depositado para a formação do eletrodo de *porta* serve como máscara durante o implante das regiões de *dreno* e *fonte*, evitando que a região do canal seja dopada. Assim, na definição das bordas que definem o comprimento do canal, as rugosidades periféricas do poli-silício acabam por ser replicadas na definição das bordas das regiões implantadas de *dreno* e *fonte*, e que limitam o canal (fig. 3.22).

As etapas de tratamento térmico posteriores tendem a difundir levemente as bordas das regiões implantadas, reduzindo as variações de dimensão do canal, mas não as eliminando por completo.

Com a crescente redução no tamanho dos dispositivos, há alguns anos o *LER* voltou a ser motivo de preocupação dos pesquisadores. Acredita-se que este efeito passe a ser significativo para tecnologias com comprimento de canal da ordem de 50 nm ou menor [56], [57], [58], [59], [60].

Várias estratégias têm sido utilizadas na modelagem e simulação deste efeito, sendo a maioria através de métodos computacionais [59], ou de análise estatística [57].

A caracterização deste efeito tem sido feita através da extração experimental do desvio-padrão da variação dimensional da borda (σ_{LER}) e da faixa de distâncias de correlação (ou “comprimentos de onda” – aplica-se a transformada de Fourier sobre a curva amplitude versus distância) em que ocorrem. Esse efeito também pode ser quantizado através da variação da dimensão do canal, no sentido de interesse (*LWR* ou *line*

width roughness), sendo que $\sigma_{LWR}^2 = 2\sigma_{LER}^2$, pois não há correlação entre as variações nas duas bordas “paralelas” de uma área.

A propagação do efeito de uma grandeza estatística, pode ser estimada por

$$\sigma_P^2 = \left| \frac{\partial \Delta P}{\partial f} \right|^2 \sigma_f^2 = 2 \left| \frac{\partial P}{\partial f} \right|^2 \sigma_f^2, \quad (3.11.1)$$

onde P é a variável em que se quer estimar a propagação da grandeza f .

Como a corrente em um transistor *MOS* é inversamente proporcional à L , e considerando-se que o *LER* não é suficiente para afetar microscopicamente o nível de inversão nas regiões de *dreno* e *fonte* de um transistor polarizado, pode-se estabelecer que

$$\frac{\sigma_{I_D}}{I_D} = \sqrt{2} \frac{\sigma_{LWR}}{L}. \quad (3.11.2)$$

O *ITRS* 2006 [95] estabeleceu que o limite para se produzir os circuitos digitais otimizados atuais é quando $3\sigma_{LER}$ está abaixo de 12% do comprimento médio efetivo do canal, o que significa uma dispersão média relativa na corrente de dreno (σ_{I_D}/I_D) de cerca de 8%. Tipicamente, pode-se considerar algo como 5% de variação média na corrente, como sendo uma margem tolerável, mas lembrando que são as exigências impostas pela aplicação que irão determinar este limite de forma mais exata.

Através das publicações voltadas à extração e análise experimental desse efeito [56], [57], [58], [60], observa-se que as tecnologias mais modernas (comprimentos de canal entre 180 e 50 nm) apresentam valores para σ_{LWR} na faixa de 9 a 3 nm, significando que já se está atingindo o limite proposto no *ITRS* 2006.

3.12 Efeitos da corrente específica I_S

Conforme já visto no item 3.2, através do modelo *ACM* para canal longo do *MOSFET*, a corrente de dreno é definida como

$$I_D = I_F - I_R = I_S (i_f - i_r), \quad (3.12.1)$$

sendo a corrente específica dada por

$$I_S = I_{SQ} \frac{W}{L} = \frac{1}{2} \mu C'_{ox} n \phi_t^2 \frac{W}{L}. \quad (3.12.2)$$

Assumindo-se que a temperatura é a mesma para os dispositivos em que se deseja avaliar o descasamento, o desvio relativo da corrente específica pode ser expresso como

$$\frac{\Delta I_S}{I_S} = \frac{\Delta I_{SQ}}{I_{SQ}} + \frac{\Delta W}{W} - \frac{\Delta L}{L} \quad (3.12.3)$$

onde I_{SQ} é denominado como a corrente específica por quadrado (ou de folha).

De forma simplificada, pode-se assumir que I_{SQ} independe da polarização, e que os desvios dos três termos de (3.12.3) não são correlacionados, de modo que

$$\frac{\sigma^2(I_S)}{I_S^2} = \frac{\sigma^2(I_{SQ})}{I_{SQ}^2} + \frac{\sigma^2(W)}{W^2} + \frac{\sigma^2(L)}{L^2}. \quad (3.12.4)$$

Este resultado é semelhante ao apresentado em [6], [7] e [8], de modo que pode-se supor que estas variâncias também seguem as relações

$$\frac{\sigma^2(I_{SQ})}{I_{SQ}^2} = \frac{A_{ISQ}^2}{WL}, \quad (3.12.5)$$

$$\sigma^2(W) = \frac{A_W^2}{L}, \quad (3.12.6)$$

$$\sigma^2(L) = \frac{A_L^2}{W}. \quad (3.12.7)$$

Substituindo-se (3.12.5), (3.12.6) e (3.12.7), a expressão (3.12.4) resulta

$$\frac{\sigma^2(I_S)}{I_S^2} = \frac{A_{ISQ}^2}{WL} + \frac{A_W^2}{LW^2} + \frac{A_L^2}{WL^2}. \quad (3.12.8)$$

Deve-se observar que (3.12.8) é uma aproximação, pois a corrente específica por quadrado (I_{SQ}) depende de grandezas que podem sofrer variações com a polarização, como n e μ .

Para o caso de transistores não muito estreitos ou curtos, tem-se que (3.12.8) fica reduzida à aproximação feita no item 3.6, ou

$$\frac{\sigma^2(I_S)}{I_S^2} = \frac{A_{ISQ}^2}{WL}. \quad (3.12.9)$$

4 CIRCUITO DE TESTE

Neste capítulo são tratados o projeto e a implementação de circuitos integrados voltados à caracterização elétrica do efeito do descasamento entre transistores *MOS*. Duas versões são apresentadas, sendo que ambas foram implementadas em diversas tecnologias disponíveis, de modo a permitir que se obtenha o volume de medidas necessário, e com exatidão suficiente, para que o modelo aqui apresentado possa ser avaliado e utilizado.

4.1 Objetivos

Quanto ao objetivo, pode-se classificar os circuitos (ou estruturas) de teste voltados à caracterização elétrica de uma tecnologia (ou processo) em dois tipos.

No primeiro, encontram-se os grupos de estruturas, implementados repetidamente em algumas quadrículas em posições estratégicas sobre cada lâmina (*wafer*), cujo objetivo é a extração posterior de medidas que possam informar sobre a estabilidade das diversas etapas de fabricação entre regiões de uma mesma lâmina (*interdie*), entre lâminas de um mesmo lote (*interwafer*) ou entre diferentes lotes (*interbatch*). Normalmente, como o objetivo dessas estruturas é apenas de verificação das várias características do processo, prima-se durante sua concepção mais pela variedade do que pela quantidade, de modo a se ter estruturas com diversidade suficiente para se cobrir todas as características ou parâmetros de interesse, repetindo-se cada estrutura algumas vezes dentro da quadrícula, apenas em número suficiente para se ter segurança da correção da medida obtida.

No segundo tipo, encontram-se as estruturas cujo objetivo é a caracterização estatística detalhada, de um ou mais parâmetros do processo ou da tecnologia de fabricação, incluindo aqui a extração de parâmetros de modelo. Em função da necessidade de confiabilidade estatística, os elementos a serem medidos de cada estrutura são repetidos em número e posições suficientes para que se possa ter uma quantidade de amostras condizente com o grau de confiabilidade desejado.

Os circuitos de teste aqui descrito se encontram nesse segundo tipo, pois devem permitir a caracterização do efeito de descasamento do comportamento CC (corrente-contínua), entre transistores MOS geometricamente idênticos, com confiabilidade estatística suficiente para que possa substanciar o modelo desenvolvido.

4.2 Caracterização de descasamento

O processo de caracterização (medição) de descasamento do comportamento elétrico entre transistores MOS é realizado obtendo-se as curvas desse comportamento nas regiões de polarização de interesse (fig. 4.1). Após ser medido um número suficiente de dispositivos similares (amostra de um grupo com mesmas dimensões), estes resultados passam por um processo de análise estatística, de modo que representem o descasamento com fidelidade suficiente.

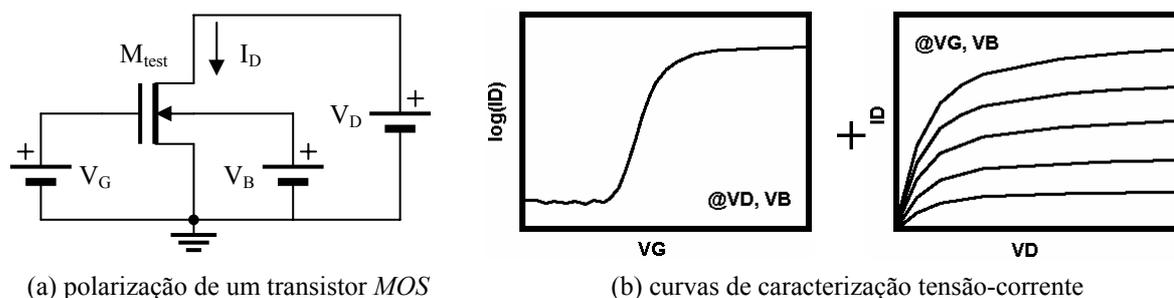
Durante a caracterização, cada transistor sob teste (M_{test}) é polarizado através dos canais de estímulo de um equipamento apropriado (*HP4145, Agilent 4156, Keithley 4200* ou similar), por onde também são medidas as correntes resultantes (no caso de interesse, somente a corrente entre *dreno* e *fonte* - fig. 4.1(a)). Esses canais são aqui denominados *unidades de estímulo e medida (UEM)*.

A natureza estocástica do descasamento faz com que seja necessário um grande volume de medidas, obtidas de grupos de dispositivos “idênticos”. Devido à dependência do descasamento com a geometria dos transistores e com sua polarização, a caracterização deve ser feita com grupos de variadas geometrias e sob uma ampla faixa de condições de polarização. Todas essas exigências tornam o processo de caracterização elétrica do descasamento entre transistores MOS , uma atividade trabalhosa, demorada e muito suscetível a erros de procedimento.

O circuito de teste, por sua vez, tem de comportar um grande número de dispositivos idênticos, agrupados com proximidade (para reduzir os efeitos de gradientes globais sobre a medida), sendo que são necessários vários arranjos destes, com geometrias diferentes, para se avaliar os efeitos geométricos sobre o descasamento. Como cada um dos dispositivos deve ser acessível individualmente, torna-se necessário o uso de alguma forma de multiplexação dos pinos (ou *pads*) de acesso ao circuito, pois estes são (geralmente) em número muito limitado. Deve-se ainda ter cuidado com a forma de acesso aos terminais

dos dispositivos, para que esta não interfira com o resultado da medida, de forma a mascarar ou deturpar a grandeza que se deseja medir.

Assim, vê-se que muitos aspectos desse tipo de circuito devem ser estudados e planejados, de forma a se ter um circuito adequado à caracterização do descasamento elétrico entre transistores *MOS*.



(a) polarização de um transistor *MOS*

(b) curvas de caracterização tensão-corrente

Figura 4.1: Polarização do transistor *NMOS* sob teste (a) e curvas que devem ser obtidas no procedimento de caracterização elétrica, visando a estimativa do descasamento (b).

4.3 Aspectos de projeto

Para embasar o projeto dos circuitos de teste, foi feita uma pesquisa buscando-se artigos relacionados à caracterização de descasamento de transistores *MOS* [83], [96] - [109]. Com base nessa pesquisa e na experiência do autor no projeto de circuitos voltados à medição de grandezas físicas, concluiu-se que as estruturas de teste voltadas esse tipo de medida devem ser planejadas, levando-se em conta os seguintes aspectos:

- tamanho da amostra: quantidade de dispositivos idênticos que formam o arranjo a ser caracterizado. Uma amostra pequena demais resulta em baixa confiabilidade estatística, e grande demais significa uma caracterização demorada e um circuito complexo;
- variedade de geometrias: quantidade de arranjos de dispositivos idênticos. Deve-se ter geometrias que permitam se estabelecer uma correlação do descasamento com as dimensões do dispositivo, indo-se desde dispositivos mínimos, até grandes, curtos e estreitos;
- variedade de tipos de dispositivos: somente tipo N ou P, ou ambos;
- *floor-planning*: em função do grande número de dispositivos, que devem ser distribuídos sobre um circuito para caracterização de descasamento, o cuidadoso planejamento do leiaute se torna fundamental, de modo que se consiga manter os

dispositivos de um mesmo grupo geométrico próximos entre si e acessíveis aos *pads*, com a utilização de trilhas de conexão curtas. Além disso, deve-se evitar o dispêndio de uma grande área de silício;

- tipo de polarização de *porta*: *direta* (a polarização *porta-fonte* (V_{GS}) do dispositivo-sob-teste (DST), é gerada diretamente por um canal do equipamento de medida) ou com *referência* (um transistor independente é polarizado em corrente, fornecendo uma tensão V_{GS} ao DST, em função do nível de inversão em que se encontra);
- forma de estímulo em tensão: *tradicional* (estímulo e medida feitos através do mesmo cabo) ou *force-sense* (ou medida *Kelvin*, quando são utilizados cabos independentes para estímulo e para medida de um mesmo canal, compensando as quedas de tensão através de um laço de realimentação);
- tipo de medida: por *complete pooling* (mede-se o comportamento de todos os dispositivos idênticos individualmente e calcula-se a média e o desvio-padrão do grupo) ou *differential pooling* (dois dispositivos caracterizados simultaneamente por vez, calculando-se o descasamento entre cada par dentro de um grupo, sendo calculado o desvio-padrão dessa diferença);
- acesso aos dispositivos: *direto* (os canais são ligados diretamente aos terminais de *porta*, *dreno* e *fonte* dos DSTs), *seleção parcial* (chaveia-se apenas os terminais de *porta*, mantendo-se os terminais de *dreno/fonte* ligados diretamente aos canais do medidor) e *seleção total* (chaveia-se quaisquer terminais dos DSTs – caso que necessariamente exige uso da técnica *force-sense*, para compensar as quedas de tensão nas chaves por onde circula uma corrente considerável);
- parâmetro a ser extraído: o tipo de curva a ser obtida, e as condições de polarização desta medida, dependem do parâmetro (ou parâmetros) experimental que se deseja obter;
- processo de medição: manual ou automatizado.

Essa classificação permite que se vislumbre os graus de liberdade que se tem no projeto de tal circuito e sua complexidade.

4.4 Quantidade de dispositivos

O tratamento estatístico dado ao resultado das medidas CC de um conjunto de dispositivos de mesma geometria considera que a distribuição do descasamento é aleatória e corresponde a um comportamento *normal*, podendo então ser caracterizado através de sua média e desvio-padrão.

De modo simplificado, para distribuições normais, pode-se estimar o grau de incerteza (desvio-padrão) embutido no desvio-padrão calculado através de uma amostra (σ_p), usando-se a seguinte expressão

$$\sigma_{\sigma_p} = \frac{\sigma_p}{\sqrt{2N_{am}}} \quad (4.4.1)$$

onde N_{am} representa o número de elementos que compõem a amostra.

Deve-se sempre lembrar que a quantidade de dispositivos a serem amostrados nesse tipo de circuito, afeta diretamente a complexidade do processo de medida e o tempo gasto no mesmo. Segundo a expressão apresentada, um grupo de 20 elementos idênticos resulta em 15% de incerteza, e 50 elementos resulta em 10%. A referência [110] sugere um mínimo de 25 a 30 elementos para que a distribuição seja simétrica, podendo ser enquadrada como uma distribuição normal.

4.5 Formas de polarização do dispositivo

A polarização *porta-substrato* pode ser feita diretamente por uma *UEM* (fig. 4.2(a)), no modo de excitação em tensão, ou através de um transistor de referência, conectado na forma de diodo e excitado por corrente (M_{ref} na fig. 4.2(b)). Esta última estratégia é mais interessante por permitir uma visualização mais direta do nível de inversão em que M_{test} se encontra.

Um dos problemas encontrados nesse tipo de caracterização diz respeito à interferência que as quedas de tensão na malha de estímulo causam na tensão real de excitação do dispositivo (fig 4.3(a)). Como, em geral, essas quedas têm característica ôhmica, pode-se reduzi-las através de alguns cuidados no leiaute das linhas de conexão por onde passa corrente, como o uso de trilhas largas e não muito longas, e de um grande número de contatos entre as camadas. Outra estratégia é o emprego de duas vias independentes por canal de caracterização, uma para estímulo e outra para monitoração,

estratégia vulgarmente conhecida como “medida a quatro pontos” ou “medida Kelvin”. Alguns equipamentos de caracterização (p. ex. *Agilent 4156* e *Keithley 4200*) permitem esse tipo de medida, denominada também *force-sense*, disponibilizando dois cabos por canal, de forma que não circule corrente pelo condutor destinado à realimentação de tensão (fig 4.3(b)).

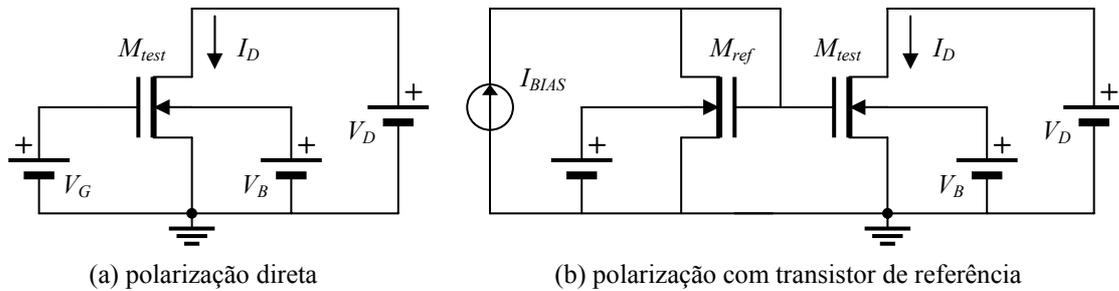


Figura 4.2: Polarização do transistor sob teste (a) direta e (b) com um transistor de referência (V_B e V'_B são uma mesma fonte).

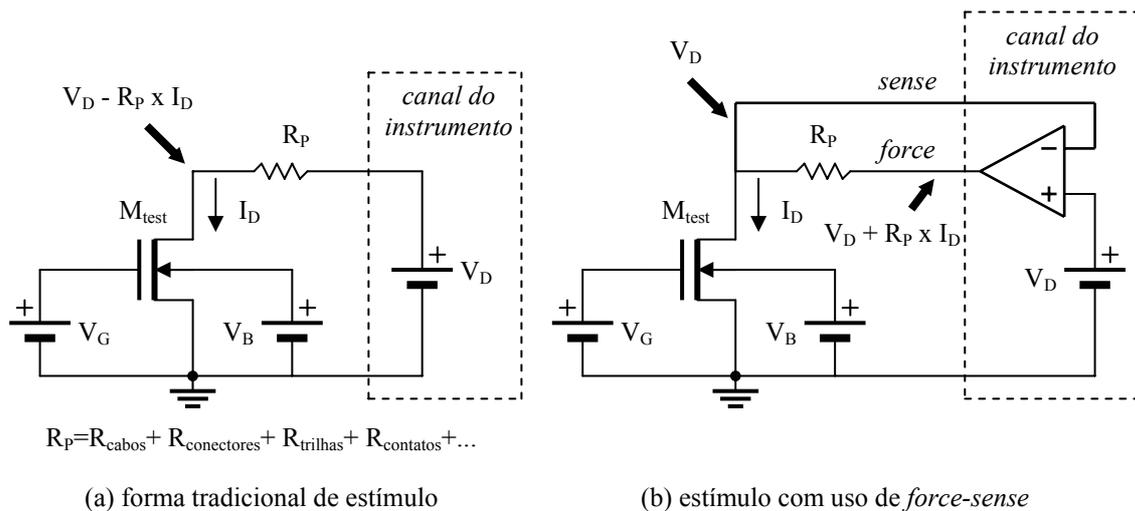


Figura 4.3: Estratégias de estímulo em tensão do dispositivo.

4.6 Formas de seleção do dispositivo

A confiabilidade estatística desse tipo de caracterização depende da quantidade de dispositivos idênticos medidos sob as mesmas condições (amostra). Como não se dispõe de um grande número de equipamentos de caracterização, ou de pontos de acesso na pastilha, é necessário que haja uma forma de se seleccionar cada um dos dispositivos de um grupo para que compartilhe o equipamento e o número limitado de pontos de acesso durante sua caracterização.

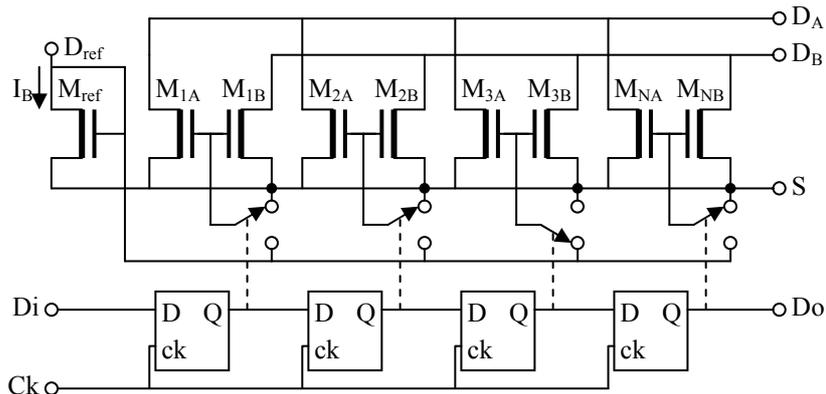


Figura 4.4: Exemplo de transistores de teste dispostos em pares (*differential pooling*) e polarizados por um transistor de referência. A seleção do par a ser medido é interna ao circuito, e realizada pelo terminal de *porta*.

Essa seleção pode ser:

- manual: cada dispositivo ou par é selecionado manualmente através de seus pinos de acesso. As estruturas de teste já implementadas usam essa estratégia, que impõe um grande ônus em termos de tempo e trabalho de caracterização, além da confiabilidade reduzida em função da possibilidade de erro humano;
- seleção na pastilha através de micro-provadores automatizados: é utilizado um equipamento composto de uma mesa de fixação para a pastilha, uma lupa de visualização e um conjunto de micro-agulhas acionadas por servo-mecanismos, controlados por um computador, que as posiciona sobre cada uma das áreas de acesso (*pads*) elétrico de cada dispositivo, estabelecendo o acesso aos seus terminais;
- seleção no encapsulamento através de uma matriz de chaveamento: é utilizado um equipamento composto por uma mesa que contém uma matriz de contatos (pinos ou soquetes), que selecionam independentemente cada pino do circuito encapsulado, através de um conjunto de relés controlados por um computador;
- seleção no circuito através de chaves *CMOS* internas: é a estratégia que exige menor infra-estrutura, pois todo o processo de seleção é feito pelo próprio circuito de teste, através um conjunto de chaves *CMOS* acionadas por uma interface digital que o conecta com um computador (figura 4.4). Entretanto, para que o efeito destas chaves tenha pouco impacto sobre as medidas, deve-se ter o cuidado de incluir a técnica *force-sense* nos caminhos por onde a corrente seja apreciável.

4.7 Circuito de teste – versão 1

O primeiro circuito de teste implementado para caracterização de descasamento, foi fabricado em três tecnologias diferentes, *TSMC 0.35*, *AMIS 0.5* e *AMIS 1.5* (fig. 4.5), através do *Mosis Educational Program (MEP)* – mais detalhes em www.mosis.org. *TSMC* faz referência aos processos da *Taiwan Semiconductor Manufacturing Company Ltd.* e *AMIS* aos processos da *AMI Semiconductor, Inc.* Este circuito é composto por 5 grupos *PMOS* e 5 *NMOS*, cada um contendo 20 transistores idênticos. A tabela 4.1 apresenta as dimensões utilizadas, todas referenciadas ao λ da tecnologia em questão.

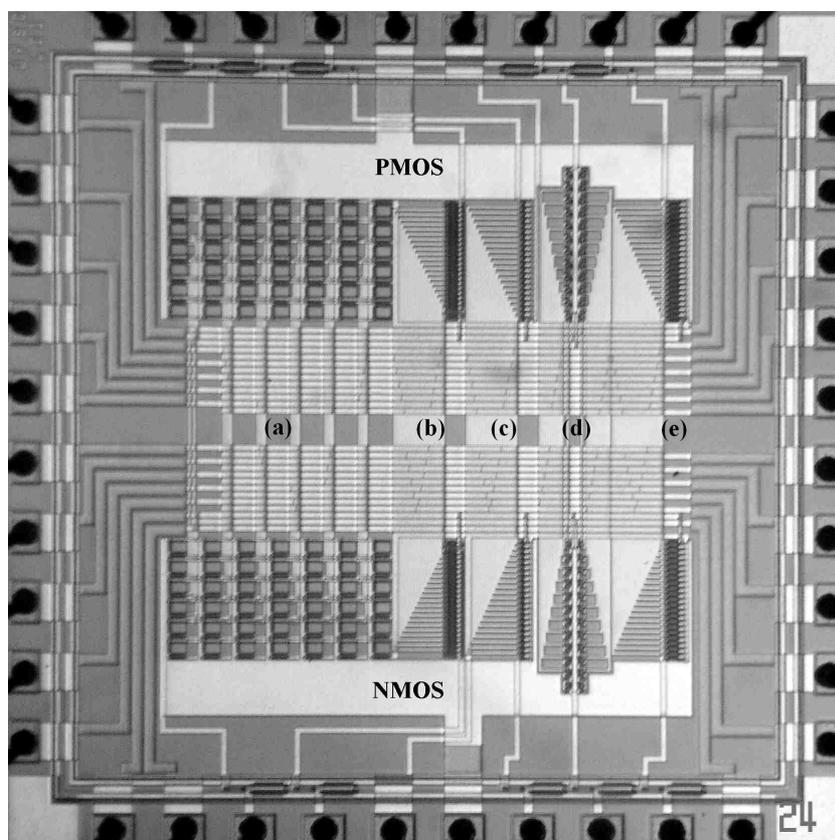


Figura 4.5: Microfotografia da primeira versão do circuito de teste, implementado na tecnologia *AMIS 1,5 μm* . É formado por cinco arranjos de geometrias diferentes: (a) grande, (b) curto, (c) pequeno, (d) médio e (e) estreito. Os arranjos são replicados em transistores *NMOS* e *PMOS*.

Utilizou-se nessas implementações as técnicas de leiaute já descritas no capítulo 1, visando reduzir os efeitos globais de descasamento, pois o objetivo da caracterização é a medida do descasamento local.

A técnica de seleção adotada foi a *manual*, de modo que estes 200 dispositivos (2 tipos x 5 grupos x 20 transistores) estão dispostos na forma de matrizes, compartilhando os 40 pinos de acesso disponíveis no encapsulamento utilizado (*DIP40*). O uso dos pinos de forma compartilhada pelos grupos, faz com que o número de pinos disponíveis no encapsulamento restrinja diretamente a quantidade de dispositivos para teste que podem ser acessados no circuito. A figura 4.6 apresenta o diagrama esquemático de dois grupos, um N e outro P (as conexões de substrato só estão representadas nos últimos dispositivos para simplificar o desenho). Cinco conjuntos como este, cada um com transistores de geometria diferentes, são ligados aos mesmos pinos de *dreno*, podendo ser selecionados através de diferentes pinos de *porta*. A limitação do número de conjuntos que podem ser unidos em paralelo, compartilhando os mesmos pinos de *dreno*, está relacionada às correntes de fuga (I_{OFF}) dos transistores que os compõem.

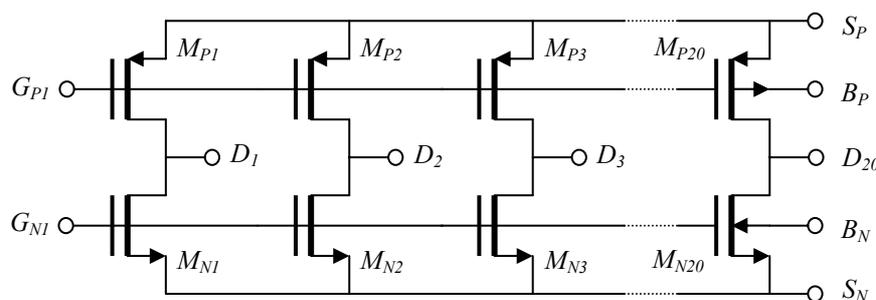


Figura 4.6: Diagrama de dois grupos (N e P) de transistores de teste.

Tabela 4.1: Dimensões das estruturas de teste (em μm), implementadas na versão 1.

Tecnologia	tipo:	pequeno	médio	grande	curto	estrito
	$W \times L [\lambda] \rightarrow$	3x2	12x8	48x32	48x2	3x32
	$\lambda [\mu\text{m}] \downarrow$	$(W_{\min} \times L_{\min})$	$(4W_{\min} \times 4L_{\min})$	$(16W_{\min} \times 16L_{\min})$	$(16W_{\min} \times L_{\min})$	$(W_{\min} \times 16L_{\min})$
<i>TSMC 0.35</i>	0,25	0,75x0,5	3x2	12x8	12x0,5	0,75x8
<i>AMIS 0.5</i>	0,35	1,05x0,7	4,2x2,8	16,8x11,2	16,8x0,7	1,05x11,2
<i>AMIS 1.5</i>	0,80	2,4x1,6	9,6x6,4	38,4x25,6	38,4x1,6	2,4x25,6

Utilizando-se a disposição multiplexada dos pinos de acesso, necessita-se:

- 20 pinos para conexão dos drenos (P e N em paralelo);
- 10 pinos para conexão das portas (5 P e 5 N);
- 2 pinos para conexão dos substratos (1 P e 1 N);

- 2 pinos para conexão das fontes (1 P e 1 N);

totalizando 34 pinos de acesso ao circuito. Os pinos excedentes foram utilizados na conexão de circuitos e dispositivos extra, que foram incorporados à pastilha.

Na implementação, utilizou-se um *padframe* (estrutura que envolve a área da pastilha e faz a conexão dos terminais do circuito, com os fios de contato do encapsulamento) especificamente desenhado para cada tecnologia, incluindo estruturas de proteção contra descargas eletrostáticas nos *pads* conectados aos terminais de *porta*.

Todos estes circuitos já foram caracterizados, extraindo-se alguns parâmetros da tecnologia e curvas de descasamento em diversas condições de operação. Em função do impacto tecnológico, o processo de menor comprimento de canal (*TSMC 0.35*) foi o mais caracterizado. Os resultados apresentados no capítulo 3 e publicados em [64], [111] e [112] foram obtidos com este circuito.

As medidas foram feitas sempre utilizando-se um dos transistores (fixo) como referência de polarização em corrente, sobrando então 19 dispositivos para caracterização. As medidas iniciais foram feitas em *complete pooling*, passando-se depois para o uso de *differential pooling*, pois os resultados obtidos eram mais coerentes. Como os transistores foram dispostos em filas, caracterizou-se sempre pares contíguos e próximos entre si.

Para reduzir o efeito do erro de descasamento dos canais do equipamento de medida (UEMs) sobre o resultado, a medida é feita em dois momentos. Implementou-se um conjunto de chaves externas, que são acionadas manualmente de forma que as UEMs que caracterizam cada transistor do par, sejam alternadas em cada momento. A seguir, é realizada a média das medidas obtidas nos dois momentos.

Apenas com cunho informativo, cita-se que a caracterização do descasamento de apenas um dos 10 grupos presentes em um destes circuitos, através do método manual de seleção do dispositivo, é atividade que necessita de 2 a 3 horas para ser desenvolvida, estando ainda muito suscetível a erros humanos.

4.8 Circuito de teste – versão 2

Dentre todas as possibilidades de topologia pesquisadas e estudadas para o circuito de teste, pode-se listar os seguintes aspectos como sendo os mais significativos no desempenho deste.

