

Escalamento e Limites dos Dispositivos MOS

Jacobus W. Swart e Marcelo A. Pavanello
CCS e FECC - UNICAMP

1. Introdução

Desde o advento dos circuitos integrados (CI's), no início da década de 60, uma contínua redução das dimensões tem sido observada. De modo geral, observa-se uma redução pela metade a cada seis anos. Simultaneamente, uma duplicação do tamanho das pastilhas (área do chip) ocorre a cada oito anos. Em adição a estes dois fatos, melhoramentos em leiautes e novas estruturas físicas resultaram em uma evolução quanto à eficiência de empacotamento[1,2].

A combinação destas três evoluções, mencionadas acima, resultou em um incremento muito intenso do número de componentes por pastilha. Este incremento corresponde a aproximadamente 2 vezes/ano até o ano 1972/1973 e 4 vezes/3 anos a partir desta data. A figura 1 ilustra a grande evolução tecnológica evidenciada na fabricação de CI's.

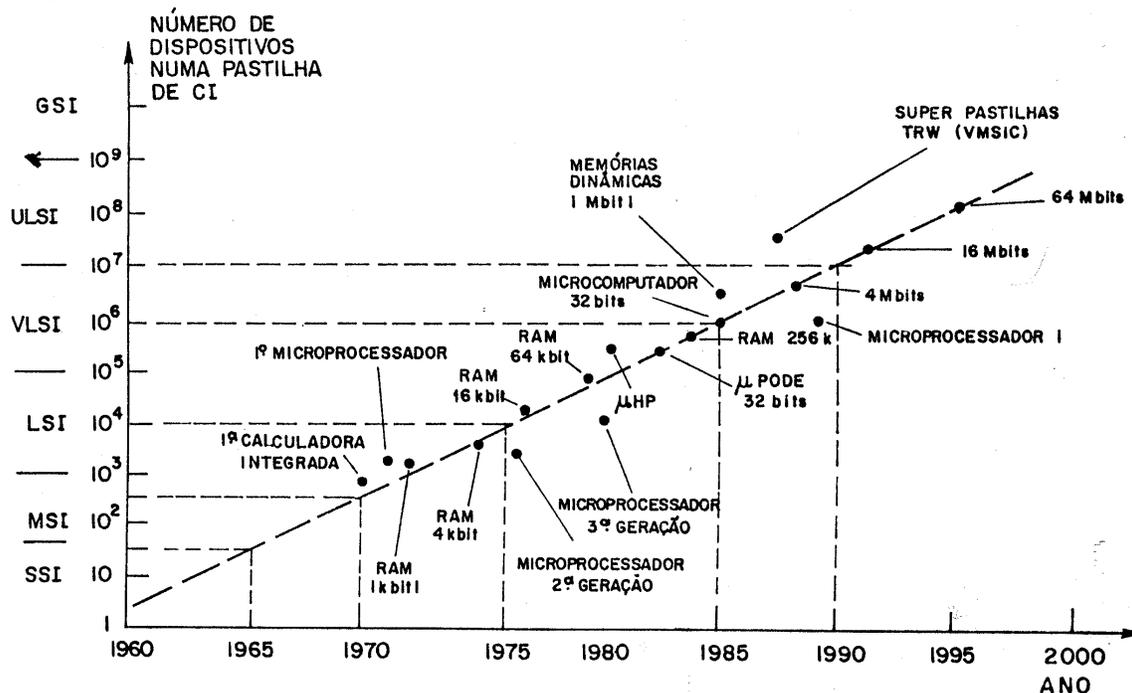


Figura 1 – Evolução do número de dispositivos por pastilha com o passar dos anos.

Esta rápida evolução na tecnologia dos CI's foi, e continua sendo, motivada por fatores econômicos e de desempenho elétrico. Como principais fatores temos:

- i) Considerando, inicialmente, uma mesma função, temos devido apenas à redução das dimensões:
 - a) Maior densidade, e portanto, maior número de pastilhas por lâminas. Como numa área de pastilha menor devemos ter um número menor de defeitos, o rendimento de fabricação deverá ser maior;
 - b) O circuito deverá ser mais veloz;
 - c) O circuito consumirá menor potência.
- ii) Por outro lado, considerando pastilhas com maior número de componentes, um mesmo sistema eletrônico pode ser fabricado com menor número de pastilhas. Isto resulta em um menor custo de montagem do sistema, menor volume e também maior confiabilidade, devido à redução no número de conexões entre as diversas pastilhas [3].

As evoluções mencionadas foram possíveis, e têm sido acompanhadas, pelo surgimento de teorias de escalamento e por pesquisas dos fenômenos limitantes em dispositivos de menores dimensões. A seguir apresentamos algumas das teorias de escalamento reportadas, analisamos várias

limitações em transistores de pequenas dimensões e discutimos os limites de escalamento dos dispositivos MOS.

2. Leis de Escalamento

As leis de escalamento podem ser utilizadas como guia para o projeto de novas gerações tecnológicas, a partir de uma geração testada e em uso. Adicionalmente, estas leis permitem prever o desempenho destas novas gerações e fazer uma análise dos limites do escalamento. Apresentamos em seguida algumas das leis de escalamento reportadas.

2.1 – Escalamento por Campo Elétrico constante

Esta lei foi proposta em 1974 por Dennard *et al.* [4]. Segundo esta lei, reduz-se todas as dimensões e polarizações com um fator de escala k e aumenta-se as concentrações de dopantes com o mesmo fator k_e , conforme mostrado na Tabela 1:

Tabela 1 – Regra de escalamento por campo elétrico constante

Parâmetro	Fator de Escala ($k_e > 1$)
Dimensões L , W , t_{ox} e x_j	$1/k_e$
Concentração de dopantes	k_e
Polarização	$1/k_e$

O fato do campo elétrico não ser alterado pelo escalamento evita variações em efeitos que dependam do campo elétrico. Desta forma, os dispositivos devem conservar características elétricas similares.

Uma análise do impacto do escalamento no desempenho dos dispositivos pode ser feita a partir de modelos básicos para a corrente elétrica que flui entre fonte e dreno (I_{DS}), substituindo-se nestas equações básicas as constantes definidas na Tabela 1:

a) em triodo:

$$I_{DS}' = \frac{\mu \epsilon_{ox} W'}{t_{ox} L'} \left(V_{GS}' - V_T - \frac{V_{DS}'}{2} \right) V_{DS}' \cong \frac{I_{DS}}{k_e} \quad (1)$$

b) em saturação:

$$I_{DS}' = \frac{\mu \epsilon_{ox} W'}{2 t_{ox} L'} (V_{GS}' - V_T)^2 \cong \frac{I_{DS}}{k_e} \quad (2)$$

Desta forma, a corrente elétrica é também escalada por um fator $1/k_e$.

Para o caso da impedância de saída, $R_{on}' = \frac{V_{DS}'}{I_{DS}'} = \frac{V_{DS}/k_e}{I_{DS}/k_e}$ (3), observa-se que a mesma não

é escalada, mantendo-se constante.

A potência dissipada, $P' = V_{DS}' I_{DS}' = \frac{V_{DS}}{k_e} \frac{I_{DS}}{k_e} = \frac{P}{k_e^2}$ (4), sofre um escalamento com $1/k_e^2$.

Por outro lado, a potência por unidade de área não sofre escalamento:

$$\frac{P'}{A'} = \frac{P/k_e}{A/k_e} = \frac{P}{A} \quad (5)$$

Como as dimensões verticais são escaladas com o mesmo fator de escala que as dimensões horizontais, as capacitâncias são escaladas também por um fator $1/k_e$:

$$C' = \epsilon \frac{A'}{d'} = \epsilon \frac{A/k^2}{d/k} = \frac{C}{k} \quad (6)$$

A velocidade de chaveamento torna-se maior com o escalamento, como indica a expressão aproximada para o tempo de atraso (t_a), o qual também é escalado com $1/k_e$:

$$t_a' = \frac{C'V'}{I'} = \frac{C/k \cdot V/k}{I/k} = \frac{t_a}{k} \quad (7)$$

A figura de mérito, produto potência *versus* tempo de atraso, sofre um escalamento $(1/k_e)^3$, como mostra a equação (8):

$$P' \cdot t_a' = \frac{P}{k^2} \cdot \frac{t_a}{k} = \frac{P \cdot t_a}{k^3} \quad (8)$$

Assim, o desempenho elétrico do dispositivo escalado é melhorado, enquanto que a potência por unidade de área permanece inalterada, evitando problemas com a temperatura. A figura 2 ilustra esquematicamente o princípio de escalamento, neste caso promovendo uma redução $k_e=2$, observando-se que as curvas características permanecem idênticas e escaladas:

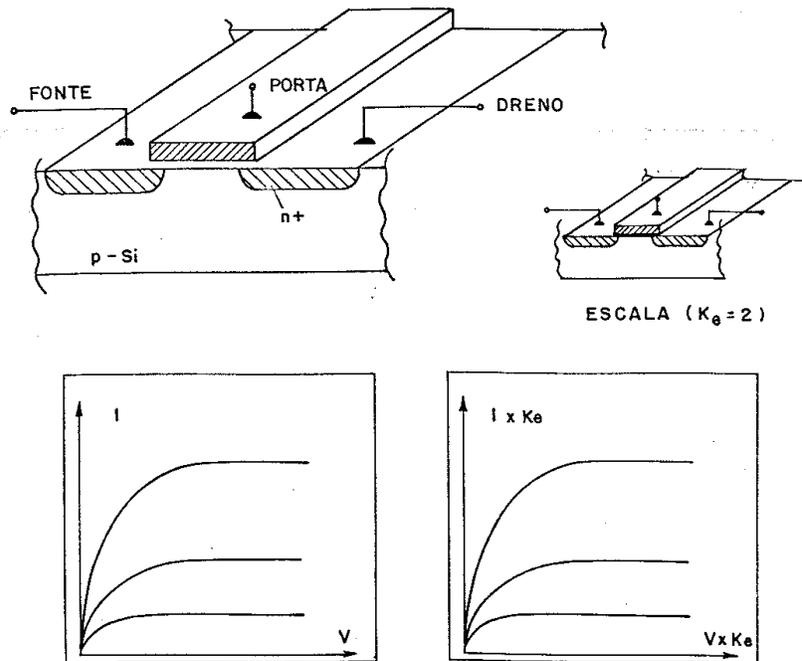


Figura 2 – Representação esquemática do princípio do escalamento.

Entretanto, o escalamento teórico descrito, sofre as seguintes ressalvas:

- 1) A mobilidade (μ) foi considerada constante com o escalamento. Sabe-se, no entanto, que o incremento do nível de dopagem do substrato requerido pelo escalamento causa uma redução na mobilidade [5] e, por conseqüência, na corrente I_{DS} ;

2) As larguras das regiões de depleção (w_D) não são escaladas como previsto para as demais dimensões. Esta discordância deve-se à não escalabilidade do potencial de barreira das junções (V_{Bi}), o qual na verdade eleva-se com o aumento da dopagem:

$$V_{Bi} = \frac{kT}{q} \ln \left(\frac{N_A N_D}{n_i^2} \right) \quad (9)$$

A equação (10) indica o escalamento de w_D , o qual ocorre idealmente apenas se $V \gg V_{Bi}$:

$$w_D' = \sqrt{\frac{2\epsilon_{Si}}{qk_e N} \left(V_{Bi} + \frac{V}{k_e} \right)} \cong \frac{w_D}{k_e} \quad (10)$$

Pode-se, no entanto, contornar o problema aumentando N por um fator maior que k_e , ou ainda, reduzindo-se a temperatura de operação para, por exemplo, a temperatura de nitrogênio líquido (77K), com o intuito de reduzir V_{Bi} . Entretanto, esta última solução é muito radical, pois incrementa significativamente a complexidade de montagem dos equipamentos.

3) A diferença de função trabalho entre metal de porta e semiconductor (ϕ_{ms}) e o potencial de Fermi (ϕ_F) não são escalados, resultando em um escalamento não ideal para a tensão de limiar (V_T):

$$V_T' = \phi_{ms} + 2\phi_F + \frac{t_{ox}}{\epsilon_{ox}} k_e \left[-Q_{ox} + \sqrt{2\epsilon_{Si} q k_e N \left(2\phi_F - \frac{V_{BS}}{k_e} \right)} \right] \quad (11)$$

Este problema pode também ser contornado escalando-se N e/ou t_{ox} com um fator diferente de k_e .

4) A corrente na região de sublimiar não pode ser escalada. Desta forma, o inverso da inclinação da curva de corrente, comumente chamado de inclinação de sublimiar (S), também não é escalado, como indicado na equação (12). Assim, a redução de V_T implicaria numa elevação da corrente de corte (I_0) indesejável. A figura 3 mostra esquematicamente a região de obtenção da inclinação de sublimiar. Nesta figura, a curva tracejada indica a elevação de I_0 provocada pelo escalamento de V_T . Como o valor de I_0 determina a frequência mínima para a restauração de informações (*refresh time*) em CI's dinâmicos e a potência DC em CI's estáticos, elevações neste valor são evitadas. Desta forma, escalar V_T como proposto inicialmente, representa uma séria dificuldade. A solução seria evitá-la enquanto possível, ou assumir compromissos. Uma possível solução para o problema é também a redução da temperatura.