- a) Processo de caracterização totalmente automatizado (fig. 4.7), onde um computador seleciona o DST, dentre todos os dispositivos disponíveis, programando um registrador interno ao circuito, e coordena o estímulo e a medida deste, programando o equipamento de caracterização elétrica, através de uma interface de controle. O uso de um registrador interno com acesso serial contribui com a redução do número de pinos de acesso necessários no encapsulamento.
- b) Uso de chaveamento interno dos terminais de *porta* e *dreno* (fig. 4.8), de modo a não haver restrições no circuito relacionadas ao número de pinos do encapsulamento (essa decisão implica na necessidade de se utilizar o método *Kelvin (force-sense)* para compensar a queda de tensão nas chaves por onde passam correntes significativas, restringindo a caracterização do circuito a equipamentos com essa facilidade).
- c) Extensão das trilhas de *force-sense* de *dreno* e *fonte* até o ponto mais próximo possível dos terminais do transistor a ser caracterizado (fig. 4.8), de modo a se compensar a maior gama de quedas de tensão possível, sem que isso implique em aumento substancial de complexidade.
- d) Distribuição posicional dos DSTs adequada para a caracterização em *differential pooling*, pois este tipo de medida evidencia as causas de descasamento *locais*, sendo menos suscetível aos efeitos *globais*.
- e) Realização das medidas de forma simultânea, em cada par de transistores, através do uso de dois canais de polarização de *dreno* (fig. 4.9), de forma a se reduzir o efeito de variações temporais (térmicas p. ex.), durante a sua caracterização. A medida é feita em dois momentos, sendo que os canais do equipamento (UEM) que caracterizam cada transistor do par são alternados em cada um deles, e depois é realizada a média das medidas obtidas nos dois momentos. Essa estratégia de reversão das UEMs reduz o impacto do erro de descasamento dos canais de medida, sobre o resultado.
- f) Uso de polarização através de um transistor de referência adicional por grupo, facilitando a caracterização do descasamento em função do nível de inversão (fig. 4.9).
- g) Monitoração da saída da interface digital, de forma a se poder conferir a validade do vetor de programação transferido.

- h) Implementação de uma interface opto-isolada entre o computador e os pinos de acesso à interface digital do circuito de teste, de modo a se evitar a injeção de ruído elétrico por parte da interface de controle.

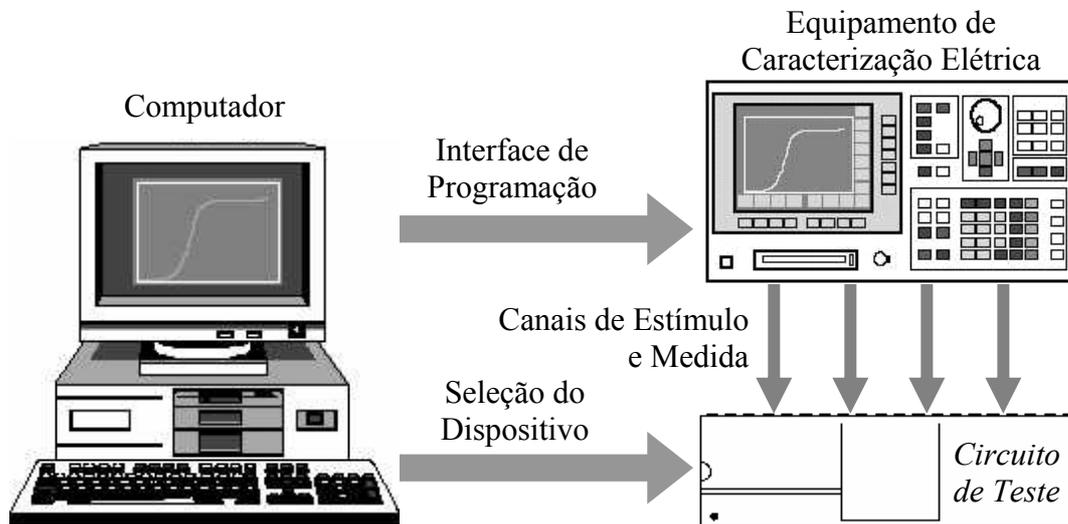


Figura 4.7: Sistema proposto para caracterização de descasamento de transistores MOS.

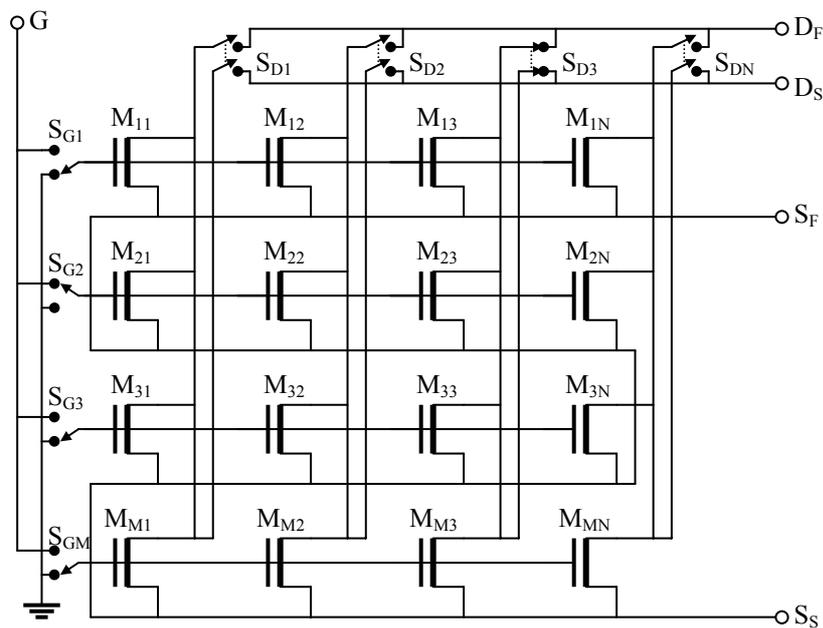


Figura 4.8: Conceito da topologia proposta em *complete pooling* (apenas tipo N desenhado), incluindo as chaves de seleção de *dreno* e de *porta*.

A proposta de topologia que segue, é a que melhor coaduna os aspectos apresentados, mantendo-se numa complexidade e consumo de área razoáveis, podendo vir a ser encapsulada até na opção de 28 pinos da *MOSIS (DIP28)*. As figuras 4.8 e 4.10 apresentam essa topologia nas versões *complete* e *differential pooling*, sendo que apenas os

transistores N são representados. Na fig. 4.8 é apresentado o conceito de chaveamento por *porta* e por *dreno*, com uma matriz composta de quatro grupos, sendo cada grupo composto por quatro transistores idênticos. Os transistores de referência não são apresentados. Na fig. 4.10 é apresentado o diagrama esquemático simplificado da topologia implementada, com uma matriz composta de três grupos, cada um com três pares de transistores idênticos. As chaves que possibilitam a reversão dos canais de medição de *dreno* (fig. 4.9) também estão representadas.

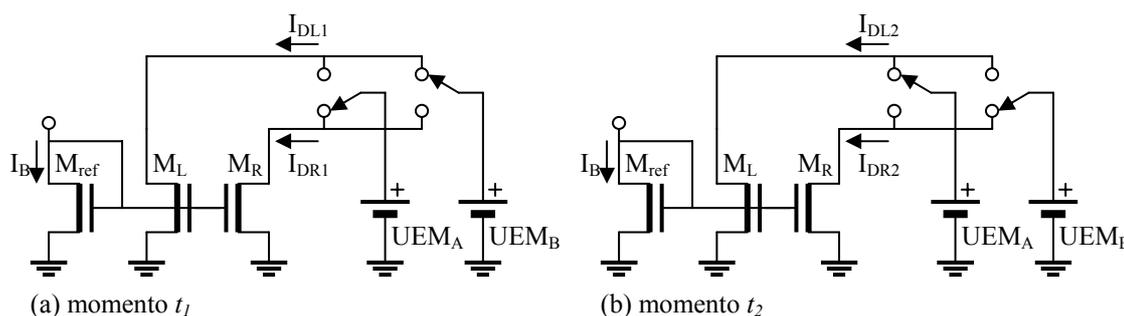


Figura 4.9: Estratégia de caracterização da corrente de dreno em *differential pooling*, em dois momentos, com o emprego da reversão das UEMs, visando à redução do efeito do erro do descasamento existente entre os canais do equipamento, sobre o descasamento entre os dispositivos a ser medido .

Pode-se observar na fig. 4.10 que um par específico “ i ” de um determinado grupo de transistores idênticos pode ser selecionado através das chaves de *porta* (S_{Gi}). A seleção do grupo de transistores idênticos “ j ” a ser caracterizado é feita através das chaves de *dreno* (S_{jLA} , S_{jRA} , S_{jLB} , e S_{jRB}). As chaves S_{Bj} selecionam o transistor de referência a ser utilizado com o grupo respectivo.

Observa-se que o circuito permite a compensação das quedas de tensão no caminho que vai dos pinos até o ponto de conexão ao dispositivo, incluindo as quedas nas chaves e fiação, só restando as quedas nos contatos entre o nível de metal e as regiões de *dreno* e *fonte* no substrato. E esta compensação é realizada para os terminais de *dreno* e de *fonte*, que é por onde circula a quase totalidade da corrente no transistor *MOS*.

O circuito necessita dos seguintes pinos de conexão:

- 2 para alimentação (V_{cc} e Gnd);
- 3 para programação (Di , Do e Ck);
- 3 terminais de *fonte* por tipo (S_F , S_S e S_{ref} , que é a fonte das referências);

- 3 terminais de *dreno* por tipo (D_{Ref} , D_S e D_F), para *Complete Pooling*, ou;
- 5 terminais de *dreno* por tipo (D_{Ref} , D_{SA} , D_{FA} , D_{SB} e D_{FB}), para *Differential Pooling*.

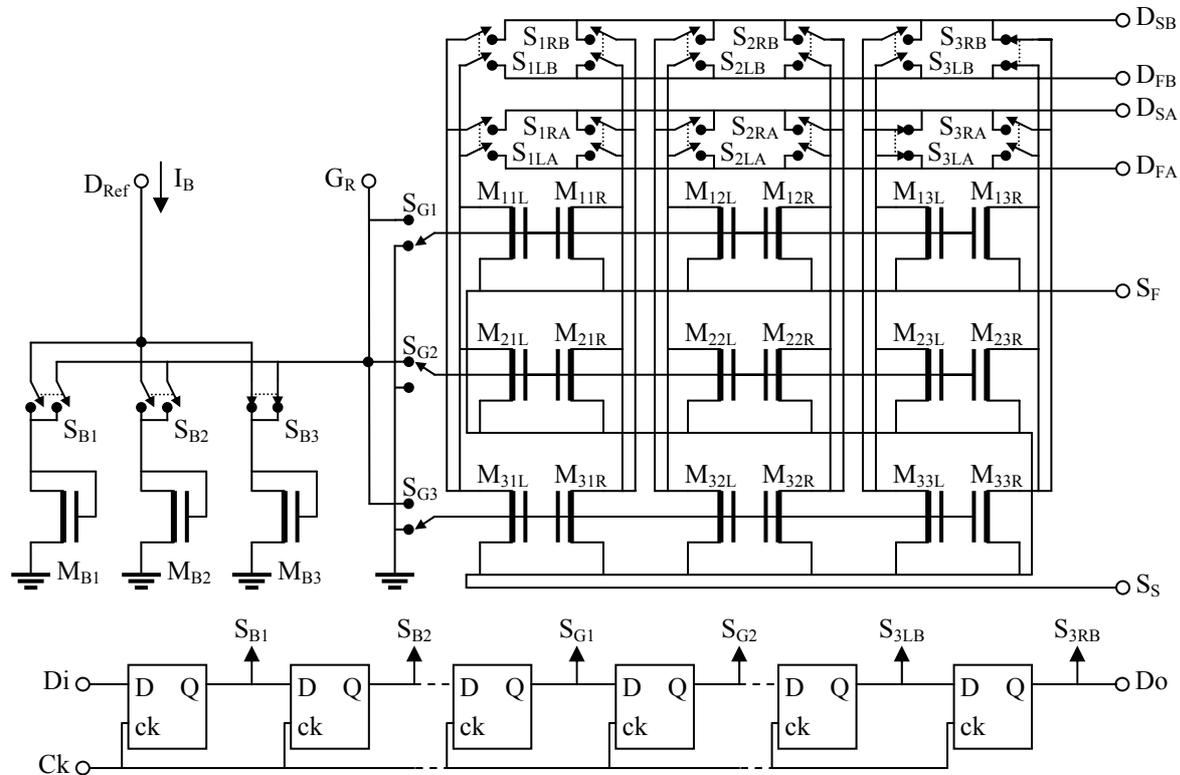


Figura 4.10: Topologia proposta em *differential pooling* (apenas transistores tipo N apresentados). As chaves de seleção de *drenos*, que possibilitam o acesso a cada um dos arranjos de transistores idênticos (na vertical), foram implementadas para permitir o uso da técnica *force-sense*. Esta técnica também é utilizada no acesso aos terminais de *fonte*. Em cada arranjo, um par específico pode ser selecionado pela chave de *porta* (S_{Gi}).

As chaves S_{Bj} permitem a seleção do transistor de referência de interesse.

Com base nessas necessidades, pode-se calcular o número total de pinos necessários no encapsulamento, para cada possibilidade de implementação:

- *Differential Pooling* com N e P juntos: 21 pinos;
- *Differential Pooling* com apenas N ou P: 13 pinos;
- *Complete Pooling* com N e P juntos: 17 pinos;
- *Complete Pooling* com apenas N ou P: 11 pinos.

Em virtude do reduzido número de pinos exigido pelo circuito, poder-se-ia também distribuir os terminais de *dreno* de grupos diferentes em pinos diferentes (visto que o circuito pode ser encapsulado com até 40 pinos através da *MOSIS*), de forma a se reduzir a

quantidade de chaves e transistores desligados (em estado *off*) conectados a um mesmo pino de acesso, pois isto restringe o valor mínimo de corrente que pode ser medido em um dispositivo, devido à perturbação causada na medida pelas suas correntes de fuga (I_{OFF}).

Uma possível estratégia para se verificar e reduzir a perturbação das correntes de fuga (das chaves e transistores desligados) sobre a corrente medida, é a de se medir, antes e depois da caracterização de cada grupo, o valor total da corrente nos pinos de interesse com todos os transistores *desligados*. Assim, tem-se uma aproximação do valor dessa corrente de fuga dentro de toda a faixa de tensões de excitação. Se ambas as curvas (de antes e depois) apresentarem pouco desvio entre si, o sistema está com boa estabilidade e pode-se simplesmente subtrair os valores médios dessas curvas das curvas de caracterização do grupo, compensando a fuga. Caso as curvas de antes e depois apresentem grande divergência, considera-se que o processo de medida foi feito em condições adversas (p. ex. grande variação térmica durante a caracterização) e este é desprezado.

Foram também implementadas algumas junções P-N (3 diodos e 1 transistor bipolar parasita vertical), distribuídas em pontos estratégicos do circuito (próximo aos transistores de teste que dissipam maior potência) e ligadas a pinos independentes, de modo que se possa avaliar o impacto do processo de caracterização na temperatura do dispositivo, através da medida da variação na tensão dessas junções, quando polarizadas com corrente constante.

4.8.1 Geometria e leiaute

Sabe-se que o descasamento entre transistores *MOS* é também função de suas dimensões físicas, dependendo de sua área, e do comprimento e largura de canal, conforme visto anteriormente. Assim, é interessante que o circuito de teste possua um conjunto de geometrias que cubra ao menos três possibilidades de cada dimensão, ou seja, *pequeno*, *médio* e *grande*, de modo a se poder comparar os efeitos dessas três dimensões relativas nos eixos de comprimento (L) e largura (W) de canal.

Empregando-se a dimensão *pequena* como a mínima possível na tecnologia, pode-se obter os efeitos de canal curto e estreito bastante destacados nas medidas, o que é interessante para se estimar o impacto destes no modelo. Já a dimensão *grande* permite que se obtenha medidas sem qualquer influência desses efeitos. Assim, decidiu-se que tanto L quanto W teriam as seguintes dimensões: $1 D_{min}$, $4 D_{min}$ e $16 D_{min}$, onde D_{min} é a dimensão

mínima para cada eixo (3λ para W e 2λ para L , nas regras escaláveis fornecidas pela *MOSIS* – na tecnologia *TSMC 0.35* o valor de λ é $0,25\mu\text{m}$ e na *TSMC 0.18* seu valor é $0,10\mu\text{m}$). A tabela 4.2 apresenta as dimensões dos transistores de teste em detalhe.

Tabela 4.2: Dimensões dos transistores de teste, implementados na versão 2. Sem parênteses estão as dimensões em μm para a tecnologia *TSMC 0.35* ($\lambda = 0,25\mu\text{m}$), e entre parênteses, para a *TSMC 0.18* ($\lambda = 0,10\mu\text{m}$).

$W \times L$ [μm]		W :		
		$1W_{min} = 3\lambda$	$4W_{min} = 12\lambda$	$16W_{min} = 48\lambda$
L :	$1L_{min} = 2\lambda$	0,75 x 0,5 (0,3 x 0,2)	3 x 0,5 (1,2 x 0,2)	12 x 0,5 (4,8 x 0,2)
	$4L_{min} = 8\lambda$	0,75 x 2 (0,3 x 0,8)	3 x 2 (1,2 x 0,8)	12 x 2 (4,8 x 0,8)
	$16L_{min} = 32\lambda$	0,75 x 8 (0,3 x 3,2)	3 x 8 (1,2 x 3,2)	12 x 8 (4,8 x 3,2)

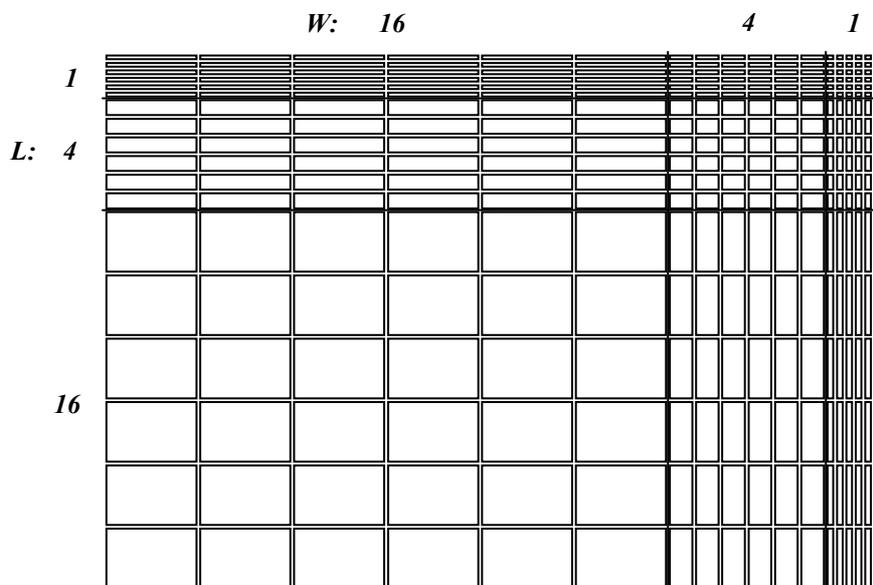


Figura 4.11: Proposta de disposição dos transistores de teste em grupos de 6x6 pares.

Um fator importante nesse projeto é a otimização do uso da área do silício. Assim, a geometria proposta é composta por conjuntos de transistores idênticos, distribuídos numa matriz com colunas de L constante e linhas de W constante, permitindo a verificação da correlação nos resultados quando uma dessas variáveis é mantida constante. No total, o conjunto resulta em 9 grupos de transistores, cada um composto por 36 pares de

transistores idênticos (*differential pooling*), e mais 9 transistores de referência, para polarização *porta-substrato*. Com uma amostra desse tamanho, espera-se incertezas na medida do desvio-padrão, da ordem de 11%.

Como os 9 grupos possuem dimensões correlacionadas, estes podem ser agrupados de forma a se otimizar a área, conforme a figura 4.11. Nessa figura, vê-se a disposição esquemática dos transistores de teste em grupos de 6x6 pares (cada retângulo representa um par), distribuídos de modo que, na vertical estão dispositivos com mesmo W e na horizontal, com mesmo L .

Na figura 4.12 vê-se a planta-baixa proposta para o circuito, mostrando a disposição dos conjuntos de transistores de teste N e P, das chaves de seleção de *dreno* e de *porta*, dos transistores de referência com suas respectivas chaves e do registrador de deslocamento.

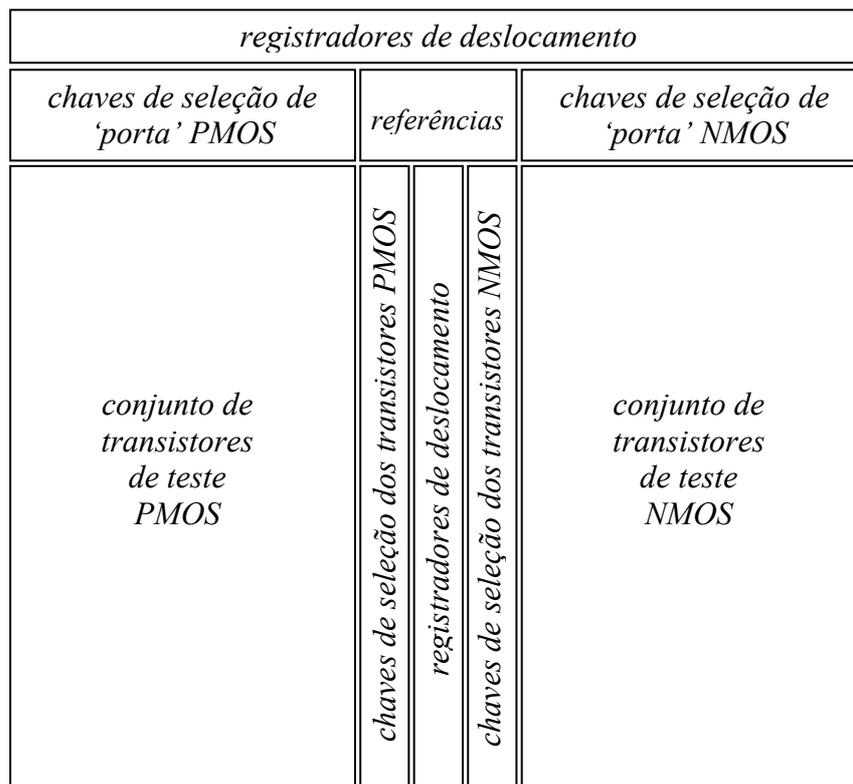


Figura 4.12: Proposta de planta-baixa para o circuito.

A figura 4.13 apresenta a disposição e o leiaute de 4 pares de transistores de teste, com as respectivas conexões. Em cada par, os transistores são desenhados respeitando-se as distâncias mínimas das regras escaláveis, de modo que estes fiquem com a menor distância possível entre si, reduzindo os efeitos globais no seu descasamento. Através de 6 grupos de 6 linhas verticais pode-se selecionar, pelo terminal de *porta*, o par de transistores

a ser caracterizado, de um grupo de 36 pares. Os 9 grupos compartilham os mesmos pinos de *dreno*, sendo cada um individualmente selecionada através de chaves. A trilha inscrita com *B* é a conexão de substrato (*body-tie*).

Estima-se que o circuito completo deva ter cerca de $3000\lambda \times 3500\lambda$, o que corresponde a $750\mu\text{m} \times 875\mu\text{m}$ na tecnologia *TSMC 0.35*.

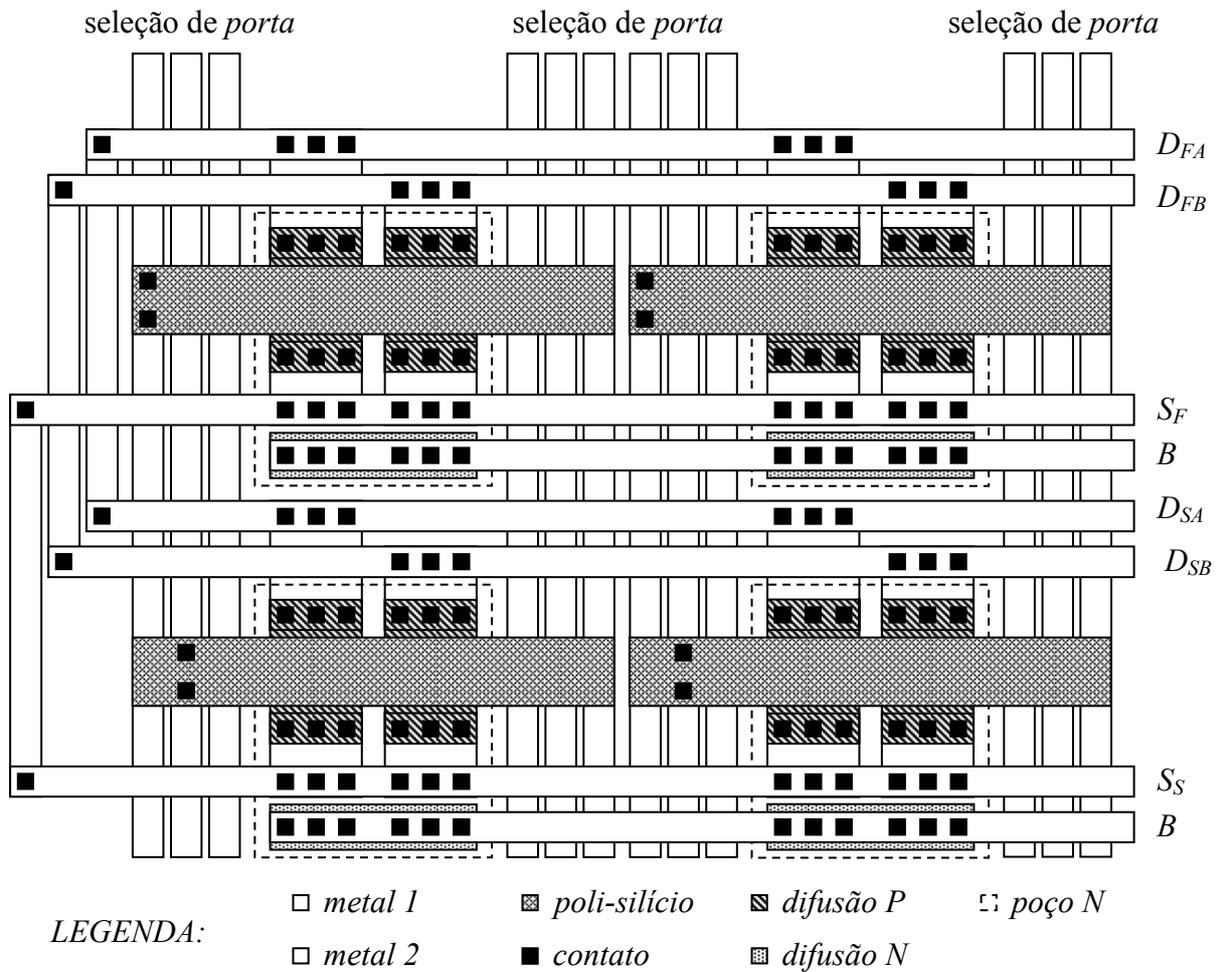


Figura 4.13: Proposta de leiaute de um grupo PMOS, mostrando a disposição dos pares de transistores de teste e conexões. O leiaute do grupo NMOS é análogo.

4.8.2 Especificação e projeto

Esse circuito foi desenvolvido visando sua implementação original na tecnologia *TSMC 0.35*, de forma que se está considerando os seus parâmetros para o projeto (tabela 4.3). Em uma segunda etapa, foi re-escalado e implementado também na *TSMC 0.18*, necessitando somente algumas verificações e simulações.

Como o dimensionamento dos dispositivos de teste já foi feito, resta nessa etapa dimensionar as chaves (transistores de passagem), o registrador de deslocamento e os *buffers* que são ligados aos pinos de entrada e saída digital.

O projeto dos circuitos de seleção (chaves) é apresentado para o caso dos transistores de teste *NMOS*, sendo o caso dos *PMOS* perfeitamente análogo. O mesmo ocorre para as simulações.

Tabela 4.3: Principais parâmetros da tecnologia *TSMC 0.35*, obtidos no site www.mosis.org, para a rodada T3BM (LO_EPI). O valor de n foi estimado através de simulações com o modelo BSIM 3V3 fornecido no mesmo site. A corrente específica por quadrado - I_{SQ} - foi calculada através dos parâmetros da tecnologia.

	μ_0 [cm ² /V.s]	C'_{ox} [fF/μm ²]	V_{T0} (trans. curto) [V]	V_{T0} (trans. grande) [V]	n	I_{SQ} [nA]
NMOS	417,44	4,54	0,62	0,57	1,25	80,0
PMOS	146,37	4,44	-0,74	-0,70	1,30	28,6

4.8.3 Seleção pela porta

A figura 4.14 apresenta o diagrama esquemático simplificado do circuito de chaveamento do terminal de *porta* que foi utilizado.

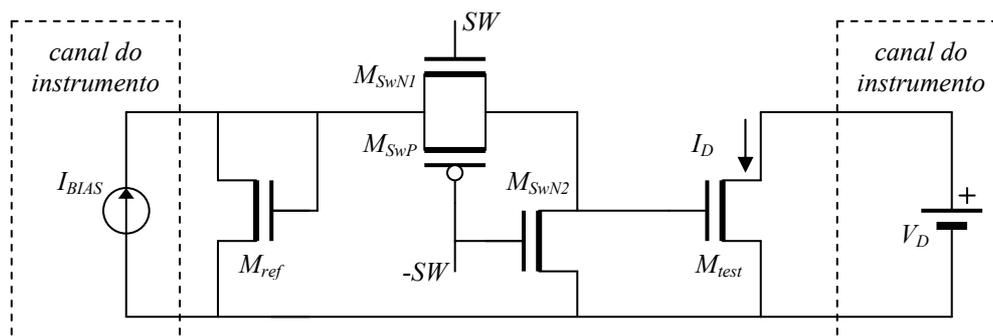


Figura 4.14: Chaveamento do transistor de teste pela porta.

Os sinais de controle SW e $-SW$ provém da saída de um dos registradores. Utiliza-se um par N-P como chave na situação em que seus terminais de *dreno* e *fonte* possam flutuar em tensão (M_{SwN1} e M_{SwP}), em relação à alimentação, e apenas um tipo de transistor onde o terminal de *fonte* tem seu potencial fixo (M_{SwN2}). Como a corrente que circula por essas chaves é praticamente nula e o circuito vai operá-las em frequências muito baixas,

estas podem ser especificadas com dimensões pequenas (optou-se por $W/L_N = 2\mu\text{m}/0,5\mu\text{m}$ e $W/L_P = 4\mu\text{m}/0,5\mu\text{m}$).

Esse circuito de seleção foi simulado para uma vasta gama de condições de operação e mostrou-se com funcionamento adequado.

4.8.4 Seleção pelo dreno

O circuito da figura 4.15 apresenta o diagrama esquemático simplificado, escolhido para o chaveamento do terminal de *dreno* dos dispositivos de teste, no modo *force-sense*.

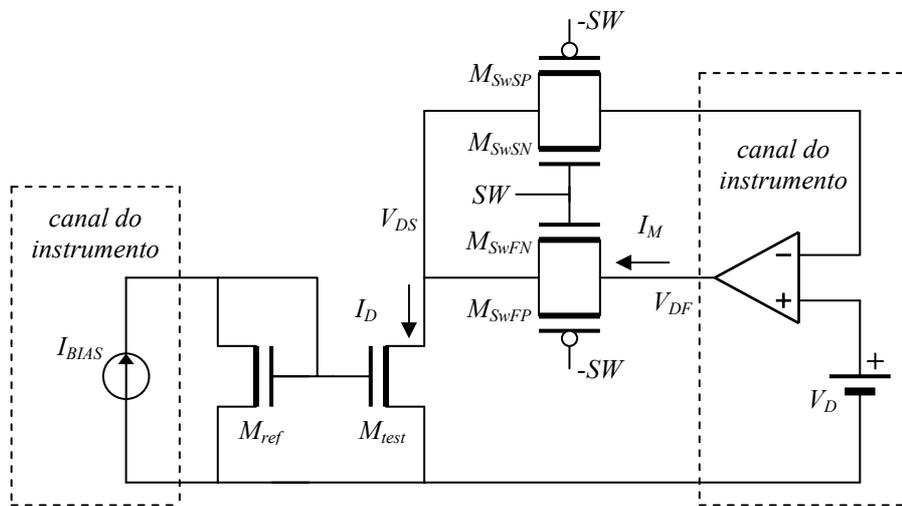


Figura 4.15: Chaveamento do transistor de teste pelo *dreno*.

A corrente que circula pelo condutor de medida (*sense*) de tensão é praticamente nula, de modo que os transistores que compõem essa chave (M_{SwSN} e M_{SwSP}) podem ter dimensões pequenas (optou-se por $W/L_N = 2\mu\text{m}/0,5\mu\text{m}$ e $W/L_P = 4\mu\text{m}/0,5\mu\text{m}$).

Como há corrente significativa circulando através dos transistores que compõem a chave do caminho de estímulo (M_{SwFN} e M_{SwFP}), ocorre uma queda de tensão sobre estes, em função da sua impedância, que deve ser avaliada.

Determinou-se como limite da região de caracterização dos dispositivos, o nível de inversão 1000, pois dificilmente se necessita de transistores casados em níveis de inversão superiores a esse. Em função disso, estima-se que a queda de tensão necessária para saturar um transistor, nesse nível de inversão, seja da ordem de 900mV [65]. Como se pretende caracterizar o descasamento também para a situação de saturação, definiu-se como limite de tensão a ser atingida no *dreno* de M_{test} o valor de 2,5V. Isso faz com que, para que não

se ultrapasse o limite da tensão de alimentação de 3,3V no ponto V_{DF} , a queda máxima aceitável na chave seja de 0,8V.

A condição de máxima corrente na chave ocorre para o transistor de teste de maior relação de aspecto, no caso o *curto* (12/0,5). Considerando-se a corrente específica por quadrado do processo ($I_{SQN} = 80$ nA e $I_{SQP} = 29$ nA), chega-se a uma corrente máxima de cerca de 2mA e 0,7mA para os transistores *NMOS* e *PMOS* respectivamente, em nível de inversão 1000.

Considerando-se que o limite de queda de tensão na chave é mais importante para valores de V_{DS} altos (tendo como limite 2,5V), e nessa situação o transistor da chave que predomina na condução de corrente é o *PMOS*, este deve ser especificado para apresentar uma queda máxima de 0,8V enquanto conduz 2mA e é submetido à tensão de V_{GS} , igual a -3,3V. Utilizando-se a expressão (4.8.1) [67], calculou-se a relação de aspecto desse transistor (M_{SwFP}) como 23.

$$I_D = \frac{W}{L} \mu C'_{ox} \left[(V_{GB} - V_{T0})(V_{DB} - V_{SB}) - \frac{n}{2} (V_{DB}^2 - V_{SB}^2) \right] \quad (4.8.1)$$

Nos transistores de passagem é empregado comprimento mínimo de canal, de modo que a razão de aspecto deve ser aumentada (sobre-dimensionada) para compensar os efeitos de canal curto e a margem de incerteza nos parâmetros nominais do processo.

O transistor *NMOS* (M_{SwFN}) só opera na condição de V_{DS} baixo, enquanto V_{DF} não estiver acima da tensão de limiar do transistor *PMOS* (M_{SwFP}). Assim, o transistor M_{SwFN} necessita de uma baixa razão de aspecto, visto que a corrente que por ele circula é sempre reduzida.

Simulando-se o circuito da figura 4.15, e considerando-se ainda, como fator de segurança, uma corrente de dreno de 5mA, chega-se a uma razão de aspecto para M_{SwFP} de 120 ($60\mu\text{m}/0,5\mu\text{m}$) e para M_{SwFN} de 24 ($12\mu\text{m}/0,5\mu\text{m}$), para uma queda máxima de 0,7V na chave.

É importante aqui observar que para o caso dos transistores de teste com menor razão de aspecto, pode-se reduzir também essa razão dos transistores de passagem respectivos na mesma proporção, em função da redução no valor da corrente máxima de dreno para o nível de inversão 1000.

Essa estratégia é também empregada na seleção do *dreno* do transistor de referência que está ligado como diodo e que deve ser conectado ao pino D_{ref} em função do grupo a ser caracterizado (figura 4.16). Para o chaveamento destes transistores, a chave de *sense* não é necessária, visto que sua excitação é feita em corrente.

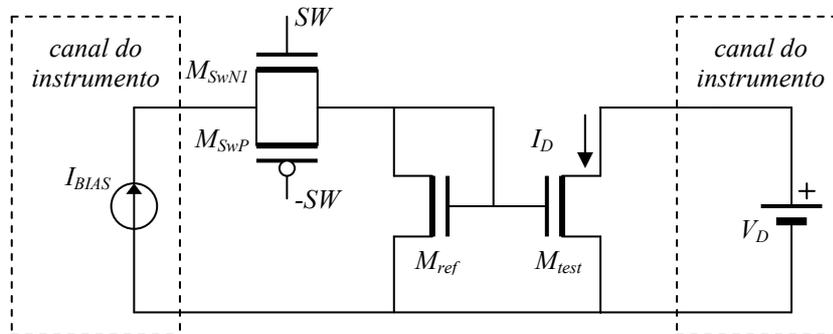


Figura 4.16: Chaveamento do transistor de referência.

A tecnologia alvo (*TSMC 0.35*) utiliza substrato tipo P para os transistores NMOS e poço tipo N para os PMOS. Isso faz com que todos os NMOS operem sob a mesma tensão de *substrato*, incluindo os transistores das chaves e dos registradores, pois este terminal está interligado pelo poço ao pino de terra (*GND*) em todo o circuito (no caso dos grupos PMOS, seus poços foram interligados ao pino de V_{CC}). Como visto no capítulo 3, é interessante que se obtenha dados sobre o comportamento do descasamento em função da polarização de substrato. Entretanto, dadas estas limitações do circuito, talvez a pequena margem de tensão que se possa aplicar ao *substrato* seja insuficiente para se medir esse efeito. De qualquer modo, os circuitos de seleção apresentados foram simulados inclusive para a condição de tensão de *substrato* negativo (exagerada, até $-4V$), pois nessa condição os transistores apresentam a maior tensão de limiar, resultando numa maior tensão *porta-fonte*, necessária para uma dada corrente de *dreno*.

4.8.5 Registrador de deslocamento

Nesse projeto, optou-se por não despendar muito tempo no desenvolvimento dessa parte, utilizando-se alguma topologia clássica e simples. Através da página da *MOSIS* na *internet*, obteve-se acesso a uma biblioteca de células digitais básicas (incluindo leiaute), de onde se extraiu uma topologia de circuito registrador tipo *flip-flop* com pequenas dimensões, composto por dois registradores tipo *latch*, ligados em configuração *mestre-escravo*.

A figura 4.17 apresenta o diagrama de um dos registradores tipo *latch*, cujo funcionamento foi simulado para as dimensões $W/L_N = 1,5\mu\text{m}/0,5\mu\text{m}$ e $W/L_P = 2,5\mu\text{m}/0,5\mu\text{m}$, tendo se comportado a contento. Nessa figura, D representa o dado de entrada, CK é o sinal de carga e Q e $-Q$ são as saídas.

Em função do modo como a matriz de transistores de teste é acionada, dividida em 9 grupos chaveados por *dreno* (4 conjuntos de chaves por grupo) por 36 pares de dispositivos chaveados por *porta* e mais 9 referências, são necessários 81 bits de registro, ou seja, 81 *flip-flops* em cascata.

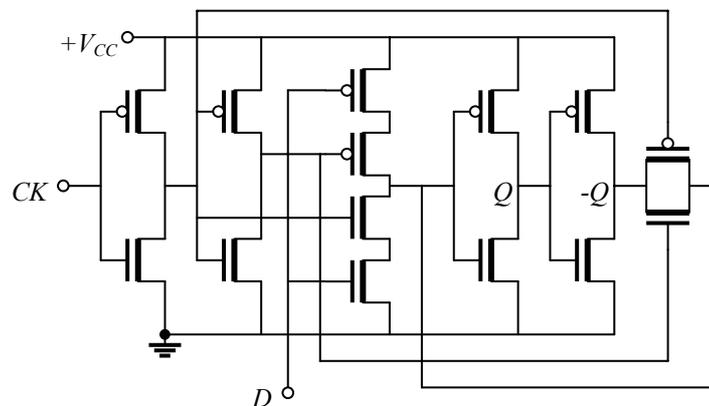


Figura 4.17: Circuito esquemático do registrador tipo *latch*.

Como são utilizados pinos separados de acesso aos terminais dos conjuntos de transistores de teste *NMOS* e *PMOS*, ambos podem compartilhar as mesmas saídas do registrador, sendo acionados em paralelo.

4.8.6 Buffers de entrada e saída

O sinal de *dado* (D_i) não necessita reforço, visto que só é empregado no primeiro registrador. Para o sinal de *carga* (CK), um circuito de reforço (*buffer*) foi previsto, pois esse sinal será aplicado a 81 registradores. Mas, visto que a interface não operará em frequências elevadas, esse circuito não necessita elevada capacidade de corrente de saída.

O sinal de saída de dado do último registrador (D_o) tem como função uma forma de se verificar, através do computador, se os dados enviados estão sendo registrados adequadamente. Assim, um *buffer* com mais capacidade de corrente foi previsto nesse caso, visto que esse sinal será lido por um pino da porta paralela do computador.

Considerando que esse sinal deve ter capacidade de carga de cerca de 2mA (para excitar facilmente uma entrada *TTL*) e apresentar um erro máximo relativo às tensões de alimentação de 0,3V (ou seja, a saída deve atingir no máximo +0,3V para o nível lógico *zero* e no mínimo +3V para o nível *um*), calculou-se que os transistores de saída desse *buffer* devem ter razões de aspecto de 14 para o *NMOS* e 42 para o *PMOS*. Considerando-se os efeitos de canal curto e uma margem de segurança adicional, eles foram dimensionados com $W/L_N = 20\mu\text{m}/0,5\mu\text{m}$ e $W/L_P = 50\mu\text{m}/0,5\mu\text{m}$.

Na implementação em *TSMC 0.35*, observou-se que ocorre um problema de escorregamento do vetor de bits através do registrador de deslocamento, caso o *tempo de subida ou descida* do sinal de *carga* seja superior a 70ns. Este problema não havia sido observado nas simulações, assim como não havia qualquer restrição temporal a este sinal nas especificações do *flip-flop* utilizado, na biblioteca da *MOSIS*. A solução para este inconveniente foi a inclusão de um *buffer* de transição rápida em série com este sinal, externamente ao circuito. Na implementação feita em *TSMC 0.18*, o *buffer* interno foi reprojetoado, passando então a ser composto por 4 estágios inversores, de modo a adequar os tempos de transição do sinal que vem da interface óptica às exigências do *flip-flop*.

4.8.7 Fabricação e resultados

Esta versão do circuito de teste foi implementada nas tecnologias *TSMC 0.35* e *TSMC 0.18*, através do programa *MEP* da *MOSIS*. Em função das dimensões do circuito e das restrições de área impostas pela *MOSIS* (pastilhas com 1,5 mm de lado), dificilmente este poderia ser fabricado em uma tecnologia com comprimento de canal superior a 0,35 μm .

A fig. 4.18 apresenta uma microfotografia da pastilha, onde podem ser observadas algumas partes do circuito. Nota-se que quase toda a área disponível foi ocupada, tendo-se inclusive de desenhar um *padframe* específico para o circuito, que ocupasse pouca área da periferia da pastilha. O circuito completo, sem o *padframe*, ocupa cerca de $4000\lambda \times 4250\lambda$, consumindo uma área aproximada de $1000\mu\text{m} \times 1050\mu\text{m}$ em *TSMC 0.35* ($\lambda = 0,25\mu\text{m}$) e de $400\mu\text{m} \times 425\mu\text{m}$ em *TSMC 0.18* ($\lambda = 0,10\mu\text{m}$).

O circuito foi encapsulado no formato *DIP28*, cerâmico, e a tabela 4.4 apresenta a descrição dos pinos utilizados. Em relação à pastilha apresentada na fig. 4.18, o pino (1)

corresponde ao *pad* central da coluna vertical da direita (ao lado do texto “chaves e referências N e P”), e o ordenamento da numeração é feito no sentido anti-horário.

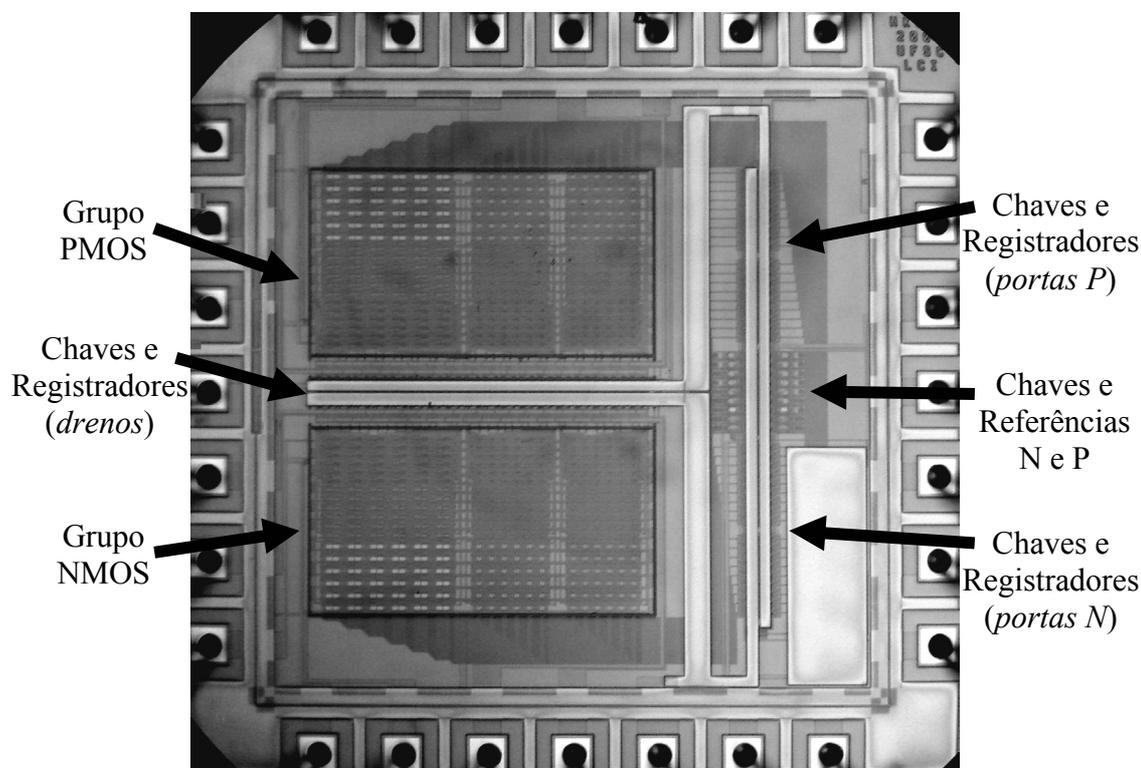


Figura 4.18: Microfotografia da segunda versão do circuito de teste, implementado na tecnologia *TSMC 0.35*, em uma pastilha com 1,5 mm de lado. Cada conjunto NMOS e PMOS é formado por 9 grupos de geometrias diferentes, cada um composto por 36 pares de transistores idênticos.

Tabela 4.4: Número do pino de acesso, nome e significado, referentes ao encapsulamento *DIP28* utilizado. A letra, entre parênteses, informa se o pino deve ser utilizado para estímulo (E) ou medida (M).

Pino	Nome	Significado
1	DRP	<i>Dreno</i> da referência PMOS, onde se aplica a corrente de polarização (E)
2	SRP	<i>Fonte</i> da referência PMOS, onde se aplica a tensão de <i>fonte</i> deste transistor (E)
3	E	Emissor do transistor bipolar PNP parasita vertical, implementado adicionalmente (o coletor está em GND) (E)
4	B	Base do transistor bipolar PNP parasita vertical (E)
5	GRP	É a tensão produzida na <i>porta</i> do transistor de referência PMOS polarizado em corrente, e que é aplicada à <i>porta</i> dos transistores de teste

		selecionados (M)
6	VCC	Alimentação positiva (E)
7	DSPA	Ligação <i>sense</i> do <i>dreno</i> do conjunto PMOS, transistor A do par (M)
8	DFPA	Ligação <i>force</i> do <i>dreno</i> do conjunto PMOS, transistor A do par (E)
9	SFP	Ligação <i>force</i> da <i>fonte</i> do conjunto PMOS (E)
10	DIOP	Diodo implementado entre este pino (anodo) e GND (catodo), próximo ao transistor de maior razão de aspecto PMOS; pode ser utilizado para se avaliar a variação térmica da pastilha durante sua caracterização (E)
11	SSP	Ligação <i>sense</i> da <i>fonte</i> do conjunto PMOS (M)
12	DSPB	Ligação <i>sense</i> do <i>dreno</i> do conjunto PMOS, transistor B do par (M)
13	DFPB	Ligação <i>force</i> do <i>dreno</i> do conjunto PMOS, transistor B do par (E)
14	CLK	Entrada de sinal de <i>clock</i> do registrador (E)
15	DIN	Entrada de sinal de <i>dados</i> do registrador (E)
16	DFNB	Ligação <i>force</i> do <i>dreno</i> do conjunto NMOS, transistor B do par (E)
17	DSNB	Ligação <i>sense</i> do <i>dreno</i> do conjunto NMOS, transistor B do par (M)
18	SFN	Ligação <i>force</i> da <i>fonte</i> do conjunto NMOS (E)
19	DION	Diodo implementado entre este pino (anodo) e GND (catodo), próximo ao transistor de maior razão de aspecto NMOS (E)
20	SSN	Ligação <i>sense</i> da <i>fonte</i> do conjunto NMOS (M)
21	DFNA	Ligação <i>force</i> do <i>dreno</i> do conjunto NMOS, transistor A do par (E)
22	DSNA	Ligação <i>sense</i> do <i>dreno</i> do conjunto NMOS, transistor A do par (M)
23	GND	Terra da alimentação (E)
24	DRN	<i>Dreno</i> da referência NMOS, onde se aplica a corrente de polarização (E)
25	SRN	<i>Fonte</i> da referência NMOS, onde se aplica a tensão de <i>fonte</i> deste transistor (E)
26	GRN	É a tensão produzida na <i>porta</i> do transistor de referência NMOS polarizado em corrente, e que é aplicada à <i>porta</i> dos transistores de teste selecionados (M)
27	DOUT	Saída de sinal de <i>dados</i> do registrador (M)
28	DIORN	Diodo implementado entre este pino (anodo) e GND (catodo), próximo ao transistor de maior razão de aspecto da referência NMOS (E)

A tabela 4.5 apresenta alguns parâmetros das duas tecnologias empregadas, resultantes dos lotes em que o circuito de teste foi fabricado, e obtidos no site *www.mosis.org*. A corrente específica por quadrado - I_{SQ} - foi calculada através dos parâmetros da tecnologia, utilizando-se $n = 1,3$, valor que tem se mostrado uma boa aproximação prática.

Tabela 4.5: Principais parâmetros das tecnologias *TSMC 0.35* (rodada T52F MM_EPI) e *TSMC 0.18* (rodada T55U MM_NON-EPI), utilizadas na fabricação do circuito.

	tipo	μ_0 [cm ² /V.s]	C'_{ox} [fF/μm ²]	V_{T0} [V] (trans. curto)	V_{T0} [V] (trans. grande)	I_{SQ} [nA]
TSMC 0.35 (T52F)	NMOS	416,53	4,51	0,59	0,52	82,5
	PMOS	141,86	4,44	-0,75	-0,75	27,7
TSMC 0.18 (T55U)	NMOS	412,01	8,47	0,52	0,43	153
	PMOS	85,96	8,27	-0,52	-0,41	31,2

Para a caracterização elétrica CC do descasamento, entre os transistores dos pares que compõem cada grupo, realizou-se a medida $I_D \times V_{DS}$ de todos os 36 pares, sob os níveis de inversão 0,01; 0,1; 1; 10; 100 e 1000, e com V_{DS} variando logarithmicamente entre 20 mV e 2 V para a tecnologia *TSMC 0.35* e entre 18 mV e 1,8 V para a tecnologia *TSMC 0.18*. Essa faixa de tensões *dreno-fonte* é suficiente para cobrir desde a região linear até a saturação em ambas tecnologias. Em cada condição de polarização, a corrente de dreno de cada um dos dois transistores de um par é medida duas vezes, devido ao processo de reversão de canais (UEMs), que tem por objetivo a redução do impacto do descasamento existente entre as UEMs, sobre as medidas. Além disso, o equipamento de caracterização realiza várias medidas do mesmo ponto, integrando-as como forma de reduzir o “ruído” presente no resultado.

A fig. 4.19 apresenta o resultado obtido neste processo para os 36 pares de uma das geometrias. Este tipo de caracterização foi realizado em todos os 9 grupos NMOS e PMOS das duas tecnologias em que o circuito de teste foi fabricado, e serviu como base para a geração das curvas de descasamento apresentadas. No gráfico relativo a $I_D \times V_{DS}$ cada marca (x) representa a corrente medida em cada um dos transistores dos 36 pares. Observa-se que, na medida em que os transistores são submetidos a níveis de inversão mais baixos (curvas inferiores), o valor da corrente medida para os 72 transistores, sob determinado

V_{DS} , apresenta maior dispersão em torno de seu valor médio (visualmente, observa-se que as marcas “x” não coincidem, estando mais dispersas verticalmente na medida em que o nível de inversão diminui). As linhas sólidas correspondem ao modelo ACM, com I_{SQ} extraído experimentalmente da tecnologia.

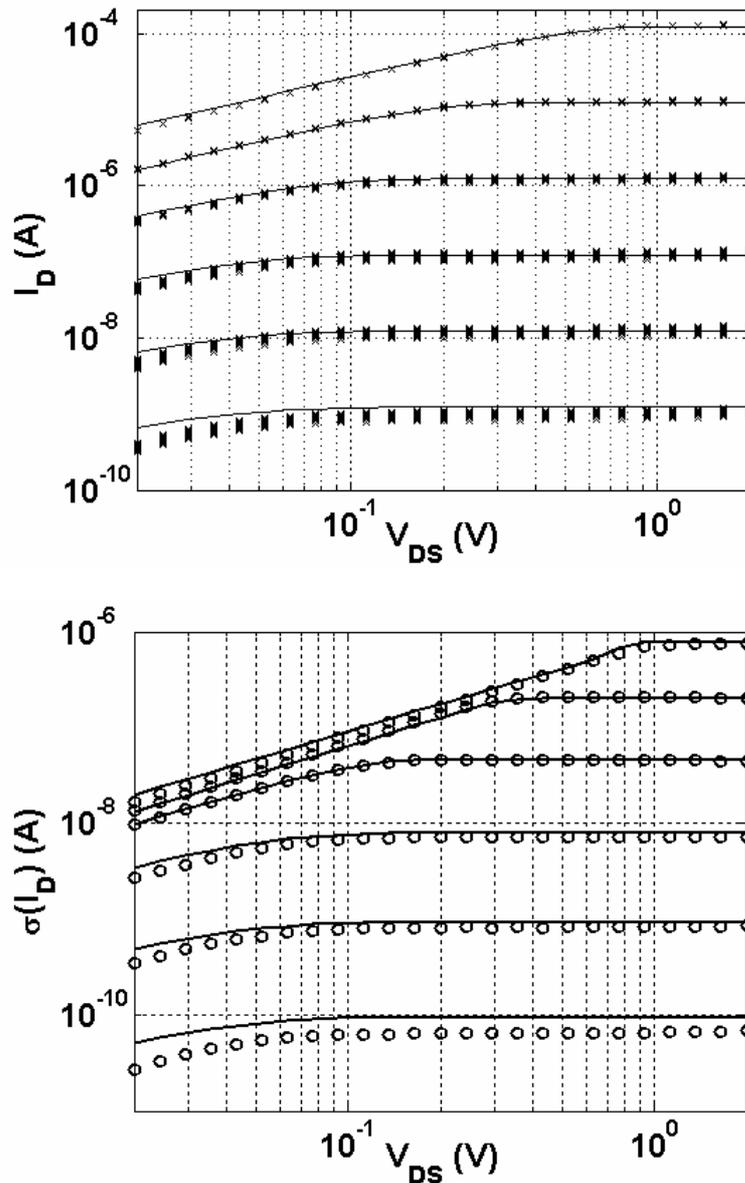


Figura 4.19: Medidas $I_D \times V_{DS}$ (x - cima) e $\sigma(I_D) \times V_{DS}$ (o - baixo) de um grupo NMOS 04x04 na tecnologia TSMC 0.35. As medidas foram realizadas com V_{DS} variando entre 20 mV e 2 V, para os níveis de inversão 0,01; 0,1; 1; 10; 100 e 1000 (de baixo para cima). As linhas contínuas representam a estimativa teórica de I_D , calculada através do modelo ACM, e do descasamento médio absoluto $\sigma(I_D)$, calculado através do modelo aqui apresentado (os parâmetros de modelo I_{SQ} , N_{oi} e B_{ISQ} foram extraídos experimentalmente da tecnologia)

No gráfico relativo a $\sigma(I_D) \times V_{DS}$ da fig. 4.19, observa-se que o descasamento absoluto médio também aumenta com o nível de inversão. Observa-se ainda que o “espaçamento” entre as curvas diminui na medida que se aumenta o nível de inversão, principalmente na região linear, fazendo com que o descasamento relativo $\sigma(I_D)/I_D$ também diminua com o aumento do nível de inversão. Cada marca (o) representa o resultado dos 36 pares medidos. As linhas sólidas correspondem ao modelo de descasamento proposto, incluindo o efeito de saturação na velocidade dos portadores, e com N_{oi} e B_{ISQ} extraídos experimentalmente.

A caracterização automatizada completa de um conjunto NMOS ou PMOS de uma única amostra do circuito consome cerca de 10 a 12 horas, gerando um arquivo texto com cerca de 550 KB de dados. O tempo necessário para este processo resulta da grande quantidade de dispositivos medidos sob muitas condições de polarização. Como o equipamento de caracterização despende mais tempo na medida de correntes baixas, a caracterização dos grupos com menor I_S é mais demorada (o valor de I_S diminui com a razão de aspecto W/L e com o valor de I_{SQ} do transistor caracterizado). No total, para cada gráfico de descasamento gerado para uma geometria, são medidos 72 transistores, cada um por 2 vezes, sob 6 níveis de inversão, e sob 25 valores de V_{DS} , resultando em 21.600 medidas. Além das correntes de dreno, diversas outras grandezas são medidas e armazenadas durante a caracterização, como forma de se ter maior controle sobre a confiabilidade do processo.

Os testes feitos utilizando-se as junções de diodos implementadas, não apontaram uma variação significativa na temperatura próxima aos transistores com maior W/L .

As fig. 4.20 a 4.22 apresentam o descasamento resultante da caracterização de todas as geometrias NMOS e PMOS na tecnologia *TSMC 0.35*. Diferentes pastilhas (amostras) apresentaram resultados semelhantes para uma mesma geometria. Foram completamente caracterizadas 10 amostras desta tecnologia.

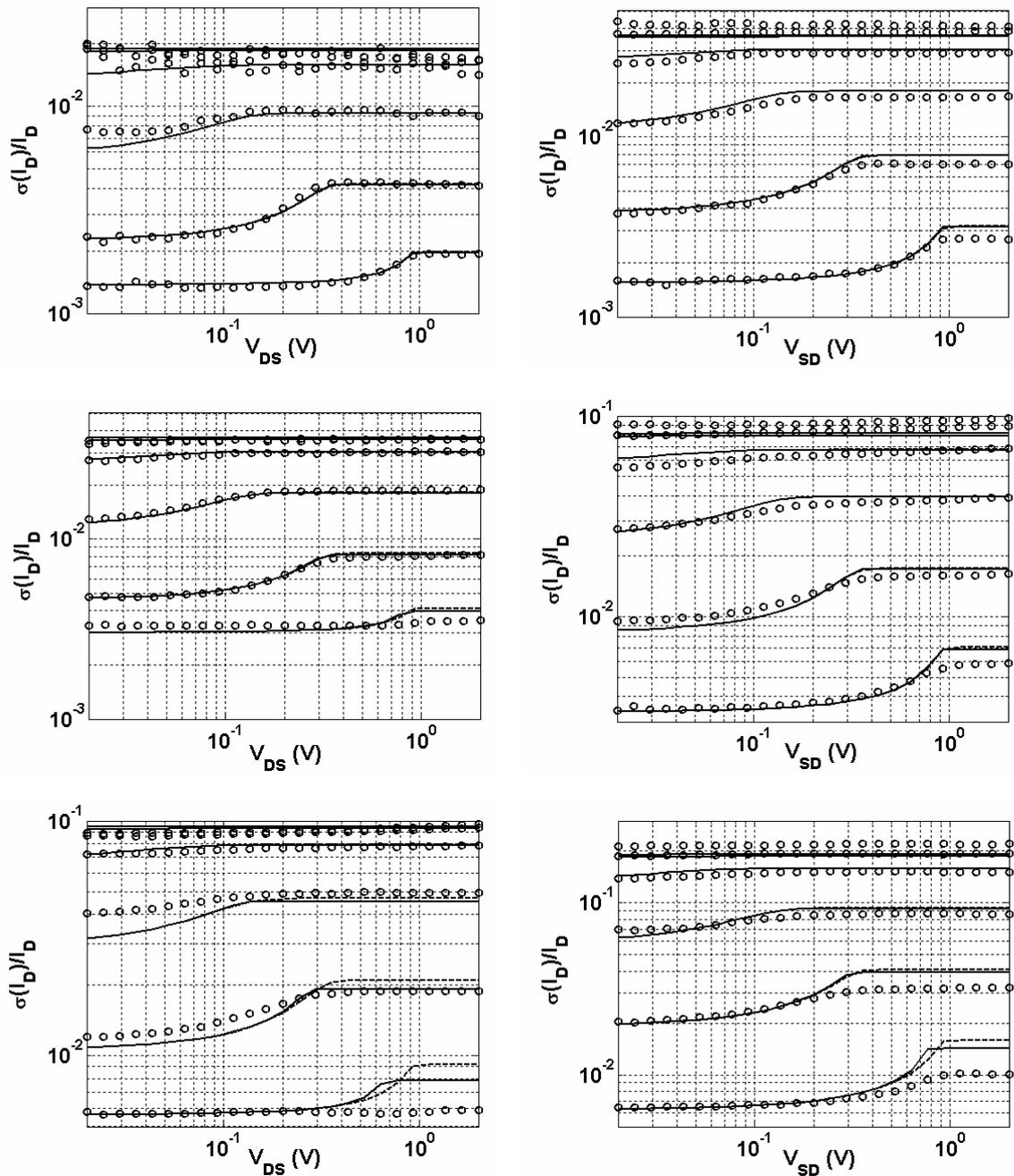


Figura 4.20: Descasamento dos grupos 16x16 (12 μm / 8 μm - cima), 16x04 (12 μm / 2 μm - centro) e 16x01 (12 μm / 0,5 μm - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia TSMC 0.35. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (--): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000.

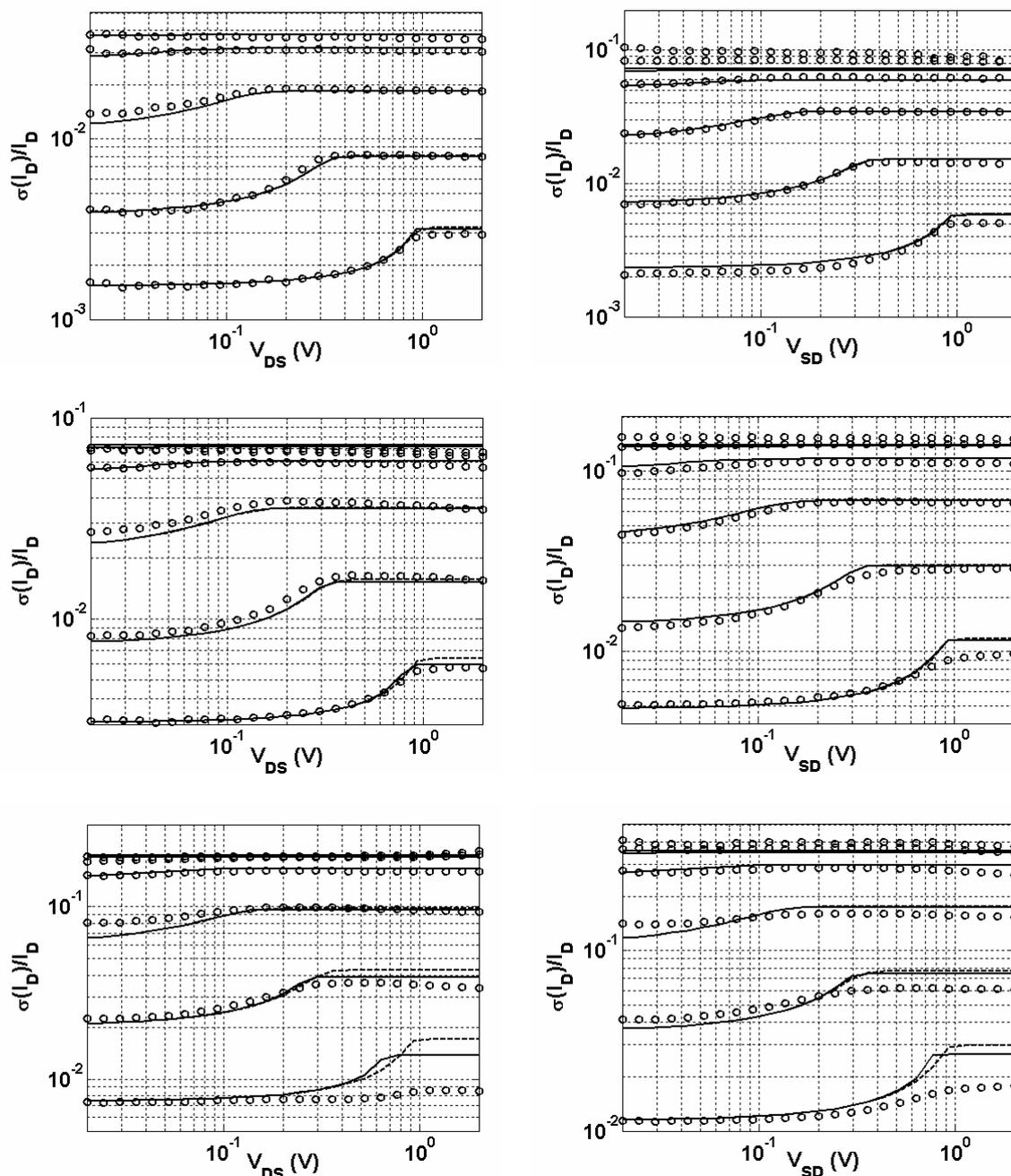


Figura 4.21: Descasamento dos grupos 04x16 (3 μ m / 8 μ m - cima), 04x04 (3 μ m / 2 μ m - centro) e 04x01 (3 μ m / 0,5 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia TSMC 0.35. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (- -): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. No grupo 04x16 NMOS o nível de inversão 0,01 não está representado.

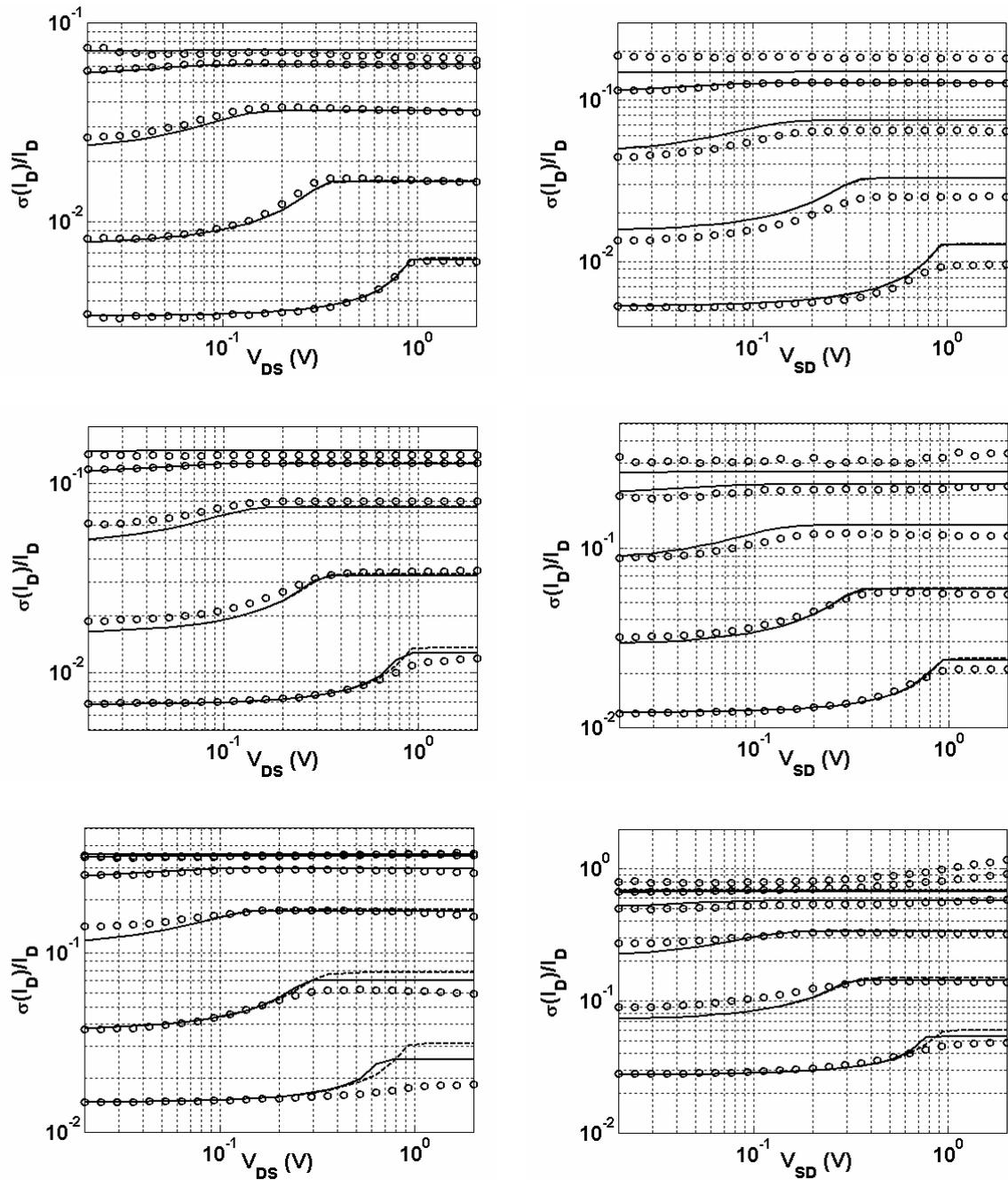


Figura 4.22: Descasamento dos grupos 01x16 (0,75 μ m / 8 μ m - cima), 01x04 (0,75 μ m / 2 μ m - centro) e 01x01 (0,75 μ m / 0,5 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.35*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. Nos grupos 01x16 e 01x04 o nível de inversão 0,01 não está representado.

O valor experimental do descasamento, sob cada condição de polarização, é obtido através da seguinte expressão

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{1}{2(NP-1)I_D^2} \sum_{i=1}^{NP} (I_{DA(i)} - I_{DB(i)})^2, \quad (4.8.2)$$

onde $I_{DA(i)}$ e $I_{DB(i)}$ são as correntes de cada um dos transistores do par i , com $NP = 36$ representando o número de pares de um grupo. O valor destas correntes resulta da média dos valores obtidos de um transistor com cada um dos dois canais de caracterização, devido ao processo de reversão utilizado.

Em algumas geometrias foram omitidas as curvas de descasamento sob níveis de inversão muito baixos (0,01 e 0,1), pois a corrente medida resultante é muito pequena, de forma que o descasamento estimado está muito mascarado pela corrente de fuga e pelos erros intrínsecos das UEMs.

Tabela 4.6: Valores dos parâmetros do modelo de descasamento utilizados nas figuras 4.20 a 4.22. Foram calculados através de ajuste feito pelo método de mínimos quadrados, aplicado às curvas de descasamento obtidas experimentalmente da rodada T52F, utilizada na fabricação do circuito de teste na tecnologia *TSMC 0.35*.

TSMC 0.35		NMOS		PMOS	
Grupo	Dimensão W/L	N_{oi} [cm^{-2}]	B_{ISQ} [%- μm]	N_{oi} [cm^{-2}]	B_{ISQ} [%- μm]
16x16	12 μm / 8 μm	3,0x10 ¹²	1,2	11,4x10 ¹²	1,0
16x04	12 μm / 2 μm	2,6x10 ¹²	1,3	13,8x10 ¹²	1,1
16x01	12 μm / 0,5 μm	3,8x10 ¹²	1,1	17,4x10 ¹²	0,4
04x16	3 μm / 8 μm	3,1x10 ¹²	0,5	11,0x10 ¹²	0,3
04x04	3 μm / 2 μm	2,7x10 ¹²	0,5	10,2x10 ¹²	0,4
04x01	3 μm / 0,5 μm	4,2x10 ¹²	0,4	15,3x10 ¹²	0,1
01x16	0,75 μm / 8 μm	2,9x10 ¹²	0,6	12,1x10 ¹²	0,5
01x04	0,75 μm / 2 μm	2,9x10 ¹²	0,6	9,8x10 ¹²	1,0
01x01	0,75 μm / 0,5 μm	3,4x10 ¹²	0,5	14,1x10 ¹²	1,0

Os parâmetros N_{oi} e B_{ISQ} do modelo resultam do ajuste de suas curvas aos resultados experimentais, pelo método dos mínimos quadrados. São apresentadas as curvas do modelo com e sem o efeito da saturação na velocidade dos portadores (ESVP). O valor da velocidade de saturação (v_{sat}) é essencialmente independente da concentração de dopantes, sendo da ordem de 10^7 cm/s para elétrons e lacunas em temperatura ambiente

[67]. Para o cálculo deste efeito no modelo, utilizou-se os valores de mobilidade a baixo campo fornecidos pela *MOSIS*. A tabela 4.6 apresenta os valores dos parâmetros empregados em cada figura apresentada.

Na medida em que o canal dos transistores diminui, pode-se notar que o modelo diverge dos resultados experimentais, principalmente na condição de saturação, quando sob níveis de inversão altos. Isso ocorre porque o único efeito de canal-curto incluído no modelo foi o efeito da saturação de velocidade de portadores (ESVP). Outros, como a modulação de comprimento do canal, ou a deformação na distribuição de potenciais no canal, induzida pelo potencial *dreno-fonte*, não foram incluídos. E, mesmo o ESVP, foi incluído através de um modelo empírico muito simples.

Também é muito provável que as flutuações (essencialmente devido à dopagem no canal) próximas ao *dreno* e à *fonte* sejam mais intensas que na região central do canal, devido à maior concentração de dopantes dessas regiões. Quando o transistor satura em inversão muito forte, a maior penetração da zona de depleção no lado do *dreno* (em saturação), reduz a influência da região mais dopada próxima ao *dreno*, em relação à dopagem média do canal, devendo contribuir para atenuar o aumento do descasamento.

A flutuação na tensão de limiar (V_T) de um *MOSFET*, decorrente do somatório dos efeitos das flutuações microscópicas de carga que ocorrem na região de depleção, pode ser estimada por (vide o *anexo B*)

$$\sigma_{V_T}^2 = \left(\frac{q}{C'_{ox}} \right)^2 \frac{1}{WL} N_{oi} \quad (4.8.3)$$

Pelo modelo de Pelgrom, a variação na tensão de limiar é dada por

$$\sigma_{V_T}^2 = \frac{A_{VT}^2}{WL}, \quad (4.8.4)$$

onde A_{VT} é o parâmetro de descasamento desta tensão, dado em mV- μm .

Assim, através de (4.8.3) e (4.8.4), pode-se estabelecer a seguinte relação entre o modelo de Pelgrom e o aqui apresentado

$$A_{VT} = \frac{q}{C'_{ox}} \sqrt{N_{oi}} \quad (4.8.5)$$

Considerando-se o valor médio obtido experimentalmente na tecnologia *TSMC 0.35* (tabela 4.6), para o parâmetro NMOS do modelo, $N_{oiN} = 3,1 \times 10^{12} \text{ cm}^{-2}$, pode-se estimar com (4.8.5) o valor do equivalente do parâmetro do modelo de Pelgrom como $A_{VTN} = 6,3 \text{ mV-}\mu\text{m}$. Para o caso PMOS, o valor médio é $N_{oiP} = 12,7 \times 10^{12} \text{ cm}^{-2}$, resultando no valor do equivalente do parâmetro do modelo de Pelgrom como $A_{VTP} = 12,8 \text{ mV-}\mu\text{m}$.

Os valores destes parâmetros não estão disponíveis nas informações técnicas fornecidas pela *MOSIS*. Entretanto, através da literatura técnica do processo *CMOS AMS 0,35 μm C35 (Austria Mikro Systeme International AG)*, que possui similaridade com o utilizado, obteve-se os parâmetros de descasamento apresentados na tabela 4.7 para três períodos diferentes do processo. Os valores apresentados foram obtidos dividindo-se por $\sqrt{2}$ os valores originais, fornecidos pela *AMS*, pois originalmente representam o descasamento médio entre dois dispositivos, e não de um único, como é o caso dos resultados experimentais aqui apresentados para o modelo. Pode-se observar que há proximidade entre os resultados extraídos para este modelo, da tecnologia *TSMC 0.35*, e os apresentados na tabela. Uma possível fonte de diferenças está relacionada às divergências tecnológicas entre os dois processos comparados.

Tabela 4.7: Valores dos parâmetros de descasamento, referentes ao modelo de Pelgrom, obtidos através de três versões da literatura técnica do processo *CMOS AMS 0,35 μm* .

Data da versão	$A_{VTN} [\text{mV-}\mu\text{m}]$	$A_{\beta N} [\%-\mu\text{m}]$	$A_{VTP} [\text{mV-}\mu\text{m}]$	$A_{\beta P} [\%-\mu\text{m}]$
23.01.2001	6,8	0,8	11,3	0,8
13.03.2003	5,8	0,1	10,5	0,4
08.05.2006	6,7	0,5	10,3	0,7

O parâmetro B_{ISQ} é equivalente ao A_{β} do modelo de Pelgrom. Os valores apresentados para a tecnologia *AMS 0.35*, para os NMOS e os PMOS, são próximos dos encontrados através dos resultados experimentais e apresentado na tabela 4.6.

Nota-se através das figuras que as medidas realizadas sobre os grupos PMOS apresentam mais imprecisão que as dos NMOS. A polarização dos NMOS é sempre feita entre $V_{CC} = +3,3 \text{ V}$ e $GND = 0$, sendo que os valores dos potenciais aplicados são referenciados ao GND . A polarização dos PMOS pode ser feita de duas maneiras: alimenta-se o circuito da forma tradicional, entre $V_{CC} = +3,3 \text{ V}$ e $GND = 0 \text{ V}$ (terra); ou

alimenta-se o circuito entre $V_{CC} = 0$ V e $GND = -3,3$ V. Nos dois casos, os valores dos potenciais aplicados são referenciados ao valor de V_{CC} . Provavelmente estes procedimentos provoquem um maior acúmulo de erros e suscetibilidade a ruídos nos potenciais gerados, em virtude de limitações do próprio equipamento de caracterização, pois os valores das grandezas geradas neste são sempre referenciados ao GND do equipamento.

As fig. 4.23 a 4.25 e a tabela 4.8 correspondem à tecnologia *TSMC 0.18*. Foram completamente caracterizadas 7 amostras desta tecnologia.

Através da comparação das curvas experimentais obtidas nas duas tecnologias, nota-se que os resultados obtidos na *TSMC 0.18* apresentam mais flutuações e erros, mesmo com esta tecnologia operando com maiores correntes para uma mesma condição de polarização (devido ao maior I_S , cerca do dobro no caso N e pouco acima no caso P), o que deveria facilitar as medidas.

Pode-se supor duas possibilidades para estes erros: a maior suscetibilidade da corrente a variações de V_{GS} nesta tecnologia (devido à menor tensão de limiar); e os maiores efeitos térmicos, pois nesta tecnologia os dispositivos são menores, resultando em maiores densidades de correntes, e operam com maiores correntes para um determinado nível de inversão, devido aos maiores valores de I_S . Além disso, nesta tecnologia as quedas ôhmicas nos contatos são maiores, pois a área destes diminuiu, devido ao escalamento do circuito. Também, devido ao escalamento, aumentou o comprimento das trilhas de conexão do circuito com os *pads*, pois o tamanho da pastilha se manteve o mesmo da tecnologia anterior. E os contatos e conexões estão submetidos a correntes maiores que na outra tecnologia.

Também, deve-se destacar que em tecnologias com comprimento de canal abaixo de aproximadamente $0,25$ μm , é comum o uso de implantes tipo “pocket” (ou “halo”), que são implantes com concentração diferenciada localizados nos extremos da região ativa, próximo ao *dreno* e à *fonte*, e têm por objetivo reduzir os efeitos de canal-curto que se agravam com a diminuição das dimensões do transistor. Nesse caso, a modelagem de um canal com concentração de dopantes homogênea como feita aqui, resulta em mais um fator de divergência entre o modelo e os resultados experimentais do descasamento.

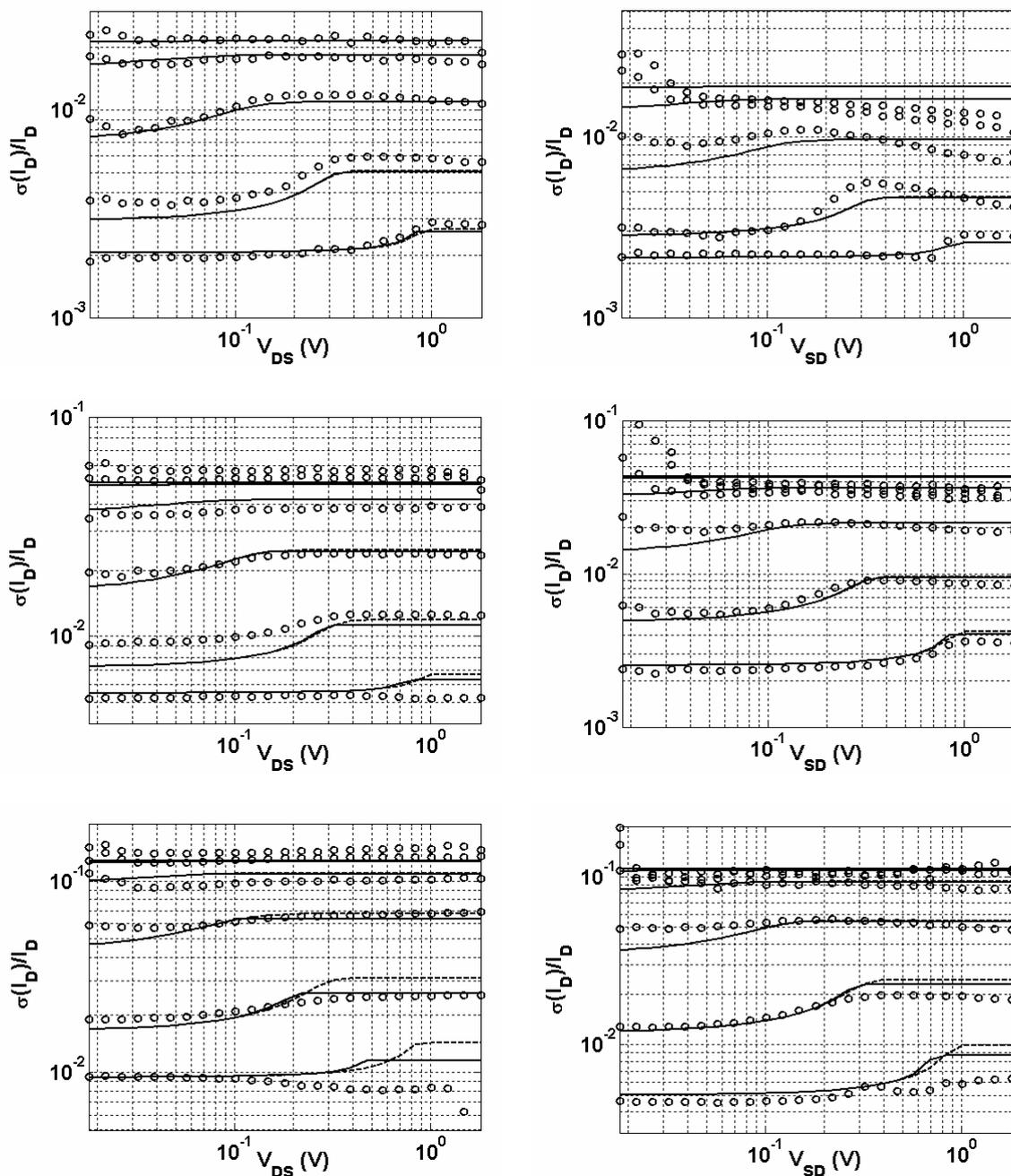


Figura 4.23: Descasamento dos grupos 16x16 (4,8 μm / 3,2 μm - cima), 16x04 (4,8 μm / 0,8 μm - centro) e 16x01 (4,8 μm / 0,2 μm - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.18*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. No grupo 16x16 o nível de inversão 0,01 não está representado.

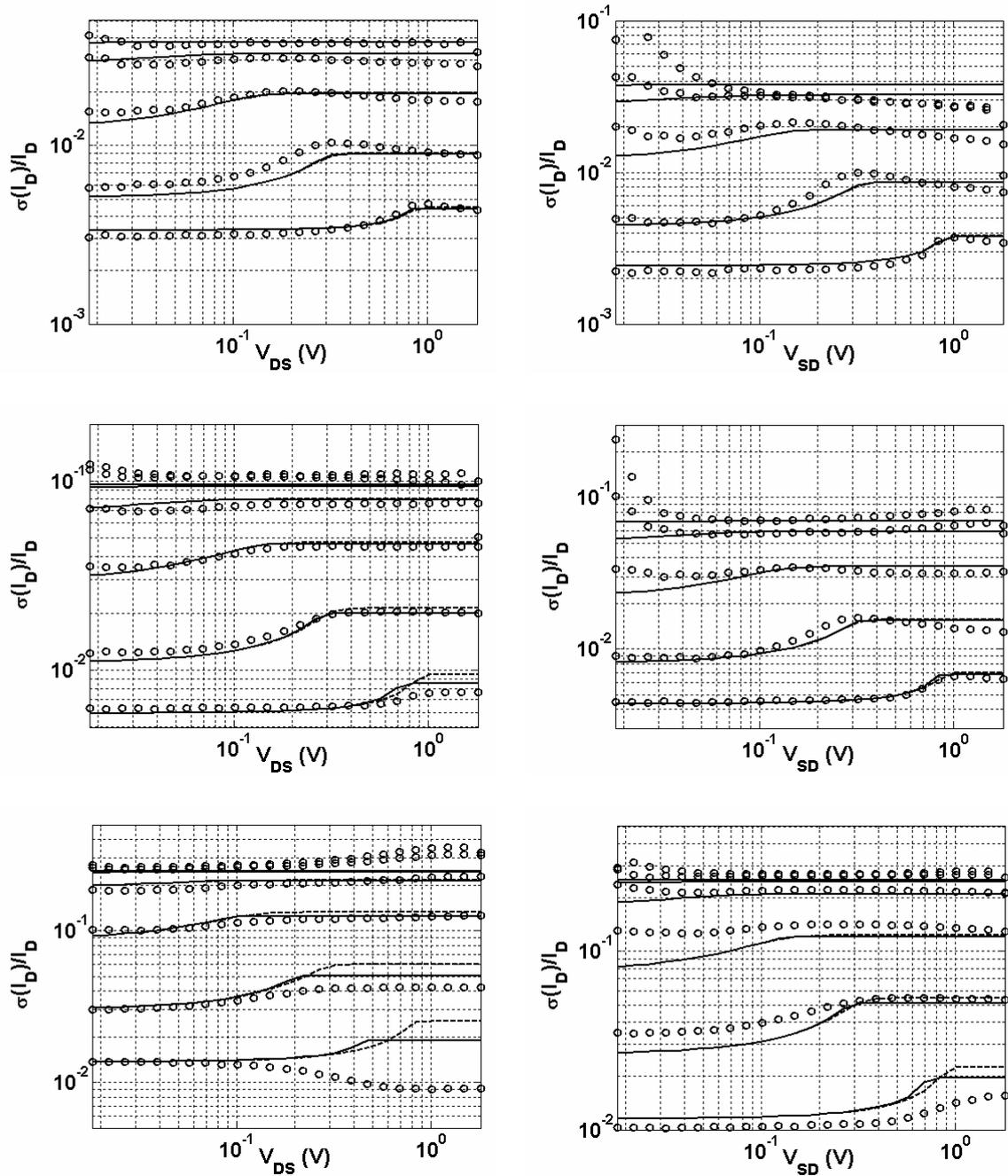


Figura 4.24: Descasamento dos grupos 04x16 (1,2 μm / 3,2 μm - cima), 04x04 (1,2 μm / 0,8 μm - centro) e 04x01 (1,2 μm / 0,2 μm - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.18*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. No grupo 04x16 e no 04x04 PMOS o nível de inversão 0,01 não está representado.

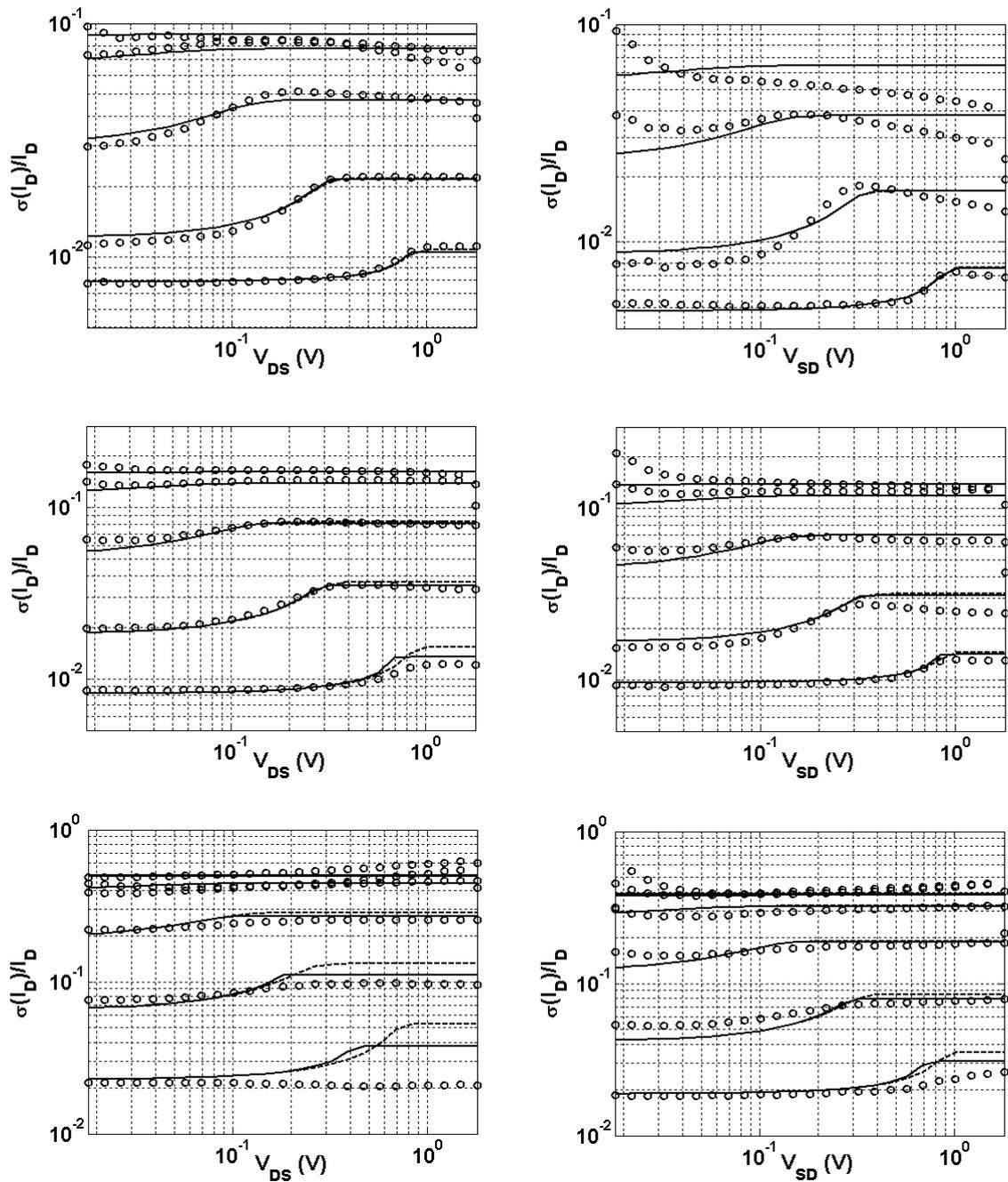


Figura 4.25: Descasamento dos grupos 01x16 (0,3 μ m / 3,2 μ m - cima), 01x04 (0,3 μ m / 0,8 μ m - centro) e 01x01 (0,3 μ m / 0,2 μ m - baixo), NMOS (esq.) e PMOS (dir.) na tecnologia *TSMC 0.18*. Círculos (o): medidas; linhas sólidas (—): modelo com ESVP; linhas segmentadas (---): modelo sem ESVP. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000. Nos grupos 01x16 e 01x04 o nível de inversão 0,01 não está representado. No grupo 01x16 PMOS o nível de inversão 0,1 também não está representado.

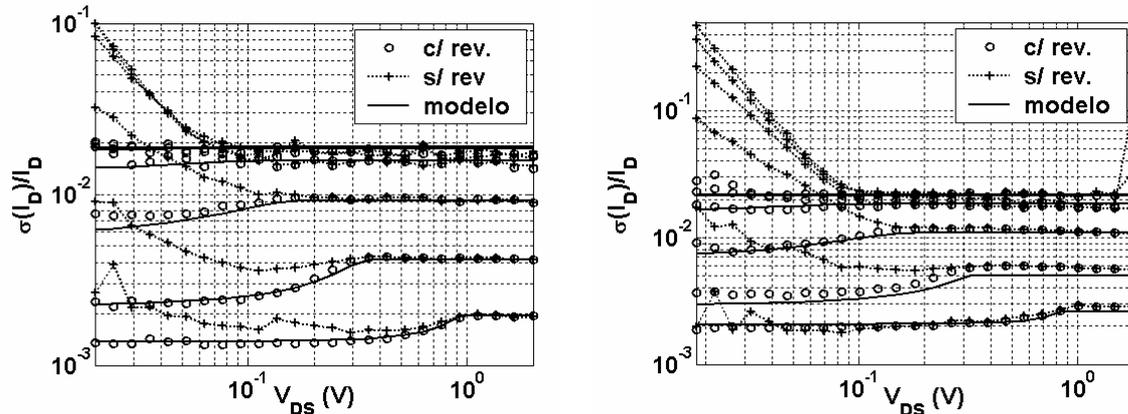


Figura 4.26: Descasamento medido dos grupos NMOS 16x16 em *TSMC 0.35* (esq.) e *TSMC 0.18* (dir.), com (o) e sem (+) o uso da reversão dos canais. As curvas representam respectivamente, de cima para baixo, os níveis de inversão: 0,01; 0,1; 1; 10; 100 e 1000.

Tabela 4.8: Valores dos parâmetros do modelo de descasamento utilizados nas figuras 4.23 a 4.25. Foram calculados através de ajuste feito pelo método de mínimos quadrados, aplicado às curvas de descasamento obtidas experimentalmente da rodada T55U, utilizada na fabricação do circuito de teste na tecnologia *TSMC 0.18*.

<i>TSMC 0.18</i>		NMOS		PMOS	
Grupo	Dimensão W/L	N_{oi} [cm^{-2}]	B_{ISQ} [%- μm]	N_{oi} [cm^{-2}]	B_{ISQ} [%- μm]
16x16	4,8 μm / 3,2 μm	0,6x10 ¹²	0,8	0,5x10 ¹²	0,8
16x04	4,8 μm / 0,8 μm	0,8x10 ¹²	1,0	0,6x10 ¹²	0,4
16x01	4,8 μm / 0,2 μm	1,0x10 ¹²	0,7	0,8x10 ¹²	0,3
04x16	1,2 μm / 3,2 μm	0,5x10 ¹²	0,6	0,5x10 ¹²	0,4
04x04	1,2 μm / 0,8 μm	0,7x10 ¹²	0,5	0,4x10 ¹²	0,4
04x01	1,2 μm / 0,2 μm	1,0x10 ¹²	0,4	1,0x10 ¹²	0,4
01x16	0,3 μm / 3,2 μm	0,8x10 ¹²	0,7	0,5x10 ¹²	0,4
01x04	0,3 μm / 0,8 μm	0,5x10 ¹²	0,3	0,4x10 ¹²	0,4
01x01	0,3 μm / 0,2 μm	1,0x10 ¹²	0,2	0,6x10 ¹²	0,3

A fig. 4.26 apresenta o descasamento do grupo NMOS 16x16 nas duas tecnologias, medido *com* e *sem* o processo de reversão dos canais utilizados para estímulo e medida dos terminais de *dreno*. Observa-se que, sem o uso desse processo, o descasamento entre as UEMs afeta os resultados principalmente na região linear, devido à maior suscetibilidade

da corrente de dreno às variações de tensão. O impacto também é maior nos menores níveis de inversão, devido às baixas correntes envolvidas.

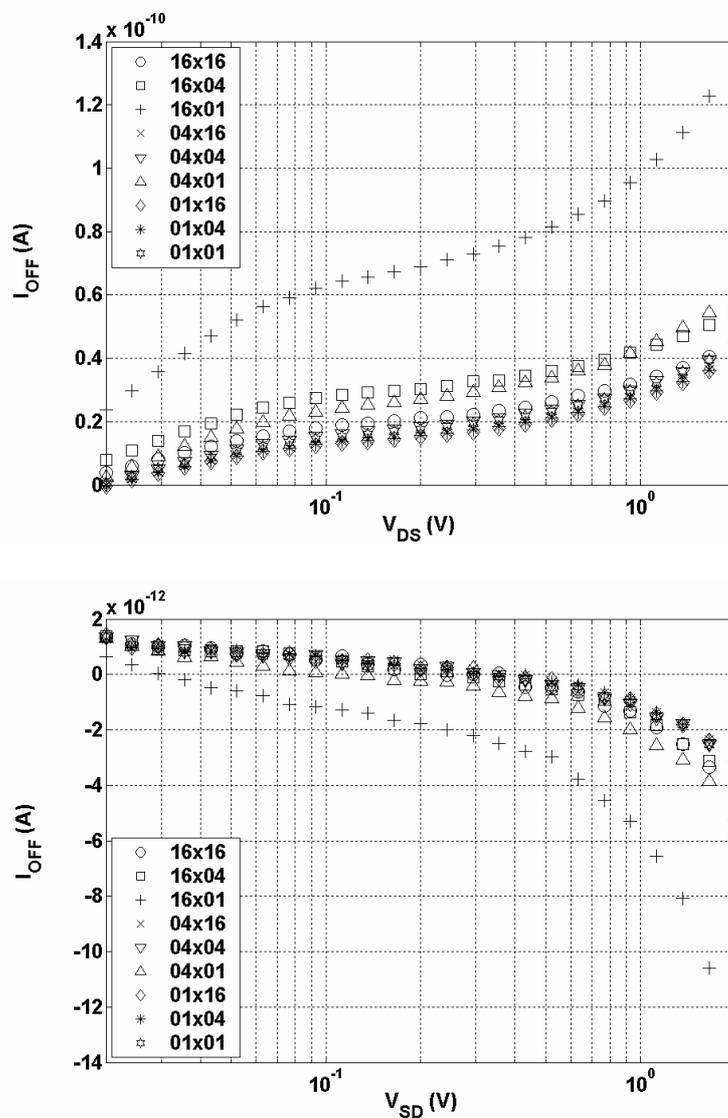


Figura 4.27: Corrente de “fuga” I_{OFF} , obtida de cada uma das nove geometrias com todos os 36 pares de transistores desativados (*off state*) através de $V_{GS} = 0V$. Resultados da tecnologia *TSMC 0.35* para o conjunto NMOS (cima) e PMOS (baixo).

As fig. 4.27 e 4.28 apresentam o comportamento das correntes de fuga (I_{OFF}) de cada grupo geométrico NMOS e PMOS, para as duas tecnologias, quando todos os seus transistores estão desligados ($V_{GS} = 0V$). Observa-se que, via de regra, esta corrente é maior para geometrias de maior razão de aspecto. Os valores na tecnologia *TSMC 0.18* são bem maiores que na *TSMC 0.35*, principalmente no caso dos PMOS. Supõe-se que isto seja devido ao valor da tensão de limiar, que é ajustado para ser menor na *TSMC 0.18*, sendo ainda proporcionalmente menor no caso dos PMOS.

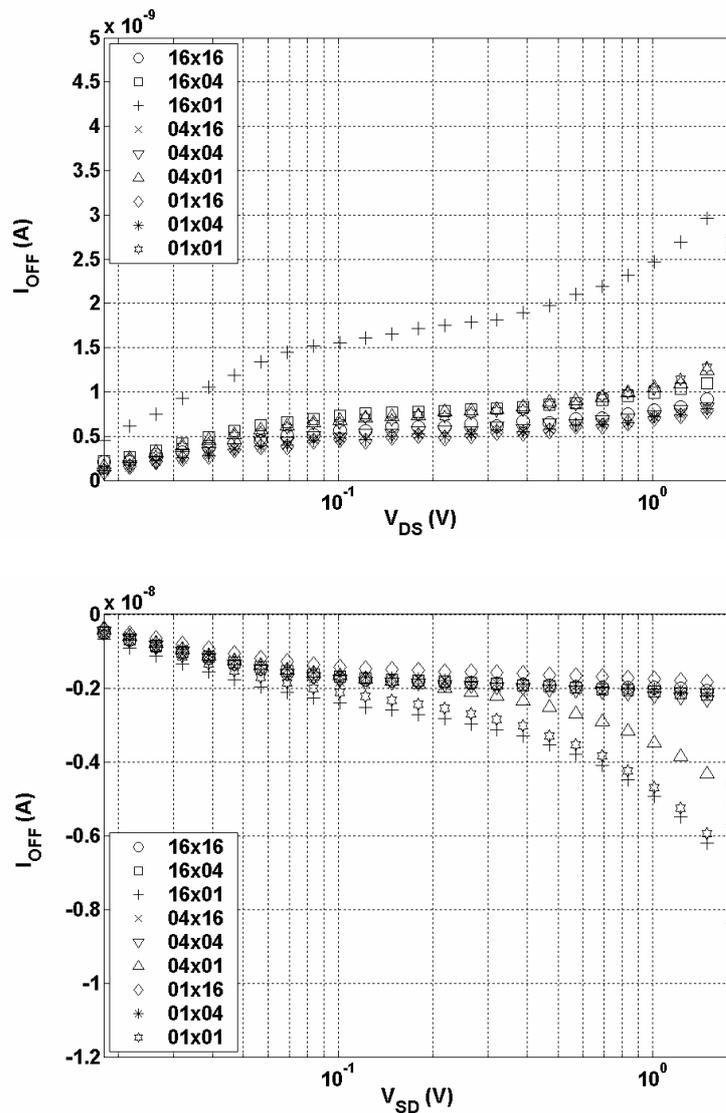


Figura 4.28: Corrente de “fuga” I_{OFF} , obtida de cada uma das nove geometrias com todos os 36 pares de transistores desativados (*off state*) através de $V_{GS} = 0V$. Resultados da tecnologia *TSMC 0.18* para o conjunto NMOS (cima) e PMOS (baixo).

O valor mais elevado da corrente I_{OFF} na tecnologia *TSMC 0.18* é um dos limitantes que dificulta a medida do descasamento sob níveis de inversão muito baixos, nos grupos de menor razão de aspecto (menor corrente absoluta).

Deve-se também lembrar que parte da corrente I_{OFF} medida é devido às chaves de seleção de dreno que, mesmo desligadas, permanecem fisicamente atreladas aos canais de medição.

A fig. 4.29 apresenta a superposição das curvas da corrente de dreno I_D , medidas de todos os transistores de um grupo geométrico, em função de V_{DS} , que varia entre 10 mV e

100 mV, mantendo os transistores predominantemente fora de saturação. Tal estratégia visa reduzir os efeitos da saturação na velocidade dos portadores sobre a corrente. A corrente é medida sob os níveis de inversão aproximados: 1, 10, 100 e 1000. Através de análise numérica aplicada sobre estas curvas, e com o uso das seguintes equações do modelo *ACM* canal-longo,

$$\frac{V_{DS}}{\phi_t} = \sqrt{1+i_f} + \ln(\sqrt{1+i_f} - 1) - [\sqrt{1+i_r} + \ln(\sqrt{1+i_r} - 1)], \quad (4.8.6)$$

$$I_D = (i_f - i_r) I_S = (i_f - i_r) \frac{W_{eff}}{L_{eff}} I_{SQ}, \quad (4.8.7)$$

estimou-se o valor da corrente específica (I_S) de cada geometria. Com o valor de I_S e as dimensões geométricas de cada grupo (L e W), calculou-se pelo método dos mínimos quadrados uma aproximação para a corrente específica por quadrado (I_{SQ}) e para os fatores de correção de comprimento (dL) e de largura (dW) da tecnologia, onde o comprimento e a largura efetivos são aproximados por $L_{eff} = L - dL$ e $W_{eff} = W - dW$.

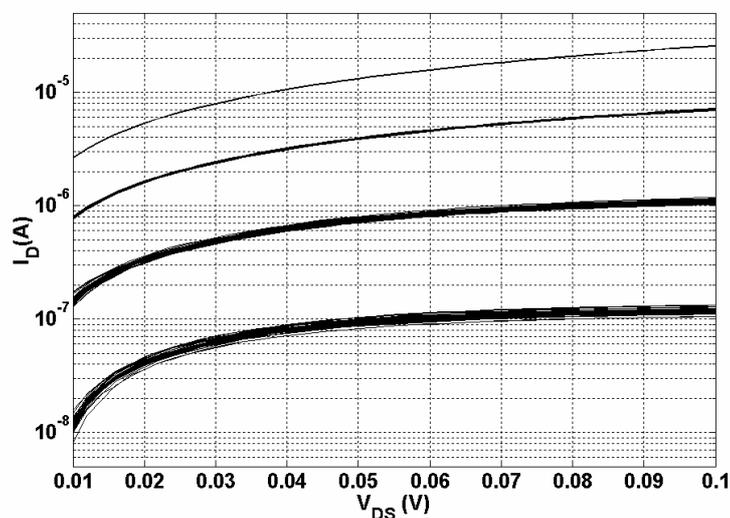


Figura 4.29: Curvas $I_D \times V_{DS}$ para os níveis de inversão 1, 10, 100 e 1000 (de baixo para cima), de todos os transistores do grupo NMOS 04x04, na tecnologia *TSMC 0.35*.

É importante aqui destacar que os valores de dL e dW obtidos por este método, representam de fato o somatório de vários efeitos (de borda, de deformação de campos elétricos, etc) que diferenciam o modelo idealizado do dispositivo real, tendo mais o sentido de parâmetros de ajuste, que de variação nas dimensões físicas. Isso se dá

principalmente para o caso da tecnologia *TSMC 0.18*, devido à concentração não uniforme de dopantes no canal (resultante dos implantes tipo “pocket” ou “halo”).

As tabelas 4.9 e 4.10 apresentam os resultados obtidos para estes parâmetros em cada uma das geometrias, nas duas tecnologias utilizadas. Os valores médios resultantes foram empregados na extração dos parâmetros do modelo de descasamento, apresentados nas tabelas 4.6 e 4.8 e utilizados nas figuras 4.20 a 4.25.

Tabela 4.9: Corrente específica (I_{SQ}) e fatores geométricos (dL e dW), obtidos através de ajuste por mínimos quadrados das curvas experimentais na tecnologia *TSMC 0.35*.

TSMC 0.35		NMOS			PMOS		
Grupo	Dimensão W/L	I_{SQ} [nA]	dW [μm]	dL [μm]	I_{SQ} [nA]	dW [μm]	dL [μm]
16x16	12 μm / 8 μm	86,9	0,13	0,07	27,2	0,16	0,05
16x04	12 μm / 2 μm	84,5	0,12	0,09	27,7	0,17	0,03
16x01	12 μm / 0,5 μm	84,6	0,11	0,09	27,2	0,16	0,05
04x16	3 μm / 8 μm	84,4	0,12	0,09	26,7	0,16	0,04
04x04	3 μm / 2 μm	82,1	0,10	0,10	28,2	0,18	0,03
04x01	3 μm / 0,5 μm	84,8	0,11	0,09	27,2	0,16	0,05
01x16	0,75 μm / 8 μm	84,7	0,11	0,09	27,2	0,16	0,05
01x04	0,75 μm / 2 μm	84,5	0,14	0,10	27,2	0,18	0,04
01x01	0,75 μm / 0,5 μm	83,8	0,10	0,10	26,9	0,15	0,07
Valores médios:		84,5	0,11	0,09	27,3	0,16	0,05

A fig. 4.30 apresenta o descasamento na corrente de *dreno* de transistores NMOS, sob um ponto-de-vista qualitativo, na tecnologia *TSMC 0.35*. Cada coluna sólida representa a corrente de dreno de um dos 36 que compõem um grupo geométrico (utilizou-se apenas um dos transistores de cada par neste ensaio). Todos os transistores de cada situação apresentada encontram-se exatamente sob a mesma condição de polarização. Percorrendo-se a figura na direção horizontal, observa-se o efeito das dimensões geométricas de cada grupo nas flutuações da corrente de dreno: os gráficos correspondem aos grupos 16x16 (a, d), 04x04 (b, e), e 01x01 (c, f). Percorrendo-se na vertical, observa-se o efeito do nível de inversão nestas flutuações: os gráficos correspondem a $i_f = 1$ (a, b, c) e $i_f = 100$ (d, e, f). Em todos os casos, os transistores encontram-se saturados com $V_{DS} = 1,1$ V.

Tabela 4.10: Corrente específica (I_{SQ}) e fatores geométricos (dL e dW), obtidos através de ajuste por mínimos quadrados das curvas experimentais na tecnologia *TSMC 0.18*.

<i>TSMC 0.18</i>		NMOS			PMOS		
Grupo	Dimensão W/L	I_{SQ} [nA]	dW [μm]	dL [μm]	I_{SQ} [nA]	dW [μm]	dL [μm]
16x16	4,8 μm / 3,2 μm	153	0,00	0,03	35,0	0,01	0,03
16x04	4,8 μm / 0,8 μm	134	0,01	0,05	32,1	0,06	0,02
16x01	4,8 μm / 0,2 μm	140	0,01	0,05	34,3	0,00	0,04
04x16	1,2 μm / 3,2 μm	153	0,00	0,03	31,7	0,01	0,05
04x04	1,2 μm / 0,8 μm	123	0,04	0,06	30,4	0,04	0,07
04x01	1,2 μm / 0,2 μm	139	0,02	0,04	36,2	0,00	0,05
01x16	0,3 μm / 3,2 μm	139	0,01	0,05	31,4	0,05	0,05
01x04	0,3 μm / 0,8 μm	139	0,00	0,05	31,0	0,02	0,09
01x01	0,3 μm / 0,2 μm	135	0,03	0,06	32,1	0,02	0,06
Valores médios:		140	0,01	0,05	32,7	0,02	0,05

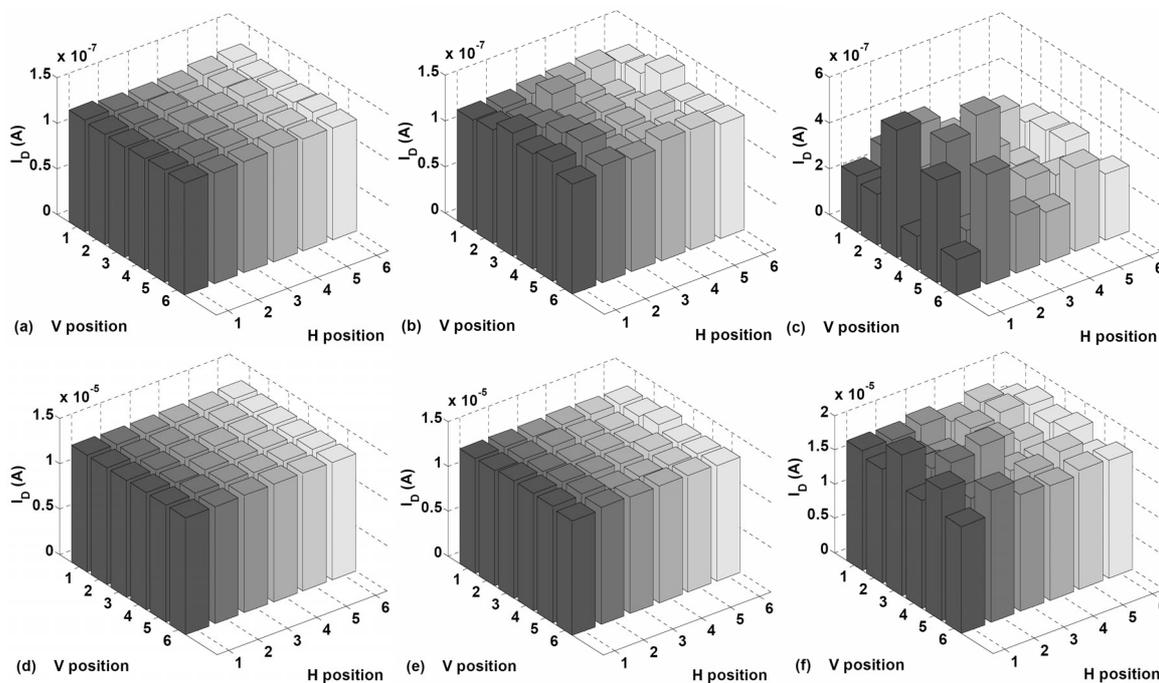


Figura 4.30: Corrente de *dreno* (I_D), medida dos grupos geométricos 16x16 (a, d), 04x04 (b, e), e 01x01 (c, f), operando em saturação sob duas condições de inversão: $i_f = 1$ (a, b, c) e $i_f = 100$ (d, e, f). Cada coluna sólida representa a corrente de um dos transistores que formam a matriz do grupo. Resultados obtidos na tecnologia *TSMC 0.35*.

Nota-se claramente que, dada a magnitude das flutuações na corrente de *dreno*, um projeto utilizando transistores de dimensão mínima (01x01) sob inversão fraca (gráfico c) só é viável quando a precisão não é um quesito importante.

O valor médio e o desvio-padrão de I_D foram calculados para cada condição, resultando: $\mu_a = 122$ nA, $\sigma_a = 2$ nA; $\mu_b = 124$ nA, $\sigma_b = 7$ nA; $\mu_c = 287$ nA, $\sigma_c = 114$ nA; $\mu_d = 12,9$ μ A, $\sigma_d = 0,066$ μ A; $\mu_e = 12,9$ μ A, $\sigma_e = 0,19$ μ A; and $\mu_f = 17,2$ μ A, $\sigma_f = 1,45$ μ A. Os transistores menores (c, f) apresentam uma corrente média mais alta devido aos efeitos de canal-curto associados à maior magnitude das flutuações.

4.9 Conclusões

Neste capítulo foram apresentadas as necessidades experimentais para a caracterização CC do descasamento entre transistores *MOS*, focando os aspectos mais relevantes deste processo. Duas versões de circuitos de teste foram apresentadas, sendo que a primeira foi implementada nas tecnologias *TSMC 0.35*, *AMIS 0.5* e *AMIS 1.5*, mas oferece pouca diversidade de dispositivos e necessita do processo de caracterização manual. A segunda versão apresentada é mais complexa e foi implementada nas tecnologias *TSMC 0.35* e *TSMC 0.18*, oferecendo maior diversidade geométrica e mais amostras por grupo, aumentando a confiabilidade dos resultados. Este circuito de teste de descasamento permite a caracterização de um grande número de transistores *MOS* muito próximos, de forma totalmente automática, e sem a necessidade do uso de equipamentos de alto custo. O uso da técnica de medida *Kelvin (force-sense)* permite o estímulo preciso de todos os terminais dos dispositivos, compensando as quedas de tensão nas conexões. A caracterização no modo *differential pooling*, medindo o descasamento entre os transistores contíguos de cada par, permite a melhor percepção do descasamento local, que é o foco deste trabalho. Neste circuito, a implementação de nove geometrias diferentes auxilia na identificação de alguns efeitos de segunda ordem. Além do descasamento, este circuito permite inúmeros tipos de caracterização elétrica, voltadas à extração de parâmetros tecnológicos, dimensionais e de modelagem. Todos os resultados obtidos de exaustivos ensaios ajudaram a corroborar o modelo proposto.

5 PROJETO DE UM CONVERSOR D/A M-2M

Neste capítulo são tratados o projeto e a implementação de um circuito conversor digital-analógico (D/A ou DAC) integrado, que utiliza uma rede de transistores *MOS* (rede M-2M) como elemento de ponderação binária. É descrita a metodologia completa de projeto, onde a inexatidão do conversor é relacionada à geometria e à polarização dos seus transistores, através do modelo de descasamento apresentado nesta tese. Os efeitos de canal-curto são também avaliados no projeto. Foram implementadas duas versões de 8 bits do conversor, com diferentes dimensões, em uma tecnologia CMOS 0,35 μm 3,3V. Sua caracterização elétrica demonstra a confiabilidade do modelo de descasamento como ferramenta de projeto. Até onde se sabe, esta é a primeira metodologia bem sucedida de projeto de conversores D/A M-2M, baseada em dados de descasamento entre transistores.

5.1 Objetivos

Neste ponto do trabalho, tem-se o modelo para o descasamento entre transistores *MOS* teoricamente substanciado, e experimentalmente comprovado. Resta então empregá-lo no projeto de um circuito de aplicação, cujo desempenho seja intrinsecamente dependente do adequado casamento entre os dispositivos que o compõem.

Os conversores analógico-digitais e digital-analógicos são elementos-chave nos circuitos para sinais mistos (*mixed-signal circuits*), que compõem a imensa maioria das aplicações eletrônica atuais.

Qualquer conversor necessita de alguma técnica para implementar a relação binária entre uma corrente ou tensão específica (grandezas analógicas) e um valor numérico correspondente (grandeza digital). Por muitos anos, entre as décadas de 60 e 80, as redes R-2R foram vastamente utilizadas para este propósito. A partir do início da década de 90, foi demonstrado [113] que o princípio de divisão proporcional de correntes das redes R-2R também poderia ser aplicado a redes formadas por transistores *MOS*, que são elementos intrinsecamente não-lineares, formando as chamadas redes M-2M.

As redes M-2M [114] não consomem área excessiva de silício, e podem ser utilizadas em aplicações para baixa tensão e baixa potência. Também, como nas tecnologias *MOS* atuais pode-se obter um casamento entre transistores com desvios inferiores a 0,1%, resoluções efetivas entre 10 e 12 bits podem ser alcançadas sem necessidade de ajustes (*trimming*).

Optou-se por implementar na mesma pastilha dois conversores, de 8 bits cada, empregando a estrutura clássica M-2M, com todos os transistores idênticos, apenas diferindo a geometria empregada em cada conversor.

A programação dos conversores é feita através de um registrador de deslocamento interno, de forma semelhante ao que é feito nos circuitos de teste. Essa estratégia permite que a mesma infra-estrutura de caracterização elétrica desenvolvida para estes circuitos seja empregada na medição dos conversores.

5.2 O divisor de corrente baseado em MOSFETs

O uso de transistores “idênticos” em paralelo, como forma de se obter correntes com uma razão determinada entre si, tem sido utilizado desde os primórdios da eletrônica de semicondutores. Em 1992, Bult e Geelen [113] demonstraram formalmente este princípio, quando aplicado a divisores de corrente formados por transistores *MOS*.

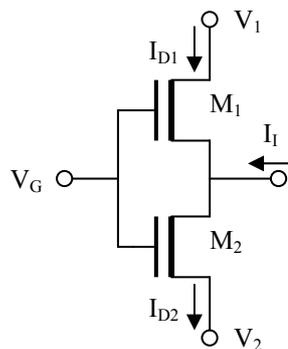


Figura 5.1: Forma geral do divisor de corrente.

Considera-se, para a fig. 5.1, que as tensões aplicadas V_G , V_1 e V_2 são constantes. Em sua forma mais geral, o princípio de divisão de corrente entre dois transistores *MOS* estabelece que, no momento em que for aplicada uma corrente externa (I_I) ao nó intermediário dos dois transistores, esta será dividida em duas partes, que fluem pelos transistores M_1 e M_2 , da seguinte forma [113]:

- as partes são constantes e proporcionais à corrente aplicada (I_I);
- independem dos valores de V_1 e V_2 ;
- independem se os transistores estão saturados ou na região linear;
- e independem do nível de inversão dos transistores.

Resumidamente, pode ser demonstrado [113] através deste princípio que

$$\frac{\Delta I_{D1}}{\Delta I_{D2}} = -\frac{W_1/L_1}{W_2/L_2}, \quad (5.2.1)$$

onde ΔI_{D1} representa o acréscimo na corrente I_{D1} e ΔI_{D2} representa o acréscimo na corrente I_{D2} , ambos resultantes da aplicação da corrente I_I .

Caso os transistores sejam idênticos e ligados em uma disposição mais usual, como na fig. 5.2, fica intuitivo este princípio, ou seja

$$I_{D1} = I_{D2} = \frac{I_D}{2}. \quad (5.2.2)$$

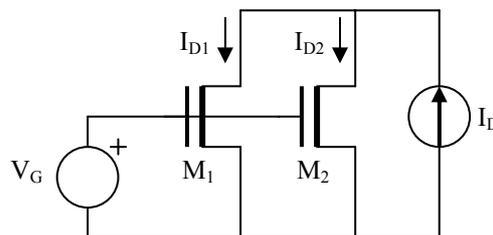


Figura 5.2: Dois transistores *NMOS* idênticos e sob mesma polarização formam um divisor de corrente por 2.

Em um transistor *MOS*, a corrente de dreno é diretamente proporcional à largura (W) e inversamente proporcional ao comprimento (L) da região efetiva do canal. Esta proporcionalidade é válida para os transistores ditos de “canal longo”, onde os efeitos devido às pequenas dimensões do canal ainda não se fazem sentir, e independe do nível de inversão ou da condição de saturação do transistor.

Partindo-se deste princípio, pode ser demonstrado que a associação série ou paralelo de transistores *MOS* idênticos (fig. 5.3), resulta numa rede cujo comportamento elétrico equivale ao de um único dispositivo, com dimensões iguais ao resultado da agregação das dimensões dos transistores que compõem a rede.

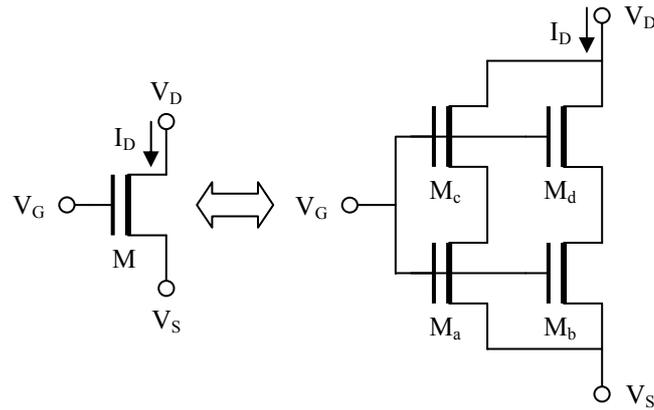


Figura 5.3: Equivalência entre um transistor *MOS* e uma associação série-paralelo de transistores.

Utilizando-se novamente o princípio de associação de transistores *MOS*, pode-se também substituir M_b da fig. 5.3 pela associação série-paralelo dos transistores idênticos M_{ba} , M_{bb} , M_{bc} e M_{bd} (fig. 5.4).

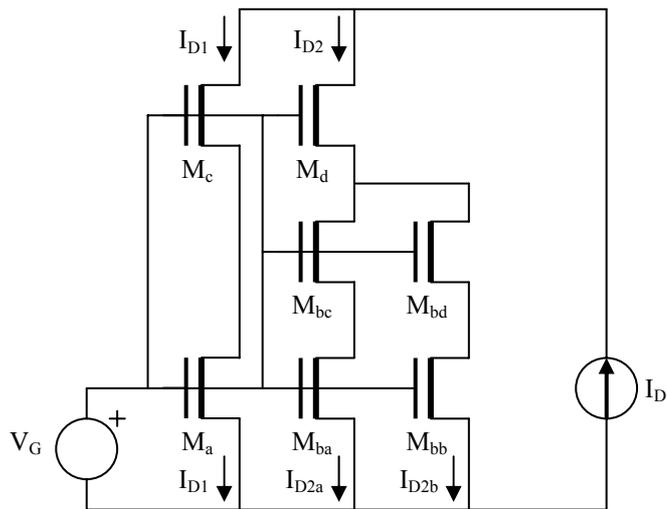


Figura 5.4: Decomposição de uma associação paralela de transistores *MOS* (fig. 5.2) em seu equivalente, através de associações série-paralelo.

Através do princípio de divisão de corrente, aplicado à fig. 5.4, pode-se afirmar que

$$I_{D1} = I_{D2} = \frac{I_D}{2}, \quad (5.2.3)$$

e que

$$I_{D2a} = I_{D2b} = \frac{I_{D2}}{2} = \frac{I_D}{4}. \quad (5.2.4)$$

Assim, fica claro que é possível se estabelecer uma relação binária entre as correntes que circulam pelos ramos de uma rede resultante da associação série-paralelo de transistores *MOS*.

5.3 O conversor D/A M-2M

A rede da fig. 5.4 pode continuar sendo expandida através da representação do transistor M_{bb} por outra associação série-paralelo, obtendo-se mais uma fração binária da corrente I_D , e assim sucessivamente, formando o circuito tradicionalmente denominado de “rede M-2M”, em analogia à rede de resistores tipo “R-2R”, cujo princípio de operação é semelhante.

A fig. 5.5 apresenta o diagrama esquemático simplificado do conversor D/A implementado, onde se pode observar a seqüência de associações série-paralelo de transistores, que formam o divisor binário da corrente aplicada ao nó I_R .

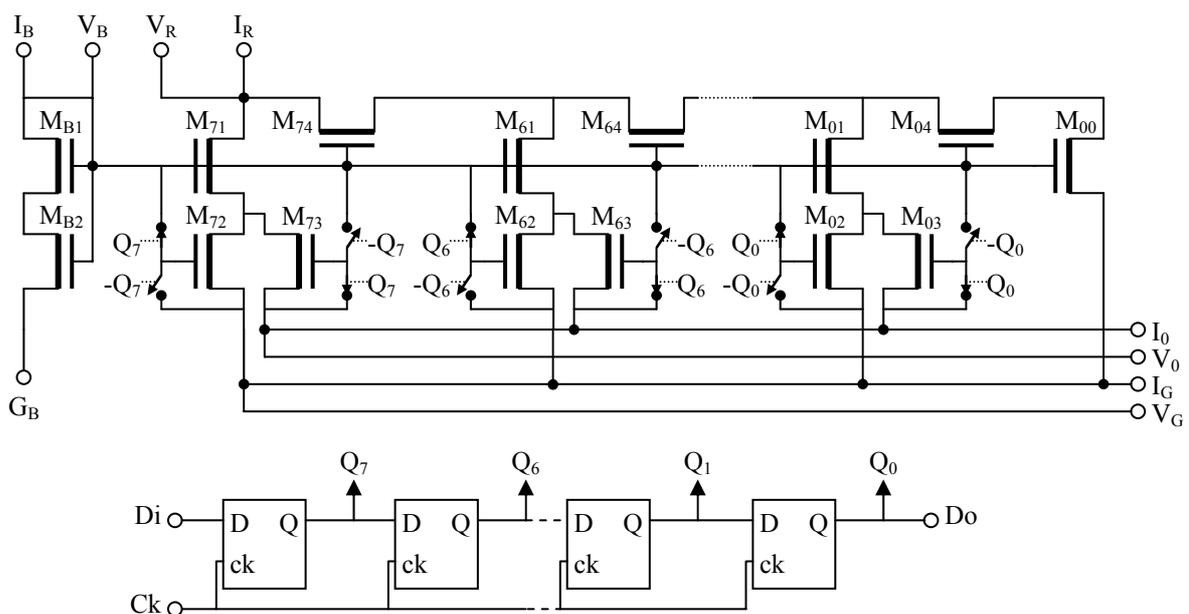


Figura 5.5: Diagrama esquemático do conversor D/A de 8 bits, cuja relação binária das parcelas de corrente é obtida através de uma rede, composta por associações série-paralelo de transistores *MOS* (rede M-2M). O valor digital, a ser convertido em analógico, é programado em um registrador de deslocamento.

Nesta figura, cada transistor inferior da associação apresentada na fig. 5.4, é implementado através de 2 transistores com os drenos em paralelo (M_{i2} e M_{i3} , $i = 7, 6 \dots 1, 0$), de modo que a fração binária da corrente I_R possa ser desviada para I_0 (nó de saída) ou

para I_G (nó de terra), conforme o estado da saída Q_i do registrador correspondente. A ambos os nós I_0 e I_G , são aplicados potenciais iguais a zero Volt. A tensão de *porta* da rede, que junto com a corrente I_R define o nível de inversão dos transistores, é estabelecida através de um circuito de polarização formado por M_{B1} e M_{B2} , e que é polarizado em corrente através do nó I_B . Essa estratégia permite que se possa avaliar o efeito do nível de inversão dos transistores da rede no desempenho do conversor. O nó G_B é conectado ao potencial de terra do circuito, ficando no mesmo nível de V_0 e V_G .

Este tipo de divisor de correntes programável tem duas vantagens principais: (a) os transistores *MOS* funcionam simultaneamente como elementos da rede divisora e como chaves, e (b) a impedância da rede é independente do número de bits e do dado programado.

Para reduzir o efeito das quedas de potencial nas conexões por onde circulam correntes significativas, estes nós foram implementados através da técnica *Kelvin (force-sense)*, onde são separados os condutores por onde flui a corrente (*force*) e por onde a tensão é medida (*sense*). Esta técnica aumenta a confiabilidade da caracterização elétrica dos conversores, pois compensa as quedas de tensão que ocorrem ao longo dessas conexões.

5.4 Metodologia de Projeto

Na topologia tradicional [114], o projeto fica restrito ao dimensionamento dos transistores da rede M-2M, que são todos de mesma geometria. Basicamente, este dimensionamento se dá em dois aspectos:

- determinação do comprimento dos transistores (L): esse fator está relacionado ao impacto dos efeitos de canal-curto no desempenho do conversor, e;
- determinação da área dos transistores (WL): esse fator está diretamente relacionado ao casamento entre eles, afetando a linearidade integral do conversor (relacionada à exatidão da representação de cada bit).

A seguir é apresentado o projeto de dois conversores M-2M de 8 bits cada, sendo um com erro estimado na faixa de $\pm\frac{1}{2}$ LSB (DAC1 - erro de 1/256), e outro na faixa de $\pm\frac{1}{4}$ LSB (DAC0 - erro de 1/512). LSB significa *least significant bit* ou “bit menos significativo”. O projeto foi realizado visando-se a tecnologia *TSMC 0.35*.

5.4.1 Determinação de L

A *saturação na velocidade de portadores* do canal é um dos efeitos de canal-curto que mais degrada a corrente de *dreno* de um MOSFET. Esse efeito se torna significativo na medida que o comprimento do canal é reduzido a um ponto onde a velocidade dos portadores não é mais proporcional ao campo elétrico longitudinal. Para reduzir o impacto desse efeito, pode-se: (a) operar todos os transistores da rede longe da condição de saturação [113], [114], ou (b) dimensionar o comprimento do canal (L) suficientemente grande, de modo que mesmo em saturação, o impacto deste efeito seja insignificante, frente à exatidão pretendida [114].

Observando-se a fig. 5.5, pode-se concluir que apenas os transistores conectados ao nó onde a corrente de referência é aplicada (I_R) são suscetíveis a entrar em saturação (M_{71} e M_{74}). Mas, infelizmente, estes são os transistores com maior impacto na exatidão do conversor, pois estão relacionados à divisão de corrente no seu “bit mais significativo” (MSB ou *most significant bit*).

O modelo empírico [91] utilizado neste projeto, para o efeito da *saturação na velocidade de portadores* sobre a mobilidade, é dado por

$$\mu_{sat} = \frac{\mu_0}{\sqrt{1 + (E_x/E_C)^2}}, \quad (5.4.1)$$

onde μ_0 corresponde à mobilidade elétrica de portadores sob baixo campo elétrico, E_C é o campo elétrico crítico, dado por $E_C = v_{sat}/\mu_0$, com v_{sat} sendo a velocidade de saturação de portadores. E_x corresponde ao valor da campo elétrico existente longitudinalmente no canal, junto à superfície, podendo ser aproximado por $E_x = V_{DS}/L$ [67], onde V_{DS} representa a tensão entre *dreno* e *fonte*.

A eq. (5.4.1) é reconhecida como uma aproximação bastante exata da dependência da mobilidade com campos elétricos elevados longitudinais no canal [91], [115]. O valor de v_{sat} é essencialmente independente da concentração de dopantes, sendo da ordem de 10^7 cm/s para elétrons e lacunas em temperatura ambiente [67], [116], o que resulta em $E_C \approx 2,3 \times 10^4$ V/cm (considerando-se $\mu_0 \approx 420$ cm²/Vs, dado fornecido pela MOSIS para os transistores NMOS da tecnologia TSMC 0.35).

De (5.4.1), pode-se estimar o comprimento de canal mínimo, necessário para que o efeito de *saturação de velocidade de portadores* resulte num desvio máximo er na mobilidade, sob determinado potencial *dreno-fonte*, ou

$$L_{\min} = \frac{V_{DS}}{E_c \sqrt{\frac{1}{(1-er)^2} - 1}}. \quad (5.4.2)$$

Considerando-se um potencial V_{DS} de 2 V (alto suficiente para saturar um transistor em um nível de inversão igual a 5000; esta tecnologia pode utilizar tensões máximas de 3,3 V), chega-se ao valor mínimo do comprimento do canal a ser utilizado no projeto, conforme o erro máximo aceitável provocado por esse efeito, o que é apresentado na tabela 5.1 a seguir.

Tabela 5.1: Comprimento mínimo do canal, em função do erro máximo aceitável.

er :	1/256 (DAC1)	1/512 (DAC0)
L_{\min} (μm):	9,81	13,89

Deve-se observar que o valor de comprimento mínimo obtido por (5.4.2) representa o comprimento do transistor equivalente, formado no conversor pela associação de uma série de dois transistores (por exemplo, na fig. 5.5, os transistores M_{71} e M_{72} , ou M_{71} e M_{73}). Assim, para o DAC1, determinou-se que todos os transistores usarão $L = 5 \mu\text{m}$.

5.4.2 Determinação da área

Na determinação da área dos transistores, utilizou-se o modelo desenvolvido para o descasamento entre transistores *MOS*. Os parâmetros deste modelo foram extraídos para a tecnologia alvo (*TSMC 0.35*) através do circuito de teste descrito no capítulo anterior, de um outro lote e fabricado cerca de 6 meses antes. A caracterização de 15 amostras deste circuito resultou nos seguintes valores extremos para os parâmetros do modelo: N_{oi} entre 2×10^{12} e $3,5 \times 10^{12} \text{ cm}^{-2}$, e B_{ISQ} entre 0,4 e 0,9 %- μm . Observa-se aqui que o N_{oi} extraído dos circuitos caracterizados apresentou uma variação de $\pm 32\%$ ao redor de um valor médio de $2,65 \times 10^{12} \text{ cm}^{-2}$, o que equivale a uma incerteza máxima no descasamento da corrente (σ_{I_D}/I_D) de $\pm 15\%$, pois este é proporcional à raiz quadrada do parâmetro.

Durante o projeto foi escolhida uma faixa de níveis de inversão para a operação do conversor, limitada entre um valor em inversão muito forte ($i_f = 2000$) e outro em inversão moderada ($i_f = 20$).

Usando-se o conjunto de parâmetros apresentado na expressão do modelo (aqui reproduzida novamente)

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{1}{WL} \left[\frac{N_{oi}}{N^{*2}} \frac{1}{i_f - i_r} \ln \left(\frac{1+i_f}{1+i_r} \right) + B_{ISQ}^2 \right], \quad (5.4.3)$$

se obteve a fig. 5.6. Nesta figura são apresentadas as curvas de descasamento de um transistor *NMOS*, para os níveis de inversão extremos da faixa escolhida, traçadas com os valores extremos dos parâmetros de modelo.

Para que um conversor D/A mantenha sua monotonicidade, e apresente seu número de bits como efetivos, é necessário que sua não-linearidade integral máxima esteja entre $\pm 1/2$ LSB, o que representa uma faixa de variação absoluta de 1 LSB.

Na rede M-2M, a corrente de saída I_0 é a soma das correntes que passam por todos os transistores M_{i3} que estiverem acionados. Isso faz com que, no final da escala (dado programado $D = 255$), a corrente de saída se aproxime da de referência, ou $I_0 = (255/256)I_R$. Assim, espera-se que o erro do conversor no final da escala seja muito pequeno, pois independentemente do descasamento entre os transistores, a corrente de saída será aproximadamente igual à de referência. Disso resulta que a incerteza do conversor, resultante do descasamento entre os transistores, aumenta com o valor de D até o meio da escala ($D = 128$), passando então a reduzir na medida em que este se aproxima do final da escala. Pode-se então afirmar que o meio da escala é o ponto de maior expectativa de erro, o que ocorre justamente quando o MSB é acionado e todos os outros bits que estavam acionados passam a desligados.

Do exposto acima, pode-se concluir que o projeto deve ser realizado de tal forma que o desvio máximo provocado pelos transistores relacionados ao MSB, seja equivalente ao desvio máximo absoluto previsto para o conversor ($1/256$). O MSB representa aproximadamente a metade da corrente máxima do conversor, sendo responsável por metade da sua escala. Assim, deve-se dimensionar os transistores que compõem este bit de modo que seu erro máximo seja da ordem $1/128$ da corrente que passa por eles, ou 0,78%.

A exatidão de uma rede M-2M aumenta com o nível de inversão de seus transistores, pois seu descasamento diminui. Assim, utilizando-se (5.4.3) com os parâmetros na situação de pior caso, na condição de saturação (maior descasamento), e para o nível de inversão igual a 20, chega-se a uma área dos transistores de $94,4 \mu\text{m}^2$. Esta área corresponde a da associação série formada por M_{71} e M_{73} , de modo que se pode utilizar $W = 10 \mu\text{m}$ para todos os transistores (no item anterior foi determinado que $L = 5 \mu\text{m}$ para todos os transistores), resultando numa área total da associação igual a $100 \mu\text{m}^2$.

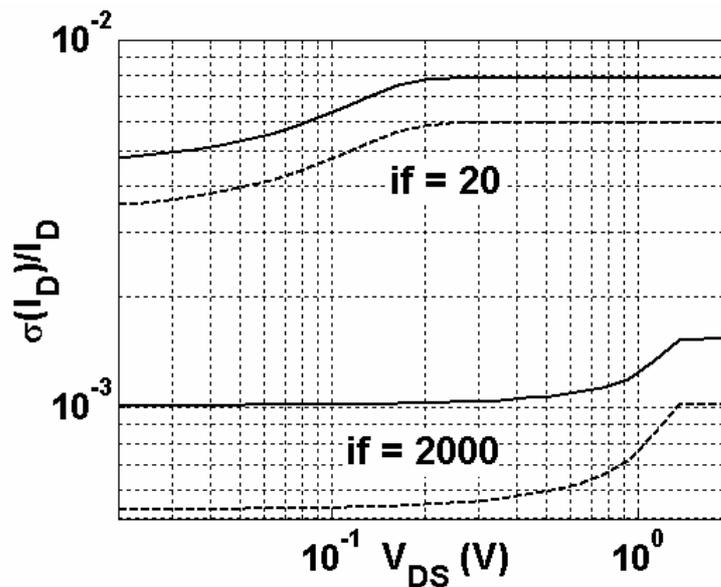


Figura 5.6: Descasamento de corrente normalizado de um transistor com $10 \mu\text{m} \times 10 \mu\text{m}$, sob dois níveis de inversão diferentes (20 e 2000), usando o modelo apresentado com os parâmetros extraídos de um lote da tecnologia TSMC 0.35. As linhas tracejadas correspondem a $N_{oi} = 2 \times 10^{12} \text{ cm}^{-2}$ e $B_{ISQ} = 0,4 \text{ \%}\cdot\mu\text{m}$. As linhas sólidas correspondem a $N_{oi} = 3,5 \times 10^{12} \text{ cm}^{-2}$ e $B_{ISQ} = 0,9 \text{ \%}\cdot\mu\text{m}$. A tensão *dreno-fonte* (V_{DS}) varia logarithmicamente entre 2 V e 20 mV.

A fig. 5.6 apresenta as curvas de (5.4.3) para um transistor com área de $100 \mu\text{m}^2$, para os valores extremos dos parâmetros e os limites especificados para o nível de inversão, desde a região linear até a saturação (V_{DS} entre 20 mV e 2 V). Pode-se observar nela que o máximo descasamento fica pouco abaixo de 0,8%, conforme determinado.

No projeto do DAC0, deseja-se a metade da inexatidão do DAC1, o que significa a metade do descasamento em corrente dos transistores do MSB. Como o descasamento é inversamente proporcional à raiz quadrada da área de um transistor, chega-se à conclusão que se deve utilizar nesse projeto transistores com área igual a 4 vezes a dos utilizados em

DAC1. Assim, utilizou-se nesse projeto $W = 20 \mu\text{m}$ e $L = 10 \mu\text{m}$ para todos os transistores, resultando numa área total da associação série igual a $400 \mu\text{m}^2$.

5.5 Implementação

Os conversores foram implementados em uma mesma pastilha e seu leiaute foi feito de modo a se reduzir as causas dos erros sistemáticos. Para tanto, empregou-se trilhas largas e com muitos contatos, nos caminhos por onde as correntes são significativas. Ainda, como já citado, empregou-se a técnica *Kelvin* nessas conexões.

Os transistores foram dispostos de forma a se ter uma matriz com distribuição homogênea, e dispositivos falsos (*dummies*) foram colocados nas proximidades dos transistores que estão em posições externas. Um anel de guarda foi desenhado ao redor da rede de transistores, de modo a evitar que correntes parasitas prejudiquem seu desempenho.

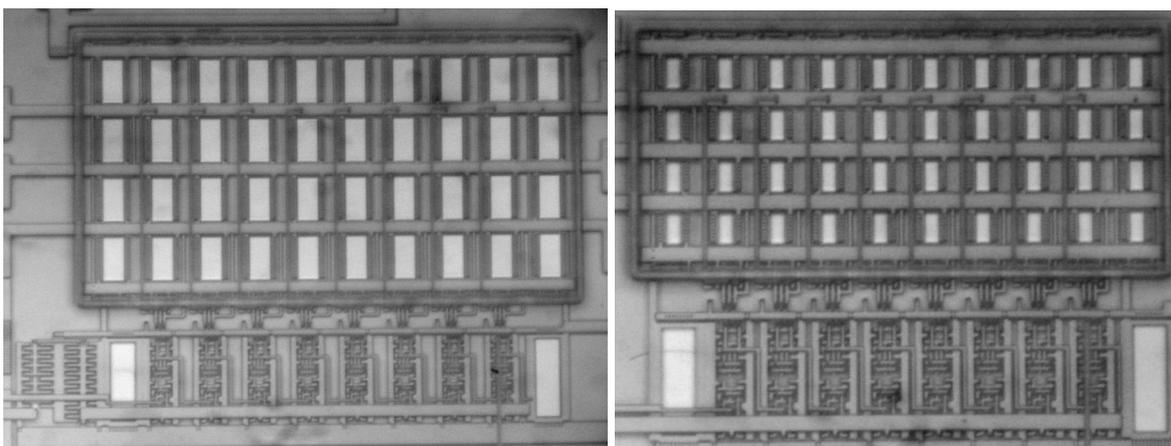


Figura 5.7: Microfotografia dos conversores fabricados: DAC0 (esq.) e DAC1 (dir.). A área superior é a rede M-2M, onde se observa claramente a matriz de 4x10 transistores, cercada pelo anel de guarda e *dummies*. Sob essa área vê-se claramente os 8 registradores, as chaves de acionamento da rede e dois capacitores de desacoplamento de alimentação. No DAC0, ao lado esquerdo dos registradores, observa-se o buffer do sinal C_k .

Foram utilizados registradores de deslocamento separados para cada conversor, de modo que se pode utilizar os conversores simultaneamente e de forma independente. Estes registradores estão conectados em série e são programados simultaneamente por um mesmo vetor de teste. Como no circuito de teste, aqui também se disponibilizou o bit de saída do último registrador (D_0), de modo que se pode verificar a se o vetor gravado

corresponde ao que foi programado. Este bit é lido pelo computador através da interface serial.

Foram implementados circuitos de reforço (*buffers*), tanto bit de saída (D_O) quanto no sinal de relógio de entrada (Ck).

A área total ocupada pela rede M-2M de DAC0, incluindo o anel de guarda e os *dummies*, ficou em $235 \mu\text{m} \times 127 \mu\text{m}$, e a de DAC1 em $185 \mu\text{m} \times 87 \mu\text{m}$. O leiaute não foi feito de forma a se otimizar a área utilizada.

5.6 Caracterização e resultados

O comportamento CC dos DACs foi obtido através da caracterização de 20 amostras do mesmo lote do processo *TSMC 0.35*. Para tanto, foi utilizado o *Agilent 4156B* (*semiconductor parameter analyzer*), com sua programação controlada por um computador PC através da sua interface GPIB, por onde também são adquiridos os dados resultantes. O mesmo computador também é responsável pela programação dos dados nos conversores, através de uma interface serial opticamente isolada (para reduzir a indução de ruído elétrico no circuito medido). Um programa específico foi desenvolvido em linguagem “C”, para o gerenciamento de toda a caracterização, e armazenamento de resultados.

Os conversores foram caracterizados em diversos níveis de inversão (i_f : 2; 6,33; 20; 63,3; 200; 633 e 2000) e com diversos fatores de relação (K_I : 0,5; 1; 1,5; 1,8; 1,9 e 2) entre a corrente de referência (I_R) e a de polarização (I_B), onde $K_I = I_R/I_B$. O fator K_I determina a condição de saturação da rede. Pela estrutura da rede M-2M utilizada, um fator K_I muito próximo a 2 (a corrente que passa pelo ramo do MSB é aproximadamente a metade de I_R), a coloca em saturação, e um fator muito baixo a coloca na região linear, fazendo com que a tensão no nó V_R seja pequena. Uma tensão muito pequena neste nó faz com que as tensões de *off-set* presentes nos nós V_0 e V_G , resultantes dos canais analógicos do equipamento de medida, interfiram mais na medida, prejudicando-a. Assim, é aconselhável o uso de valores de K_I próximos, mas inferiores, a 2. Além disso, o programa de controle executa uma etapa de estimativa desse *off-set* antes de cada caracterização, de modo a compensá-lo em parte.

Cada conversor foi caracterizado para todos os dados de entrada D (0 a 255), e seu erro (Err) foi calculado através da comparação da corrente de saída (I_0) medida, com um

valor ideal obtido através da reta $(D/256)I_R$. Este procedimento foi executado para todas as 20 amostras.

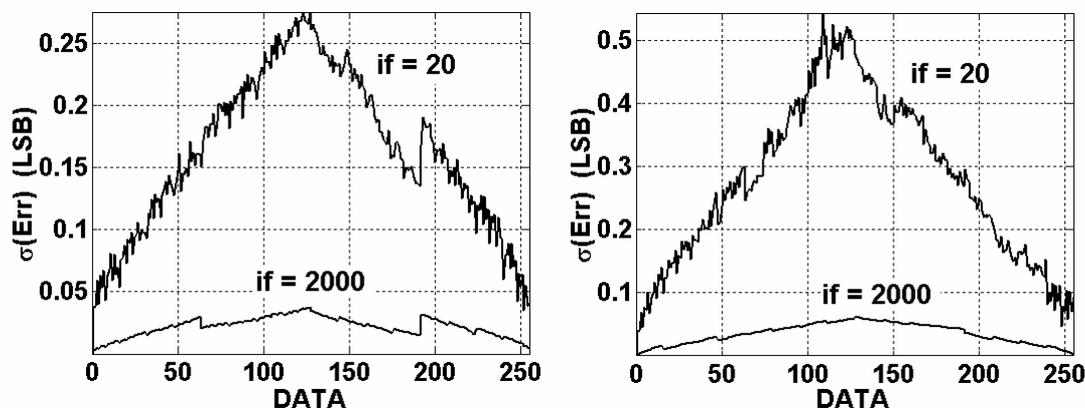


Figura 5.8: Desvio-padrão do erro medido das 20 amostras de DAC0 (esq.) e DAC1 (dir.), para todos os dados de entrada, e normalizado para 1 LSB. As medidas foram realizadas sob os níveis de inversão 20 e 2000, e com $K_I = 1,9$.

A fig. 5.8 apresenta o desvio-padrão dos erros obtidos de DAC0 e DAC1, calculado para cada dado de entrada, sob os dois níveis de inversão extremos (20 e 2000) e com $K_I = 1,9$. Desta figura pode-se observar que a incerteza média quadrática está próxima do valor projetado ($\frac{1}{4}$ LSB para DAC0 e $\frac{1}{2}$ LSB para DAC1). O erro máximo ocorre no entorno do centro da escala, o que mostra a predominância da incerteza do MSB sobre os outros bits. Da fig. 5.6 observa-se que as curvas de descasamento para os níveis de inversão 20 e 2000 estão separadas quase 1 década. Assim, pode-se esperar aproximadamente a mesma relação para os erros dos conversores, conforme pode ser visto na fig. 5.8. Também, pode-se observar que o erro de DAC0 é cerca da metade do de DAC1, para os dois níveis de inversão, conforme projetado.

A fig. 5.9 apresenta as curvas de erro dos conversores que apresentaram o mínimo e o máximo erro (extremos) de DAC0 e DAC1, sob os níveis de inversão 20 e 2000. Pode-se observar por essas curvas que, nos casos de máximo erro, a causa foi o descasamento maior do MSB (centro da faixa).

Como este não é um projeto para uso industrial, foi todo dimensionado tendo-se como meta a distribuição dos erros das amostras fabricadas dentro da faixa de 1σ , mas tendo-se ciência que isso implica em que apenas 68,3% das amostras estarão dentro da faixa de erro proposta. Isso foi feito para facilitar o confronto dos resultados experimentais com o modelo teórico de descasamento utilizado. Em projetos industriais, geralmente

dimensiona-se os circuitos para que fiquem dentro da faixa de 2 a 3 σ , aumentando a eficiência da produção (*yield*) para valores entre 95,4% e 99,7%, respectivamente.

Deve-se também lembrar que foram caracterizadas 20 amostras de cada conversor, o que significa que os resultados estatísticos obtidos experimentalmente têm uma incerteza intrínseca da ordem de 15%.

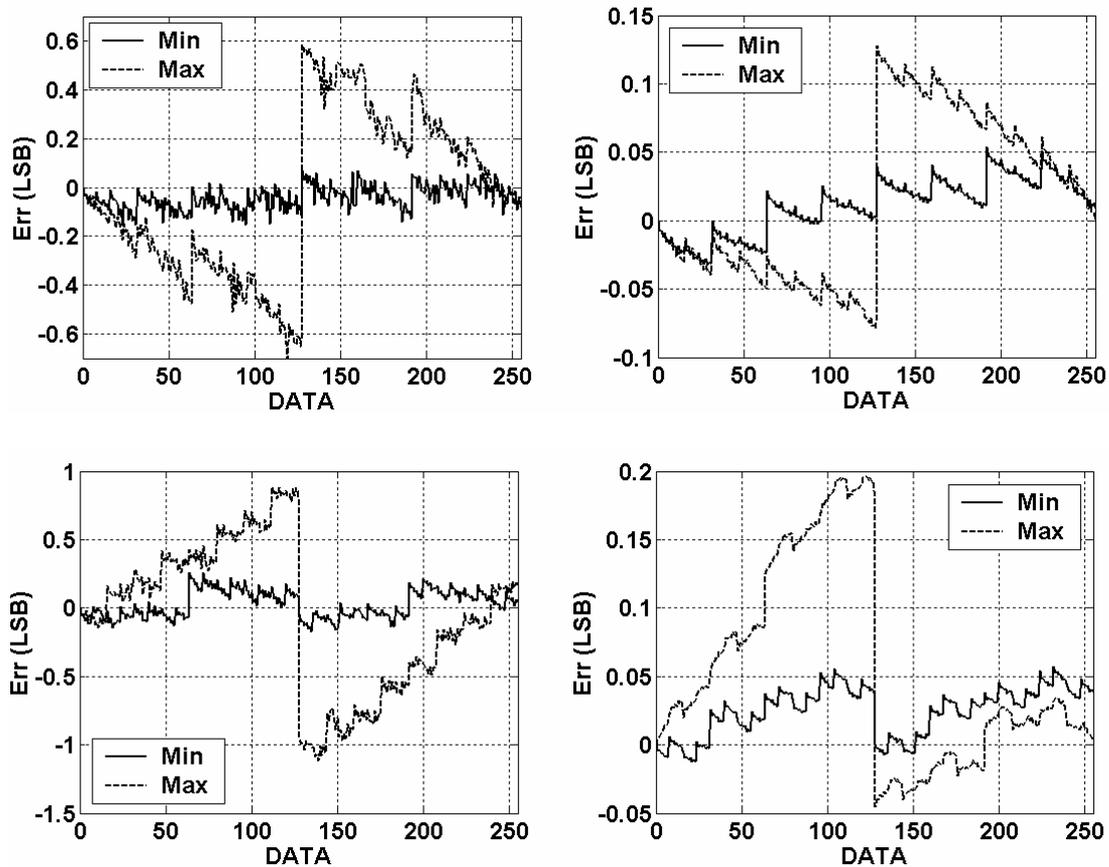


Figura 5.9: Amostras de DAC0 (sup.) e DAC1 (inf.) que apresentaram os valores mínimo e máximo de erro medido, sob os dois níveis de inversão extremos, 20 (esq.) e 2000 (dir.). As medidas estão normalizadas para 1 LSB e foram realizadas com $K_I = 1,9$.

5.7 Conclusões

Neste capítulo foi desenvolvido um conversor D/A M-2M, utilizando-se o modelo de descasamento como ferramenta de projeto, de forma a se dimensionar os transistores que o compõem com base em seu erro máximo especificado e em sua polarização. Os parâmetros do modelo para a tecnologia alvo, foram obtidos através da caracterização do circuito de teste já descrito em um capítulo anterior. O efeito de saturação da velocidade dos portadores foi também levado em conta durante o projeto, para a determinação do

comprimento do canal. Dois conversores foram projetados para exatidões diferentes, e fabricados na tecnologia *TSMC 0.35*.

As medidas experimentais foram obtidas de 20 amostras de cada conversor, de modo a se compor os resultados estatísticos para este lote. Tais resultados demonstram que a metodologia de projeto aqui apresentada foi bem sucedida, prevendo satisfatoriamente o descasamento dos transistores, e seu impacto no erro dos conversores.

6 CONCLUSÕES

O descasamento entre transistores *MOS* tem se tornado um problema fundamental na indústria de semicondutores atual, na medida em que as tecnologias estão se aproximando, cada vez mais, dos limites impostos pela natureza discreta da matéria. Isso ocorre devido à progressiva redução dimensional dos dispositivos, princípio que tem permitido o aumento no desempenho dos circuitos eletrônicos. Diversos modelos teóricos para o descasamento na tecnologia *MOS* foram propostos desde a década de '80, sendo que geralmente estes pecam ou pela simplicidade, sendo válidos apenas sob condições de operação específicas, ou por resultarem em expressões muito complexas, o que torna necessário o uso de pesados recursos computacionais.

Nesta tese de doutorado, o estudo do descasamento *CC* entre transistores *MOS* tipo *bulk* foi apresentado, incluindo a proposta de um modelo consistente e compacto, baseado em uma nova abordagem física para este efeito. Nesta abordagem as flutuações locais na concentração de dopantes no substrato são contabilizadas, integrando-se seus efeitos sobre a corrente de *dreno* ao longo do canal do dispositivo, mas se respeitando seu comportamento intrinsecamente não-linear. O modelo resultante necessita de apenas dois parâmetros experimentais para prever o descasamento estocástico entre transistores, em função de sua geometria e polarização, oferecendo ótima exatidão sob qualquer condição de operação, desde a inversão fraca até a forte, e desde a região linear até a saturação, fato que representa um grande avanço em relação aos modelos até então disponíveis. Além disso, substitui com vantagens as simulações *Monte Carlo*, que necessitam de muitos recursos computacionais e consomem muito tempo.

Alguns efeitos secundários e suas implicações sobre o modelo foram analisados, incluindo a saturação na velocidade dos portadores, a modulação na profundidade da região de depleção, a depleção no poli-silício da *porta*, e a rugosidade das bordas limítrofes do canal.

Um amplo estudo sobre estruturas e circuitos para caracterização estatística do descasamento foi realizado. Duas versões de circuitos de teste foram implementadas em diversas tecnologias, de modo a oferecer o suporte experimental necessário para o modelo. A primeira versão oferece grupos de dispositivos idênticos, que são individualizados manualmente durante o processo de caracterização. A segunda versão foi desenvolvida para permitir sua caracterização elétrica de forma totalmente automatizada, tornando-a mais rápida, precisa e confiável. Nesta versão o dispositivo-sob-teste é selecionado em um grupo através de chaves implementadas internamente, programáveis através de um computador externo que organiza o procedimento de caracterização, controlando também o equipamento de medida e servindo para o armazenamento dos resultados. Esta versão disponibiliza uma grande quantidade de dispositivos em variadas geometrias, resultando em resultados estatísticos mais precisos do descasamento, quando comparados à versão anterior. O grande volume de dados experimentais obtidos serviu para corroborar o modelo apresentado, além de permitir a extração dos parâmetros empíricos necessários ao modelo.

O projeto e a caracterização de dois conversores digital-analógicos (DAC), com diferentes especificações, são apresentados como fechamento da tese. Neste tipo de circuito, o descasamento entre seus transistores é a causa fundamental da sua não-linearidade, um dos principais parâmetros de desempenho de um DAC. Uma metodologia de projeto é descrita, baseada no modelo proposto e incluindo também o efeito de saturação na velocidade dos portadores. Todas as amostras fabricadas foram caracterizadas e os resultados experimentais obtidos corroboram a metodologia apresentada.

Espera-se com este trabalho lançar uma nova luz sobre a compreensão do descasamento estocástico entre transistores *MOS*, oferecendo uma poderosa ferramenta de projeto para a previsão desse efeito, sem a necessidade de complexas simulações discretas ou estatísticas.

Como trabalhos futuros pensa-se na inclusão de outros efeitos secundários ao modelo, e na possibilidade de sua adaptação a outros tipos de dispositivos e tecnologias em uso (SOI, FinFETs, etc). A análise dos efeitos da flutuação de dopantes no sentido da largura do transistor também pode resultar em uma melhoria significativa do modelo. Ainda, pensa-se no desenvolvimento de uma nova versão mais compacta do circuito de teste, que possa ser incorporado às lâminas com baixo consumo de área, de forma a tornar o processo de verificação e extração de parâmetros de descasamento mais simples e rápido.

Anexo A – Consistência de um modelo de descasamento

Em função da estrutura física de construção do transistor *MOS*, é amplamente aceito que o mesmo pode ser representado através de associações equivalentes dos tipos paralelo ou série, conservando seu comportamento *CC* (corrente-contínua) inalterado.

Por associação paralela, considera-se aqui a situação em que N transistores de mesmo comprimento L , têm seus quatro terminais equivalentes conectados em paralelo, resultando no comportamento *CC* equivalente ao de um único transistor de comprimento L e largura equivalente à soma das larguras dos N transistores. A figura A.1(a) apresenta essa associação (o terminal de substrato só é apresentado para o último transistor para facilitar a visualização).

Por associação série, considera-se aqui a situação em que M transistores de mesma largura W , têm seus terminais de *porta* e substrato conectados em paralelo e cada terminal de *dreno* sucessivamente ligado ao terminal de *fonte* do transistor seguinte (fig. A.1(b)), resultando no comportamento *CC* equivalente ao de um único transistor de largura W e comprimento equivalente à soma dos comprimentos dos M transistores.

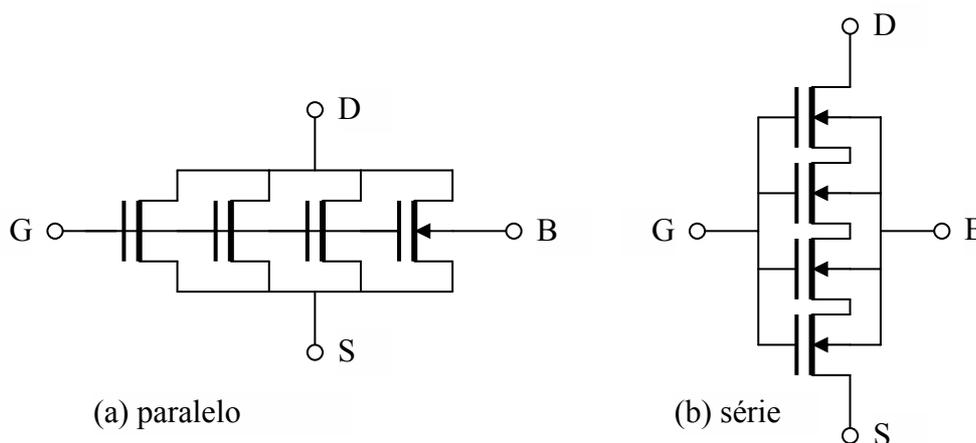


Figura A.1: Associações paralelo e série de transistores *MOS*.

Considera-se que um modelo de dispositivo é consistente se representações diferentes, embora equivalentes, do dispositivo através desse modelo resultarem em uma

previsão de comportamento equivalente. Se representações equivalentes do dispositivo resultarem em comportamentos diferentes previstos por esse modelo, considera-se que ele é inconsistente [49]. Em outras palavras, um modelo de descasamento apresenta consistência para associações do tipo série ou paralelo quando o valor do descasamento estimado com seu uso para um transistor equivaler ao estimado para uma associação equivalente de transistores em série ou paralelo.

A.1 Consistência de associação paralela

Considere as figuras A.2(a), onde a região ativa de um transistor é representada também como a associação em paralelo das regiões ativas de dois transistores de menor largura, A.2(b), que corresponde ao seu circuito equivalente de polarização *CC* e A.2(c), que apresenta o circuito das variações (pequenos sinais) equivalente.

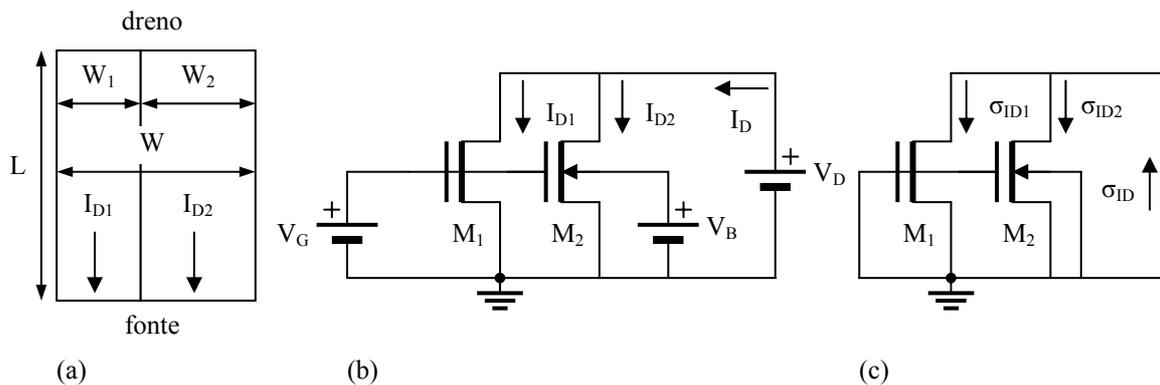


Figura A.2: Equivalente paralelo de um transistor *MOS* e seus circuitos de polarização e de variações.

No circuito de polarização, vê-se claramente que a corrente total da associação é o resultado da soma das correntes que fluem pelos dois dispositivos.

Considerando-se que, no circuito das variações, os desvios-padrão das correntes que fluem pelos dois dispositivos não apresentam qualquer correlação entre si, pode-se calcular o desvio-padrão resultante da corrente total como

$$\sigma_{I_D}^2 = \sigma_{I_{D1}}^2 + \sigma_{I_{D2}}^2 . \quad (\text{A.1})$$

Assim, um modelo de descasamento é consistente para associações do tipo paralelo quando a relação acima for respeitada pelo mesmo.

A.2 Consistência de associação série

Considere as figuras A.3(a), onde a região ativa de um transistor é representada também como a associação em série das regiões ativas de dois transistores de menor comprimento, A.3(b), que corresponde ao seu circuito equivalente de polarização *CC* e A.3(c), que apresenta o circuito das variações (pequenos sinais) equivalente.

No circuito de polarização, a componente de variação da corrente de *dreno* resulta da combinação de efeitos dos dois transistores em série.

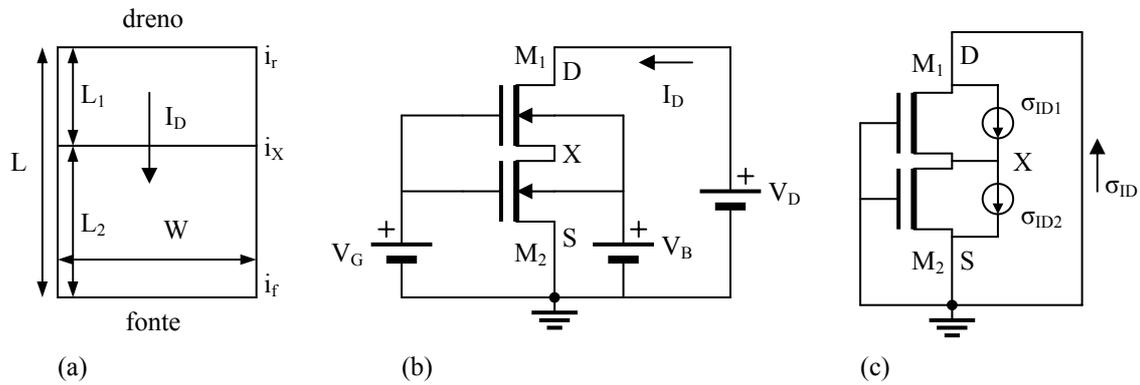


Figura A.3: Equivalente série de um transistor *MOS* e seus circuitos de polarização e de variações.

Considerando-se que, no circuito das variações, os desvios-padrão das correntes que fluem por cada um dos dispositivos podem ser representados por *fontes* de corrente em paralelo com dispositivos ideais (ou seja, representando-se separadamente a parcela que varia e o dispositivo sem variações), e que estes desvios-padrão não apresentam qualquer correlação entre si, pode-se calcular o desvio-padrão resultante da corrente total através do somatório dos efeitos das duas *fontes* sobre as transcondutâncias dos dispositivos separadamente, ou seja

$$\sigma_{I_D}^2 = \left(\sigma_{I_{D1}} \frac{g_{md2}}{g_{ms1} + g_{md2}} \right)^2 + \left(\sigma_{I_{D2}} \frac{g_{ms1}}{g_{ms1} + g_{md2}} \right)^2. \quad (\text{A.2})$$

Através do modelo *ACM*, tem-se a transcondutância vista pelo *dreno* ou *fonte* dada por

$$g_{ms(d)} = \frac{2I_{SQ}}{\phi_t} \frac{W}{L} \left(\sqrt{1 + i_{f(r)}} - 1 \right) \quad (\text{A.3})$$

permitindo o cálculo das transcondutâncias dos dois dispositivos como

$$g_{ms1} = \frac{2I_{SQ}}{\phi_t} \frac{W}{L_1} (\sqrt{1+i_{f1}} - 1) = \frac{2I_{SQ}}{\phi_t} \frac{W}{L_1} (\sqrt{1+i_X} - 1) \quad (A.4)$$

$$g_{md2} = \frac{2I_{SQ}}{\phi_t} \frac{W}{L_2} (\sqrt{1+i_{r2}} - 1) = \frac{2I_{SQ}}{\phi_t} \frac{W}{L_2} (\sqrt{1+i_X} - 1) \quad (A.5)$$

lembrando que $i_{f1} = i_{r2} = i_X$, pois as duas regiões ($S1$ e $D2$) apresentam a mesma densidade de carga de inversão.

Substituindo-se (A.4) e (A.5) em (A.2), chega-se a

$$L^2 \sigma_{I_D}^2 = L_1^2 \sigma_{I_{D1}}^2 + L_2^2 \sigma_{I_{D2}}^2 \quad (A.6)$$

cuja relação deve ser respeitada para que haja consistência no modelo de descasamento para associações do tipo série.

A.3 Consistência dos modelos apresentados

A consistência de associação paralela normalmente é atingida, pois considerando-se que dois dispositivos sejam ligados em paralelo, é lógico se supor que o efeito final seja o somatório dos comportamentos dos dispositivos. Desse modo, e afirmando-se que tanto o modelo proposto quanto o de Pelgrom possuem tal consistência, este caso não será analisado em mais detalhe.

A consistência de associação série não é tão intuitiva assim e é analisada em detalhe para os dois modelos.

A.3.1 Verificação do modelo proposto

A soma dos efeitos do descasamento dos dois transistores, segundo (A.6) é dada por

$$L_1^2 \sigma_{I_{D1}}^2 + L_2^2 \sigma_{I_{D2}}^2 \quad (A.7)$$

ou seja, substituindo-se os desvios padrão pela expressão do modelo proposto aqui tem-se

$$L_1^2 I_D^2 \frac{N_{oi}}{WL_1 N^{*2}} \frac{1}{i_X - i_r} \ln \left(\frac{1+i_X}{1+i_r} \right) + L_2^2 I_D^2 \frac{N_{oi}}{WL_2 N^{*2}} \frac{1}{i_f - i_X} \ln \left(\frac{1+i_f}{1+i_X} \right). \quad (A.8)$$

Lembrando-se que a corrente média que passa pelos dois dispositivos é a mesma, tem-se

$$I_D = I_{D1} = I_{D2} \quad \text{ou} \quad \frac{i_f - i_r}{L} = \frac{i_X - i_r}{L_1} = \frac{i_f - i_X}{L_2}. \quad (\text{A.9})$$

Substituindo-se a relação (A.9) em (A.8), chega-se a

$$L^2 I_D^2 \frac{N_{oi}}{WL N^{*2}} \left(\frac{1}{i_f - i_r} \ln \left(\frac{1+i_X}{1+i_r} \right) + \frac{1}{i_f - i_r} \ln \left(\frac{1+i_f}{1+i_X} \right) \right) = L^2 \sigma_{I_D}^2 \quad (\text{A.10})$$

de onde se conclui que o modelo proposto é consistente para associações do tipo série.

A.3.2 Verificação do modelo de Pelgrom

Considerando-se apenas a parcela referente às variações na tensão de limiar de (2.9), o modelo de Pelgrom pode ser expresso com o auxílio do modelo *ACM* em função dos níveis de inversão direto e reverso, chegando-se a

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{A_{VT}^2}{WL} \left(\frac{g_m}{I_D} \right)^2 = \frac{A_{VT}^2}{n^2 \phi_t^2 WL} \frac{4(\sqrt{1+i_f} - \sqrt{1+i_r})^2}{(i_f - i_r)^2}. \quad (\text{A.11})$$

A soma dos efeitos do descasamento dos dois transistores em série, segundo (A.6) é dada por

$$L_1^2 \sigma_{I_{D1}}^2 + L_2^2 \sigma_{I_{D2}}^2 \quad (\text{A.12})$$

onde, substituindo-se os desvios padrão pela expressão dada em (A.11) tem-se

$$I_D^2 \frac{A_{VT}^2}{W} \frac{4}{n^2 \phi_t^2} \left[L_1 \frac{(\sqrt{1+i_X} - \sqrt{1+i_r})^2}{(i_X - i_r)^2} + L_2 \frac{(\sqrt{1+i_f} - \sqrt{1+i_X})^2}{(i_f - i_X)^2} \right]. \quad (\text{A.13})$$

Considerando-se a relação (A.9) e substituindo-a em (A.13), vê-se que é impossível sua redução a

$$L^2 I_D^2 \frac{A_{VT}^2}{n^2 \phi_t^2 WL} \frac{4(\sqrt{1+i_f} - \sqrt{1+i_r})^2}{(i_f - i_r)^2} = L^2 \sigma_{I_D}^2 \quad (\text{A.14})$$

de modo que o modelo de Pelgrom se mostra inconsistente para associações do tipo série. A expressão (A.13) só se aproxima de (A.14) para a condição de inversão fraca e na região linear, divergindo na região de saturação sob inversão forte.

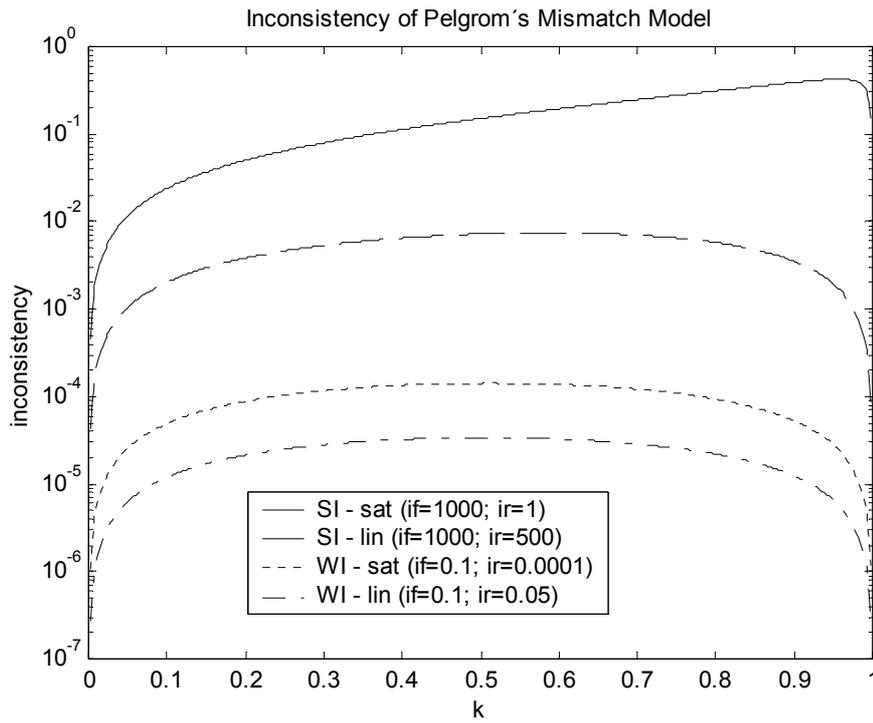


Figura A.4: Inconsistência do modelo de Pelgrom para inversão forte (SI) e fraca (WI), nas regiões linear e saturada.

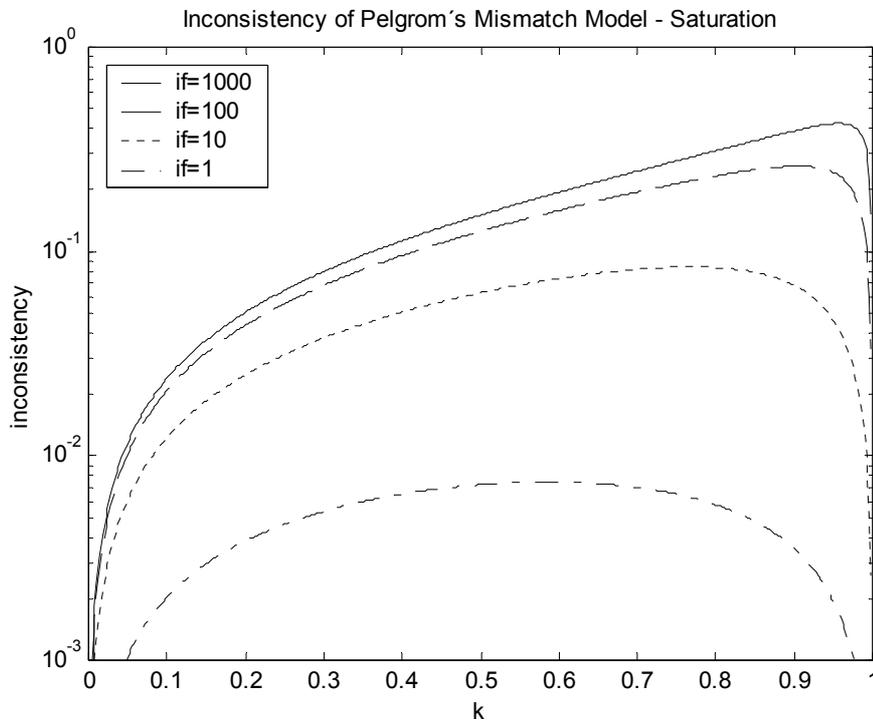


Figura A.5: Inconsistência do modelo de Pelgrom para diferentes níveis de inversão em saturação.

As figuras A.4 e A.5 apresentam a inconsistência do modelo Pelgrom para diversas condições de operação e considerando a seguinte troca de variáveis em (A.13):

$$L_1 = (1-k)L \text{ e } L_2 = kL, \text{ com } 0 < k < 1.$$

Nessas figuras, o parâmetro *inconsistency*, é dado por

$$inconsistency = \frac{L_1^2 \sigma_{I_{D1}}^2 + L_2^2 \sigma_{I_{D2}}^2}{L^2 \sigma_{I_D}^2} - 1. \quad (A.15)$$

Observa-se que, na situação de saturação e nível de inversão 1000, o parâmetro de inconsistência chega ao máximo de 42,2% para a situação em que $k=0,958$, e atinge 15,1% quando $k=0,5$. Isso significa que o modelo de Pelgrom prevê o descasamento de uma associação série com um erro de +15% para a condição de inversão forte quando os dois transistores têm o mesmo comprimento.

Anexo B – Flutuação local na carga de depleção e seus efeitos sobre o campo elétrico e tensão de limiar

O desenvolvimento a seguir apresenta o modo como as flutuações de carga que ocorrem na região de depleção, como resultado das flutuações na concentração de dopantes no sentido vertical (eixo y), devem ser quantizadas para se avaliar seu efeito médio líquido, visto pela interface óxido-semicondutor.

O princípio de conservação de cargas, quando aplicado ao capacitor *MOS*, resulta na expressão (3.6.4), já apresentada no capítulo 3 e reproduzida aqui novamente,

$$\Delta Q'_B(x) + \Delta Q'_G(x) + \Delta Q'_I(x) = 0, \quad (\text{B.1})$$

onde $\Delta Q'_G = -C'_{ox} \Delta \phi_S$ é a flutuação na carga de *porta*, e $\Delta Q'_I = -C'_i \Delta \phi_S$ é a flutuação na carga de inversão. A variação na carga de depleção $\Delta Q'_B$ resulta da soma de duas componentes, sendo a primeira igual a $-C'_b \Delta \phi_S$ e associada à flutuação no potencial de superfície, considerando-se que não há flutuações no número de impurezas, e a segunda designada $\Delta Q'_{IMP}$ e associada com a flutuação no número de impurezas ionizadas na região de interesse, considerando-se o potencial de superfície constante. Assim, a variação na carga de depleção é dada por

$$\Delta Q'_B = -C'_b \Delta \phi_S + \Delta Q'_{IMP}. \quad (\text{B.2})$$

B.1 Campo elétrico e carga de depleção equivalente

A fig. B.1 apresenta um corte esquemático da região ativa de um *MOSFET*, no sentido da profundidade do substrato (y), onde ficam evidentes a espessura do óxido (t_{ox}) e o substrato. A partir da interface óxido-substrato ($y = 0$), ocorre a região de depleção, que se estende até a profundidade y_d . A dimensão Δx representa uma pequena porção do

comprimento do canal L . A dimensão não apresentada corresponde à largura W do canal. Nesta figura, há uma flutuação na carga de depleção em uma fina fatia desta região na posição y , definida pelo volume $(W\Delta x\Delta y)$, e que é causada pela flutuação local na concentração de dopantes.

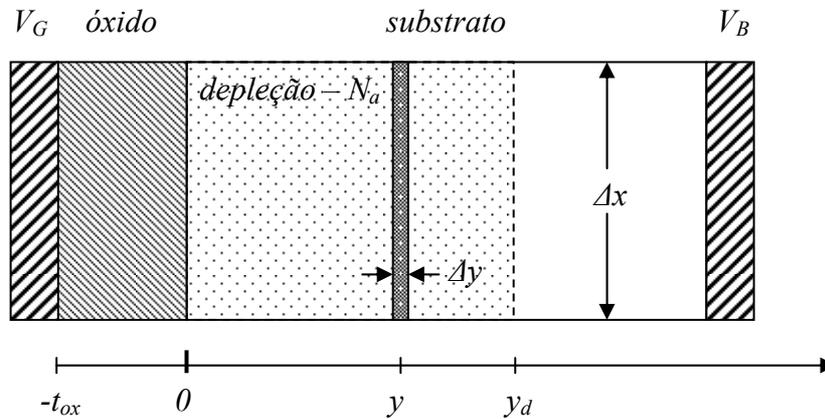


Figura B.1: Corte vertical da região ativa do MOSFET, onde ficam evidentes a espessura do óxido (t_{ox}) e o substrato. A partir da interface óxido-substrato ($y = 0$), ocorre a região de depleção, que se estende até a profundidade y_d . Em uma pequena fatia da região de depleção (Δy) há uma flutuação na sua carga, causada pela flutuação local na concentração de dopantes. A dimensão Δx representa uma pequena porção do comprimento do canal. A dimensão não apresentada corresponde à largura W do canal.

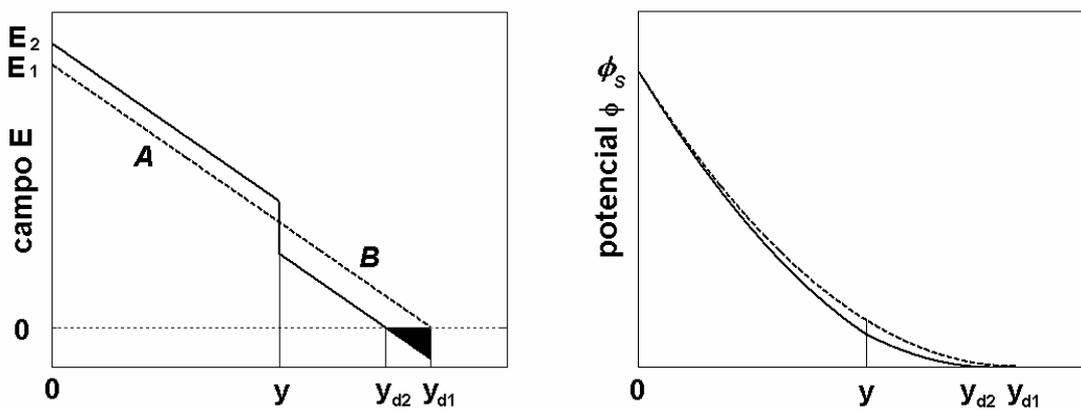


Figura B.2: Comportamento do campo elétrico E e do potencial ϕ , no sentido da profundidade do substrato. As linhas tracejadas correspondem à situação da concentração de dopantes homogênea, e as linhas sólidas, à situação onde há uma flutuação na carga de depleção na posição y .

A fig. B.2 apresenta o comportamento do campo elétrico E e do potencial ϕ , no sentido da profundidade do substrato (y). As linhas tracejadas correspondem à situação da

concentração de dopantes homogênea, sem flutuações de valor. Neste caso, tem-se um campo elétrico dado por $E_1(y)$, e o potencial de superfície correspondente ϕ_S , sendo que a região de depleção se estende até a profundidade y_{d1} .

Quando ocorre uma flutuação na densidade de carga $\delta Q'$ na fatia de espessura Δy , o campo elétrico sofre nesta posição uma descontinuidade de valor igual a

$$\delta E = \frac{\delta Q'}{\epsilon_{Si}} = \frac{q}{\epsilon_{Si}} \delta N_a \Delta y, \quad (\text{B.3})$$

onde N_a é a concentração média de dopantes e ϵ_{Si} é a permissividade elétrica do silício. O campo elétrico resultante está representado nesta figura por $E_2(y)$ (linha sólida).

A distribuição do potencial elétrico ao longo do eixo y também é modificada, conforme pode ser visto nesta figura. Como se está considerando que o potencial de superfície fica inalterado para essa situação, as duas curvas, correspondentes às situações *com* e *sem* a flutuação de carga, devem apresentar o mesmo valor ϕ_S . Assim, a descontinuidade no perfil do campo elétrico ocorre parte acima do campo original (δE_a) e parte abaixo (δE_b), de forma que $\delta E = \delta E_a + \delta E_b$, ou

$$\delta E_b = \delta E - \delta E_a = \frac{q}{\epsilon_{Si}} \delta N_a \Delta y - \delta E_a, \quad (\text{B.4})$$

o que provoca uma pequena alteração na profundidade total da região de depleção, de y_{d1} para y_{d2} .

O potencial resulta da integração do campo elétrico ao longo da sua direção. Como o valor ϕ_S tem de ser o mesmo para as duas situações, tem-se que

$$\begin{aligned} \int_0^{y_{d1}} E_1(y) dy &= \int_0^{y_{d2}} E_2(y) dy = \int_0^y [E_1(y) + \delta E_a] dy + \int_y^{y_{d2}} [E_1(y) - \delta E_b] dy = \dots \\ &= \int_0^{y_{d2}} E_1(y) dy + \underbrace{\delta E_a \int_0^y dy}_A - \underbrace{\delta E_b \int_y^{y_{d2}} dy}_B, \end{aligned} \quad (\text{B.5})$$

onde as integrais identificadas por A e B , correspondem às áreas das diferenças entre as duas curvas, indicadas nesta figura.

Como se está considerando uma flutuação de carga pequena, em relação ao seu valor médio, tem-se que δE (e conseqüentemente δE_a e δE_b) é pequeno em relação ao valor

de E . Isso faz com que a diferença entre y_{d1} e y_{d2} seja pequena, de modo que se pode aproximar $y_{d1} \approx y_{d2}$, e que daqui para frente será simplesmente referenciada como y_d . Essa aproximação faz com que a área triangular sombreada entre estas duas posições seja acrescida indevidamente ao cálculo do potencial, mas com pouco efeito, pois é muito pequena frente à área total sob a curva do campo (e que corresponde a ϕ_s).

Assim, aplicando-se $y_d = y_{d1} \approx y_{d2}$ a (B.5), resulta

$$\delta E_a \int_0^y dy = \delta E_b \int_y^{y_d} dy, \quad (\text{B.6})$$

que, quando integrada em conjunto com (B.4), resulta em

$$\delta E_a = \frac{q}{\varepsilon_{Si}} \delta N_a \left(1 - \frac{y}{y_d} \right) \Delta y, \quad (\text{B.7})$$

que é o valor da variação do campo elétrico próximo à interface óxido-semicondutor, decorrente da flutuação na concentração de dopantes δN_a , que ocorre na posição y .

A lei de *Gauss* estabelece que a integração de todo o campo elétrico que atravessa perpendicularmente uma superfície fechada é proporcional à carga líquida equivalente envolvida por essa superfície. Assim, se considerarmos o volume da região de depleção envolto por uma superfície fechada, (B.7) corresponde à variação no campo elétrico que atravessa essa superfície, próximo à interface óxido-semicondutor, em função da flutuação δN_a ocorrida. Então, pode-se afirmar que a flutuação equivalente na carga de depleção percebida na interface, decorrente apenas da flutuação local na concentração de impurezas, é dada por

$$\delta(\Delta Q'_{IMP}) = \varepsilon_{Si} \delta E_a = q \delta N_a \left(1 - \frac{y}{y_d} \right) \Delta y, \quad (\text{B.8})$$

e cuja variância pode ser dada por

$$\sigma^2[\delta(\Delta Q'_{IMP})] = q^2 \sigma^2(\delta N_a) \left(1 - \frac{y}{y_d} \right)^2 \Delta y^2. \quad (\text{B.9})$$

A relação entre a flutuação na concentração média de dopantes que ocorre dentro da fatia considerada, e a flutuação no número de impurezas (n_a) que ocorre neste volume é

$$\delta N_a = \frac{\delta n_a}{W\Delta x\Delta y}, \quad (\text{B.10})$$

de modo que sua variância pode ser escrita como

$$\sigma^2(\delta N_a) = \frac{\sigma^2(\delta n_a)}{(W\Delta x\Delta y)^2}. \quad (\text{B.11})$$

Considerando-se que a distribuição de impurezas segue uma distribuição de Poisson, condição geralmente utilizada na modelagem do substrato e já comprovada experimentalmente [20], [21], pode-se afirmar que

$$\sigma^2(\delta n_a) = \overline{\delta n_a} = n_a = W\Delta x\Delta y N_a, \quad (\text{B.12})$$

que em conjunto com (B.11) e (B.9), resulta

$$\sigma^2[\delta(\Delta Q'_{IMP})] = \frac{q^2}{W\Delta x} N_a \left(1 - \frac{y}{y_d}\right)^2 \Delta y. \quad (\text{B.13})$$

Assim, a variância da carga de depleção equivalente percebida na interface, em função do somatório dos efeitos de todas as flutuações microscópicas que ocorrem em fatias infinitesimais sobrepostas, no sentido da profundidade da região de depleção, pode ser expressa como

$$\sigma^2(\Delta Q'_{IMP}) = \sum_y \sigma^2[\delta(\Delta Q'_{IMP})] = \frac{q^2}{W\Delta x} \int_0^{y_d} N_a \left(1 - \frac{y}{y_d}\right)^2 dy. \quad (\text{B.14})$$

B.2 Tensão de limiar

A tensão de limiar V_T de um transistor *MOS* é classicamente definida como a tensão que deve ser aplicada entre a *porta* e o *substrato* (V_{GB}), de modo que o potencial de superfície na interface óxido-semicondutor seja $\phi_S = 2\phi_F$, onde ϕ_F representa o nível de Fermi dos portadores majoritários.

Sob essa condição de ϕ_S constante, uma flutuação na concentração de dopantes δN_a , que ocorra na posição y , provoca uma variação do campo elétrico dentro do óxido δE_{ox} , que é proporcional à variação do campo elétrico próximo à interface do semicondutor δE_a . Esta variação é percebida no terminal de *porta* como uma variação na tensão de limiar δV_T , que com o auxílio de (B.7) pode ser expressa como

$$\delta V_T = t_{ox} \delta E_{ox} = t_{ox} \frac{\varepsilon_{Si}}{\varepsilon_{ox}} \delta E_a = \frac{\varepsilon_{Si}}{C'_{ox}} \delta E_a = \frac{q}{C'_{ox}} \delta N_a \left(1 - \frac{y}{y_d} \right) \Delta y, \quad (\text{B.15})$$

onde ε_{ox} é a permissividade elétrica do óxido.

Utilizando-se as mesmas considerações anteriores, sobre o comportamento estatístico da flutuação na concentração de dopantes, a variância da tensão de limiar, decorrente do somatório dos efeitos de todas as flutuações microscópicas que ocorrem em fatias infinitesimais sobrepostas, no sentido da profundidade da região de depleção, pode ser expressa como

$$\sigma^2(V_T) = \left(\frac{q}{C'_{ox}} \right)^2 \frac{1}{WL} \int_0^{y_d} N_a \left(1 - \frac{y}{y_d} \right)^2 dy = \left(\frac{q}{C'_{ox}} \right)^2 \frac{1}{WL} N_{oi}, \quad (\text{B.16})$$

onde N_{oi} é uma simplificação da integral, sendo o principal parâmetro do modelo de descasamento aqui proposto, e que leva em conta a influência do perfil vertical da concentração de impurezas na carga de depleção. Nesta estimativa considera-se todo o comprimento do transistor, ou $\Delta x = L$.

Este desenvolvimento desprezou os efeitos da flutuação na concentração de dopantes sobre o nível de Fermi, cujo impacto na tensão de limiar é muito pequeno [116], se comparado a (B.16).

Anexo C – Publicações

Ao longo das atividades de pesquisa realizadas durante o período de doutoramento, obteve-se a publicação dos artigos científicos listados a seguir.

Artigos em revistas internacionais:

- C. Galup-Montoro, M. C. Schneider, H. Klimach, and A. Arnaud, “A compact model of MOSFET mismatch for circuit design”, *IEEE Journal of Solid-State Circuits*, vol. 40, n. 8, pp. 1649 – 1657, Aug. 2005.
- H. Klimach, A. Arnaud, C. Galup-Montoro, and M.C. Schneider “MOSFET mismatch modeling: a new approach”, *IEEE Design & Test of Computers*, vol. 23, n. 1, pp. 20 – 29, Jan.-Feb. 2006.

Artigos em congressos internacionais:

- C. Galup-Montoro, M. C. Schneider, A. Arnaud and H. Klimach, “Self-consistent models of DC, AC, noise and mismatch for the MOSFET”, *Proc. Workshop on Compact Modeling, Nanotechnology Conference and Trade Show, Nanotech 2004*, vol. 2, pp. 494-499, Feb. 2004.
- H. Klimach, A. Arnaud, M. C. Schneider and C. Galup-Montoro, “Consistent model for drain current mismatch in MOSFETs using the carrier number fluctuation theory”, *Proc. IEEE International Symposium on Circuits and Systems, ISCAS 2004*, vol. 5, pp. 113-116, Mar. 2004.
- C. Galup-Montoro, M. C. Schneider, A. Arnaud and H. Klimach, “All-region MOS model of mismatch due to random dopant placement”, *Proc. Workshop on Compact Modeling, Nanotechnology Conference and Trade Show, Nanotech 2005*, pp. 127 - 130, Feb. 2005.
- H. Klimach, C. Galup-Montoro, and M. C. Schneider, “A setup for automatic MOSFET mismatch characterization under a wide bias range”, *Proc. Workshop on Compact Modeling, Nanotechnology Conference and Trade Show, Nanotech 2007*, v. 3. p. 662-665, San Jose, 2007.
- C. Galup-Montoro , M. C. Schneider , A. I. A. Cunha , F. Rangel de Sousa , H. Klimach, and O. Franca Siebel, “The advanced compact MOSFET (ACM) model for circuit analysis and design”, *Proc. IEEE Custom Integrated Circuits Conference, CICC 2007*, pp. 519-526, San Jose, Sept. 2007.
- H. Klimach, M. C. Schneider, and C. Galup-Montoro, “An M - $2M$ digital-to-analog converter design methodology based on a physical mismatch model”, *IEEE International Symposium on Circuits and Systems, ISCAS 2008*, Seattle, May 2008.

Artigos em congressos nacionais:

- H. Klimach, A. Arnaud, M. C. Schneider, and C. Galup-Montoro, “Characterization of MOS transistor current mismatch” *Proc. Symposium on Integrated Circuits and Systems Design, SBCCI 2004*, pp. 33-38, Galinhas, Sept. 2004.
- H. Klimach, M. C. Schneider, and C. Galup-Montoro, “A test chip for automatic MOSFET mismatch characterization” *Proc. Symposium on Integrated Circuits and Systems Design, SBCCI 2006*, pp. 83-88, Ouro Preto, Sept. 2006.

Outras publicações:

Este trabalho foi adaptado e incluído como capítulo do livro:

- C. Galup-Montoro and M. C. Schneider *MOSFET Modeling for Circuit Analysis and Design*, World Scientific, 2007.

Publicações periféricas, mas relacionadas ao assunto:

- J. P. Martinez Brito, S. Bampi, and H. Klimach, “A design methodology for matching improvement in bandgap references”, *Proc. IEEE International Symposium on Quality Electronic Design, ISQED 2007*, pp. 586 – 594, March 2007.
- J. P. Martinez Brito, S. Bampi, and H. Klimach, “A 4-bits trimmed CMOS bandgap reference with an improved matching modeling design”, *Proc. IEEE International Symposium on Circuits and Systems, ISCAS 2007*, pp. 1911 – 1914, May 2007.

REFERÊNCIAS

- [1] B. Razavi “CMOS technology characterization for analog and RF design,” *IEEE J. Solid-State Circuits*, vol. 34, no. 3, pp. 268 – 276, March 1999.
- [2] R. W. Keyes “Effect of randomness in the distribution of impurity ions on FET thresholds in integrated electronics,” *IEEE J. Solid-State Circuits*, pp. 245-247, Aug. 1975.
- [3] B. Hoeneisen and C. A. Mead “Fundamental limitations in microelectronics - I. MOS technology,” *Solid-State Electron.*, vol. 15, pp. 819-829, Aug. 1972.
- [4] J.L. McCreary “Matching properties, and voltage and temperature dependence of MOS capacitors,” *IEEE J. Solid-State Circuits*, vol. 16, pp. 608 – 616, Dec 1981.
- [5] J-B. Shyu, G. C. Temes, and K. Yao “Random errors in MOS capacitors,” *IEEE J. Solid-State Circuits*, vol. 17, pp. 1070 – 1076, Dec. 1982.
- [6] J-B. Shyu, G. C. Temes, and F. Krummenacher “Random error effects in matched MOS capacitors and current sources,” *IEEE J. Solid-State Circuits*, vol. 19, no. 6, pp. 948 – 955, Dec. 1984.
- [7] K. R. Lakshmikumar, R. A. Hadaway, and M. A. Copeland “Characterization and modeling of mismatch in MOS transistors for precision analog design,” *IEEE J. Solid-State Circuits*, vol. 21 , no. 6 , pp.1057 – 1066, Dec. 1986.
- [8] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers “Matching properties of MOS transistors,” *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433 – 1440, Oct. 1989.
- [9] J. Pineda-Gyvez and H. P. Tuinhout “Threshold voltage mismatch and intra-die leakage current in digital CMOS circuits,” *IEEE J. Solid-State Circuits*, vol. 39, no. 1, pp. 157 – 168, Jan. 2004.
- [10] M.J.M. Pelgrom, H. P. Tuinhout, and M. Vertregt “Transistor matching in analog CMOS applications,” *Technical Digest 1998 International Electron Devices Meeting*, pp. 915-918, 1998.
- [11] D. Burnett, J. Higman, A. Hoefler et alli “Variation in natural threshold voltage of NVM circuits due to dopant fluctuations and its impact on reliability,” *Digest of the International Electron Devices Meeting*, pp. 529 – 532, Dec. 2002.
- [12] A.J. Bhavnagarwala, X. Tang, and J.D. Meindl “The impact of intrinsic device fluctuations on CMOS SRAM cell stability,” *IEEE J. Solid-State Circuits*, vol. 36, no. 4, pp. 658 – 665, April 2001.
- [13] P. Kinget and M. Steyaert “Impact of transistor mismatch on the speed-accuracy-power trade-off of analog CMOS circuits,” *Proc. IEEE Custom Integrated Circuit Conference, CICC 1996*, pp. 333-336, 1996.
- [14] H.P. Tuinhout “Design of matching test structures,” *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 21 – 27, March 1994.
- [15] C. Michael and M. Ismail “Statistical modeling of device mismatch for analog MOS integrated circuits,” *IEEE J. Solid-State Circuits*, vol. 27, no. 2, pp. 154 – 166, Feb. 1992.

- [16] R.W. Gregor "On the relationship between topography and transistor matching in an analog CMOS technology," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 275-282, Feb. 1992.
- [17] H. Su, C. Michael, and M. Ismail "Yield optimization of analog MOS integrated circuits including transistor mismatch," *Proc. IEEE Int. Symp. on Circuits and Systems, ISCAS 1993*, vol. 3, pp. 1801-1804, May 1993.
- [18] E. Felt, A. Narayan, and A. Sangiovanni-Vincentelli "Measurement and Modeling of MOS Transistor Current Mismatch In Analog IC's," *Proc. IEEE Int. Conf. on Computer-Aided Design*, pp. 272 – 277, November 1994.
- [19] K. Nishinohara, N. Shigyo, and T. Wada "Effects of microscopic fluctuations in dopant distributions on MOSFET threshold voltage," *IEEE Trans. Electron Devices*, vol. 39, no. 3, pp. 634-639, March 1992.
- [20] T. Mizuno, J. Okumtura, and A. Toriumi "Experimental study of threshold voltage fluctuations using an 8k MOSFET's array," *Digest of Technical Papers of the Symposium on VLSI Technology*, pp. 41 – 42, May 1993.
- [21] T. Mizuno, J. Okumtura, and A. Toriumi "Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFET's," *IEEE Trans. Electron Devices*, vol. 41, no. 11, pp. 2216-2221, Nov. 1994.
- [22] T. Mizuno "Influence of statistical spacial-nonuniformity of dopant atoms on threshold voltage in a system of many MOSFET's," *Jpn. J. Appl. Physics*, vol. 35, pp. 842-848, 1996.
- [23] M. Steyaert, J. Bastos, R. Roovers, P. Kinget, W. Sansen, B. Graindourze, A. Pergoot, and E. Janssens, "Threshold voltage mismatch in short-channel MOS transistors," *Electronics Letters*, vol. 30, no. 18, pp. 1546-1548, Sept. 1994.
- [24] H. Elzinga "On the impact of spatial parametric variations on MOS transistor mismatch," *Proc. IEEE Int. Conf. on Microelectronic Test Structures*, pp. 173 – 177, March 1996.
- [25] P. A. Stolk and D. B. M. Klaassen "The effect of statistical dopant fluctuations on MOS device performance," *Proc. Int. Electron Devices Meeting*, pp. 627 - 630, 1996.
- [26] W. Shyh-Chyi, P. Kuo-Hua, and M. Dye-Jyun "A CMOS mismatch model and scaling effects," *IEEE Electron Device Letters*, vol. 18, no. 6, pp. 261-263, June 1997.
- [27] S. J. Lovett, L. Wall, M. Welten, A. Mathewson, and B. Mason "Sensitivity of MOS transistor mismatch to device dimensions and suggestions on how to improve matching performance," *Proc. IEEE Colloquium on Improving the Efficiency of IC Manufacturing Technology*, pp.11/1 - 11/5, 1995.
- [28] J. Bastos, M. Steyaert, A. Pergoot, and W. Sansen "Mismatch characterization of submicron MOS transistors," *Analog Integrated Circuits and Signal Processing*, vol. 12, pp. 95-106, 1997.
- [29] X. Tang, V.K. De, and J.D. Meindl "Intrinsic MOSFET parameter fluctuations due to random dopant placement," *IEEE Trans. on Very Large Scale Integration Systems*, vol. 5, no. 4, Dec. 1997.
- [30] Y. Yasuda, M. Takamiya, and T. Hiramoto "Separation of effects of statistical impurity number fluctuations and position distribution on Vth fluctuations in scaled MOSFET's," *IEEE Trans. Electron Devices*, vol. 47, no. 10, pp. 1838-1842, Oct. 2000.
- [31] T. Tanaka, T. Usuki, T. Futatsugi et alli "Direct measurement of Vth fluctuation caused by impurity positioning," *Digest of Technical Papers of the Symposium on VLSI Technology*, pp. 136 – 137, June 2000.

- [32] D.J. Frank and H.-S.P. Wong "Simulation of stochastic doping effects in Si *MOSFET's*," *International Workshop on Computational Electronics*, Book of Abstracts. pp. 2 – 3, May 2000.
- [33] T. Tanaka, T. Usuki, T. Futatsugi et alli "V_{th} fluctuation induced by statistical variation of pocket dopant profile," *Proc. Int. Electron Devices Meeting*, pp. 271 - 274, Dec. 2000.
- [34] R. Rios, Wei-Kai Shih, A. Shah et alli "A three-transistor threshold voltage model for halo processes," *Proc. Int. Electron Devices Meeting*, pp. 113 - 116, Dec. 2000.
- [35] J. McGinley, O. Noblanc, C. Julien et alli "Impact of pocket implant on *MOSFET* mismatch for advanced CMOS technology," *Proc. IEEE Int. Conference on Microelectronic Test Structures*, pp. 123-126, March 2004.
- [36] A. Asenov and S. Saini "Polysilicon gate enhancement of the random dopant induced threshold voltage fluctuations in sub-100 nm *MOSFETs* with ultrathin gate oxide," *IEEE Trans. Electron Devices*, vol. 47, no. 4, pp. 805-812, April 2000.
- [37] R. Difrenza, J. C.Vildeuil, P. Llinares, and G. Ghibaudo, "Impact of grain number fluctuations in the *MOS* transistor gate on matching performance," *Proc. IEEE Int. Conference on Microelectronic Test Structures*, pp. 244-249, 2003.
- [38] K. Takeuchi "Channel size dependence of dopant-induced threshold voltage fluctuation," *Digest of Technical Papers of the Symposium on VLSI Technology*, pp. 72 – 73, June 1998.
- [39] T. Mizuno "Novel statistical fluctuation of dopant concentration and its influence on scaled *MOS* device performance," *Int. Workshop on Statistical Metrology*, pp. 16 – 19, June 1997.
- [40] C. Lallement, C. Enz, and M. Bucher "Simple solutions for modeling the non-uniform substrate doping," *IEEE Int. Symp. on Circuits and Systems*, vol. 4, pp. 436-439, May 1996.
- [41] K. Takeuchi, T. Tatsumi, and A. Furukawa "Channel engineering for the reduction of random-dopant-placement-induced threshold voltage fluctuation," *Int. Electron Devices Meeting, Technical Digest*, pp. 841-844, 1997.
- [42] F. Forti and M. E. Wright "Measurements of *MOS* current mismatch in the weak inversion region," *IEEE J. Solid-State Circuits*, vol. 29, no. 2, pp. 138-142, Feb. 1994.
- [43] T. Serrano-Gotarredona and B. Linares-Barranco "CMOS transistor mismatch model valid from weak to strong inversion," *Proc. IEEE European Conference on Solid-State Circuits*, pp. 627-630, 2003.
- [44] M. J. Chen, J. S. Ho, and T. H. Huang "Dependence of current match on back-gate bias in weakly inverted *MOS* transistor and its modeling," *IEEE J. Solid-State Circuits*, vol. 31, no. 2, pp. 259-262, Feb. 1996.
- [45] M. J. Chen, J. S. Ho, and D. Y. Chang "Optimizing the match in weakly inverted *MOSFETs* by gated lateral bipolar action," *IEEE Trans. Electron Devices*, vol. 43, no. 5, pp. 766-773, May 1996.
- [46] R. Difrenza, P. Llinares, E. Granger, H. Brut, and G. Ghibaudo "Effect of substrate voltage and oxide thickness on NMOSFET matching characteristics for a 0.18 μm CMOS technology," *Proc. IEEE Int. Conference on Microelectronic Test Structures*, pp. 7-10, 2001.
- [47] P.G. Drennan and C.C. McAndrew "A comprehensive *MOSFET* mismatch model," *Proc. Int. Electron Devices Meeting*, pp. 167 - 170, Dec. 1999.
- [48] P.G. Drennan and C.C. McAndrew "Understanding *MOSFET* mismatch for analog design," *IEEE J. Solid-State Circuits*, vol. 38, no. 3, pp. 450-456, Mar. 2003.

- [49] M-F. Lan and R. Geiger “Impact of model errors on predicting performance of matching-critical circuits,” *Proc. IEEE Midwest Symp. on Circuits and Systems*, pp. 1324-1328, 2000.
- [50] M-F. Lan and R. Geiger “Modeling of random channel parameter variations in MOS transistors,” *Proc. IEEE Int. Symp. on Circuits and Systems*, vol. I, pp. 85-88, 2001.
- [51] G. Lannaccone and E. Amirante “Quantum and semiclassical modeling of the threshold voltage dispersion due to random dopants in deep submicron MOSFET’s,” *Proc. IEEE Conference on Nanotechnology*, pp. 26 – 28, Aug. 2002.
- [52] R. Difrenza, P. Llinares, and G. Ghibaudo “The impact of short channel and quantum effects on the MOS transistor mismatch,” *Solid-State Electronics*, vol. 7, no. 7, pp. 1161-1165, July 2003.
- [53] D. Vasileska and S. Ahmed “Narrow-width SOI devices: the role of quantum-mechanical size quantization effect and unintentional doping on the device operation,” *IEEE Trans. Electron Devices*, vol. 52, no. 2, pp. 227 - 236, Feb. 2005.
- [54] J.A Croon, M. Rosmeulen, S. Decoutere, W. Sansen, and H. E. Maes “An easy-to-use mismatch model for the MOS transistor,” *IEEE J. Solid-State Circuits*, vol. 37, no. 8, pp. 1056-1064, Aug. 2002.
- [55] H. Yang, V. Macary, J. L. Huber, Won-Gi Min, B. Baird, and Jiangkai Zuoet “Current mismatch due to local dopant fluctuations in MOSFET channel,” *IEEE Trans. Electron Devices*, vol. 50, no. 11, pp. 2248-2254, Nov. 2003.
- [56] Shiyong Xiong; J. Bokor, Qi Xiang, P. Fisher, O. Dudley, Paula Rao, and Haihong Wang “Is gate line edge roughness a first-order issue in affecting the performance of deep sub-micro bulk MOSFET devices,” *IEEE Trans. Semiconductor Manufacturing*, vol. 17, no. 3, pp. 357–361, Aug. 2004.
- [57] A. Asenov, S. Kaya, and A. R. Brown “Intrinsic parameter fluctuations in decananometer MOSFETs introduced by gate line edge roughness,” *IEEE Trans. Electron Devices*, vol. 50, no. 5, pp. 1254–1260, May 2003.
- [58] H. Fukutome, Y. Momiyama, T. Kubo, Y. Tagawa, T. Aoyama, and H. Arimoto “Direct evaluation of gate line edge roughness impact on extension profiles in sub-50-nm n-MOSFET’s,” *IEEE Trans. Electron Devices*, vol. 53, no. 11, pp. 2755-2763, Nov. 2006.
- [59] Shiyong Xiong and J. Bokor “A simulation study of gate line edge roughness effects on doping profiles of short-channel MOSFET devices,” *IEEE Trans. Electron Devices*, vol. 51, no. 2, pp. 228 - 232, Feb. 2004.
- [60] N. Gunther, E. Hamadeh, D. Niemann, I. Pesic, and M. Rahman “Modeling intrinsic fluctuations in decananometer MOS devices due to gate line edge roughness (LER),” *Proc. IEEE Int. Symp. on Quality of Electronic Design, ISQED 2005*, pp. 510 – 515.
- [61] J. Croon, W. Sansen, and H. Maes *Matching properties of deep-submicron MOS transistors*, Springer, Dordrecht, 1st ed., 2005.
- [62] Y. Ashizawa and H. Oka “Effect of discrete dopant distribution on MOSFETs scaling into the future,” *Proc. IEEE Int. Conf. on Simulation of Semiconductor Processes and Devices, SISPAD 2005*, pp. 31 – 34, Sept. 2005.
- [63] Daihyun Lim, Jonghae Kim, J. Plouchart, Daeik Kim, Choongyeun Cho, and D. Boning “Performance and yield optimization of mm-wave PLL front-end in 65nm SOI CMOS,” *Proc. IEEE Int. Symp. on Radio Frequency Integrated Circuits, RFIC 2007*, pp. 525 – 528, June 2007.

- [64] C. Galup-Montoro, M.C. Schneider, H. Klimach, and A. Arnaud “A compact model of *MOSFET* mismatch for circuit design,” *IEEE J. Solid-State Circuits*, vol. 40, no. 8, pp. 1649-1657, Aug. 2005.
- [65] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro “An *MOS* transistor model for analog circuit design,” *IEEE J. Solid-State Circuits*, vol.33, no.10, pp. 1510-1519, Oct.1998.
- [66] C. Galup-Montoro, M. C. Schneider, and A. I. A. Cunha “A current-based *MOSFET* model for integrated circuit design,” Chapter 2 in *Low-Voltage/Low-Power Integrated Circuits and Systems*, edited by E. Sánchez-Sinencio and A. Andreou, New Jersey: IEEE Press, 1998.
- [67] Y. P. Tsividis *Operation and Modeling of the MOS Transistor*, Oxford University Press, 2nd ed., 1999.
- [68] Y. Byun, K. Lee, and M. Shur “Unified charge control model and subthreshold current in heterostructure field effect transistors,” *IEEE Electron Device Letters*, vol. 11, no. 1, pp. 50-53, Jan. 1990.
- [69] V. Ambegaokar *Reasoning about Luck: Probability and its Uses in Physics*, Cambridge University Press, 1996.
- [70] H. Veendrick *Deep-Submicron CMOS ICs*, Kluwer, Dordrecht, 2nd ed., 2000.
- [71] A. Asenov “Random dopant induced lowering and fluctuations in sub-0.1 μm *MOSFETs*: a 3-D *atomistic* simulation study,” *IEEE Transactions on Electron Devices*, vol. 45, no. 12, pp. 2505 – 2513, Dec. 1998.
- [72] G. Reimbold “Modified *1/f* trapping noise theory and experiments in *MOS* transistors biased from weak to strong inversion – influence of interface states,” *IEEE Trans. Electron Devices*, vol. 31, no. 9, pp. 1190-1198, Sept. 1984.
- [73] F. Berz “Theory of low frequency noise in Si *MOSTs*,” *Solid-State Electronics*, no. 13, pp. 631-647, 1970.
- [74] R. Wilson “The dirty little secret: Engineers at design forum vexed by rise in process variations at the die level,” *EE Times*, p. 1, Mar. 25, 2002.
- [75] J. A. Croon, H. P. Tuinhout, R. Difrenza, J. Knol, A. J. Moonen, S. Decoutere, H. E. Maes, and W. Sansen “A comparison of extraction techniques for threshold voltage mismatch,” *Proc. IEEE Int. Conf. Microelectronic Test Structures*, 2002, pp. 235–240.
- [76] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S.-H. Lo, G. A. Sai-Halasz, R. G. Viswanathan, H.-J. C. Wann, S. J. Wind, and H.-S. Wong “*CMOS* scaling into the nanometer regime,” *Proc. of the IEEE*, vol. 85, no. 4, pp. 486–504, Apr. 1997.
- [77] Y. Taur “An analytical solution to a double-gate *MOSFET* with undoped body,” *IEEE Electron Device Lett.*, vol. 21, no. 5, pp. 245–247, May 2000.
- [78] S. Cristensson, I. Lundstrom, and C. Svensson “Low frequency noise in *MOS* transistors,” *Solid-State Electron.*, vol. 11, pp. 797–812, 1968.
- [79] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng “A physics-based *MOSFET* noise model for circuit simulators,” *IEEE Trans. Electron Devices*, vol. 37, pp. 1323-1333, May 1990.
- [80] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro “Derivation of the unified charge control model and parameter extraction procedure,” *Solid-State Electronics*, vol. 43, no. 3, pp. 481-485, Mar. 1999.
- [81] A. Arnaud and C. Galup-Montoro “A compact model for flicker noise in *MOS* transistors for analog circuit design,” *IEEE Trans. Electron Devices*, vol. 50, no. 8, pp. 1815-1818, Aug. 2003.

- [82] C. C. Enz, F. Krummenacher, and E. A. Vittoz “An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications”, *Analog Integrated Circuits Signal Process.*, vol. 8, pp. 83-114, 1995.
- [83] M. Quarantelli, S. Saxena, N. Dragone et alli “Characterization and modeling of MOSFET mismatch of a deep submicron technology,” *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 238 – 243, March 2003.
- [84] M. J. M. Pelgrom “Low-power CMOS data conversion,” Chap. 14 in *Low-Voltage/Low-Power Intergrated Circuits and Systems*, ed. E. Sánchez-Sinencio, A. G. Andreou, IEEE Press, New York, 1999.
- [85] D. Foty *MOSFET modeling with SPICE*, Prentice Hall PTR, Upper Saddle River, 1997.
- [86] M. Pelgrom and M. Vertregt “CMOS technology for mixed signal ICs,” *Solid-State Electronics*, vol. 41, no. 7, pp. 967-974, July 1997.
- [87] W. Shyh-Chyi, P. Kuo-Hua, and M. Dye-Jyun “A CMOS mismatch model and scaling effects,” *IEEE Electron Device Letters*, vol. 18 , no. 6 , pp. 261-263, June 1997.
- [88] R. Difrenza, P. Llinares, S. Taupin et alli “Comparison between matching parameters and fluctuations at the wafer level,” *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 241 – 246, April 2002.
- [89] C. Galup-Montoro, M. C. Schneider, A. Arnaud, and H. Klimach “Self-consistent models of DC, AC, noise and mismatch for the MOSFET,” *Proc. Nanotechnology Conference and Trade Show*, vol. 2, pp. 494-499, Feb. 2004.
- [90] H. Klimach, A. Arnaud, M. C. Schneider, and C. Galup-Montoro “Consistent model for drain current mismatch in MOSFETs using the carrier number fluctuation theory,” *Proc. IEEE Int. Symp. on Circuits and Systems*, vol. 5, pp. 113-116, Mar. 2004.
- [91] C. Sodini, Ping-Keung Ko, and J. Moll “The effect of high fields on MOS device and circuit performance,” *IEEE Trans. Electron Devices*, vol. 31, no. 10, pp. 1386 - 1393, Oct. 1984.
- [92] C. T. Sah and H. C. Pao “The effects of fixed bulk charge on the characteristics of metal-oxide-semiconductor transistors,” *IEEE Trans. Electron Devices*, vol. 13, no. 4, pp. 393 - 409, Apr. 1966.
- [93] W. Shockley, J. A. Copeland, and R. P. James “The impedance field method of noise calculation in active semiconductor devices,” in *Quantum Theory of Atoms, Molecules and the Solid-State*. Academic, New York, pp. 537-563, 1966.
- [94] C. Galup Montoro and M. C. Schneider *MOSFET modeling for circuit analysis and design*, World Scientific, 1st ed., 2007.
- [95] International Technology Roadmap for Semiconductors, Process Integration, Devices and Structures chapter, 2006 update: <http://www.itrs.net>.
- [96] M. Conti, P. Crippa, F. Fedecostante et alli “A modular test structure for CMOS mismatch characterization,” *Proc. Int. Symp. on Circuits and Systems*, vol. 5, pp. V-569 – V-572, May 2003.
- [97] M. Conti, P. Crippa, S. Orcioni et alli “A new test structure for short and long distance mismatch characterization of submicron MOS transistors,” *Proc. IEEE Midwest Symp. on Circuits and Systems*, vol. 2 , pp. 656 – 660, Aug. 2001.
- [98] J. Einfeld, U. Schaper, U. Kollmer et alli “A new test circuit for the matching characterization of npn bipolar transistors,” *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 127 – 131, March 2004.
- [99] S.B. Yeo, J. Bordelon, S. Chu et alli “A robust and production worthy addressable array architecture for deep sub-micron MOSFETs matching characterization,” *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 229 – 234, April 2002.

- [100] Y. Ta-Hsun, J.C.H. Lin, S.-C. Wong et alli "Mismatch characterization of 1.8 V and 3.3 V devices in 0.18 μm mixed signal CMOS technology," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 77 – 82, March 2001.
- [101] U. Schaper, C. Linnenbank, and R. Thewes "A novel approach for precise characterization of long distance mismatch of CMOS-devices," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 148 – 152, March 2000.
- [102] U. Schaper, C.G. Linnenbank, and R. Thewes "Precise characterization of long-distance mismatch of CMOS devices," *IEEE Trans. on Semiconductor Manufacturing*, vol. 14 , no. 4 , pp. 311 – 317, Nov. 2001.
- [103] L. Portmann, C. Lallement, and F. Krummenacher "A high density integrated test matrix of MOS transistors for matching study," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 19 – 24, March 1998.
- [104] H. Elzinga "Using test structures to assess the impact of critical process steps on MOS transistor matching," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 119 – 122, March 1998.
- [105] H.P. Tuinhout and M. Vertregt "Test structures for investigation of metal coverage effects on MOSFET matching," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 179 – 183, March 1997.
- [106] H.P. Tuinhout and M. Vertregt "Characterization of systematic MOSFET current factor mismatch caused by metal CMP dummy structures," *IEEE Trans. on Semiconductor Manufacturing*, vol. 14 , no. 4 , pp. 302 – 310, Nov. 2001.
- [107] H.P. Tuinhout, A. Bretveld, and W.C.M. Peters "Measuring the span of stress asymmetries on high-precision matched devices," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 117 – 122, March 2004.
- [108] K. Terada and M. Eimitsu "A test circuit for measuring MOSFET threshold voltage mismatch," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 227 – 231, March 2003.
- [109] K. Terada and K. Fukeda "Further study of V_{TH} -mismatch evaluation circuit," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 155 – 159, March 2004.
- [110] A. Pergoot, B. Graindourze, E. Janssens et alli "Statistics for matching," *Proc. Int. Conf. on Microelectronic Test Structures*, pp. 193 – 197, March 1995.
- [111] H. Klimach, A. Arnaud, M. C. Schneider, and C. Galup-Montoro "Characterization of MOS transistor current mismatch," *Proc. Symp. on Integrated Circuits and Systems Design*, pp. 33-38, Sept. 2004.
- [112] H. Klimach, C. Galup-Montoro, M.C. Schneider, and A. Arnaud "MOSFET mismatch modeling: a new approach," *IEEE Design & Test of Computers*, vol. 23, no. 1, pp 20-29, Jan. 2006.
- [113] K. Bult and G. Geelen "An inherently linear and compact MOST-only current division technique," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, pp. 1730 - 1735, Dec. 1992.
- [114] C. Hammerschmied and Qiuting Huang "Design and implementation of an untrimmed MOSFET-only 10-bit A/D converter with -79-dB THD," *IEEE J. Solid-State Circuits*, vol. 33, no. 8, pp. 1148 - 1157, Aug. 1998.
- [115] A. Roy and C. Enz "Compact modeling of thermal noise in the MOS transistor," *IEEE Trans. Electron Devices*, vol. 52, no. 4, pp. 611 - 614, Apr. 2005.
- [116] T. Mizuno "New channel engineering for sub-100 nm MOS devices considering both carrier velocity overshoot and statistical performance fluctuations," *IEEE Trans. Electron Devices*, vol. 47, no. 4, pp. 756-761, Apr. 2000.