$$S = \frac{\Delta V_{GS}}{\Delta \log(I_{DS})} = \frac{kT}{q} \ln(10) \left(1 + \frac{C_D + C_{it}}{C_{ox}} \right) \quad (12)$$

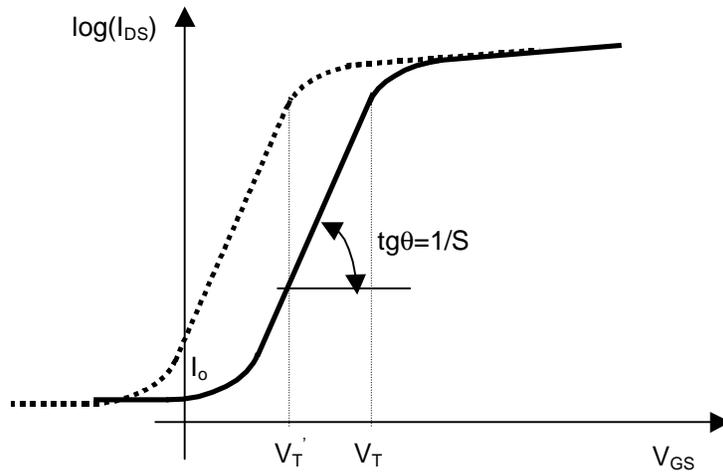


Figura 3 – Comportamento da corrente de sublimar de um transistor MOS.

2.2 – Escalamento por Tensão Constante e por Tensão Quase Constante

Estas duas leis foram apresentadas por Chatterjee *et al.* em 1980 [6]. Nestas leis, as dimensões horizontais e as dopagens são escaladas de forma idêntica à lei de escalamento com campo elétrico constante (EC). Na lei por tensão constante (VC), as polarizações não são escaladas e na lei por tensão quase constante (VQC), as polarizações são escaladas por $\frac{1}{\sqrt{k_e}}$, ou seja, por um fator intermediário entre os casos EC e VC. Com o intuito de não degradar a confiabilidade dos dispositivos por ruptura do óxido de porta, optou-se por escalar a espessura desta por um fator menor no caso VC, de modo que este campo elétrico seja o mesmo nas leis VC e VQC. Na tabela 2 apresentam-se os fatores de escalamento de acordo com o proposto por estas duas leis.

Tabela 2 – Leis de escalamento por tensão constante e por tensão quase constante

Parâmetro	Fator de Escala	
	Tensão constante	Tensão quase constante
Dimensões W, L, x_j	$\frac{1}{k_e}$	$\frac{1}{k_e}$
t_{ox}	$\frac{1}{\sqrt{k_e}}$	$\frac{1}{k_e}$
Concentração de dopantes, N	k_e	k_e
Polarização	1	$\frac{1}{\sqrt{k_e}}$

Estas duas leis tiveram as seguintes motivações para o seu surgimento:

- A não escalabilidade da tensão de limiar e da corrente de sublimar;
- A dificuldade em escalar variações nos parâmetros de processo na mesma proporção do escalamento dos valores dos mesmos parâmetros;
- A necessidade de padronização dos valores de polarização dos CI's;
- Manutenção das margens de ruído apropriadas nos sinais lógicos;
- Manutenção da compatibilidade com outras famílias de CI's, tais como a TTL.

Os parâmetros de desempenho dos dispositivos, escalados segundo estas duas leis, estão comparados ao dos dispositivos escalados segundo a lei por campo elétrico constante na Tabela 3.

Tabela 3 – Comparação do desempenho dos dispositivos escalados segundo as leis EC, VC e VQC.

Parâmetro	Lei de Escalamento		
	EC	VC	VQC
I_{DS}	$1/k_e$	$\sqrt{k_e}$	1
C	$1/k_e$	$1/k_e^{3/2}$	$1/k_e$
t_a (CV/l)	$1/k_e$	$1/k_e^2$	$1/k_e^{3/2}$
P	$1/k_e^2$	$\sqrt{k_e}$	$1/\sqrt{k_e}$
$P \cdot t_a$	$1/k_e^3$	$1/k_e^{3/2}$	$1/k_e^2$
P/A	1	$k_e^{5/2}$	$k_e^{3/2}$

Nota-se que dispositivos escalados segundo a lei VQC apresentam desempenho intermediário ao dos casos de escalamento segundo EC e VC. As leis VC e VQC resultam em CI's mais rápidos que no caso da lei EC, porém o consumo de potência, e de potência por unidade de área aumentam. Conseqüentemente, o escalamento por VC e VQC não poderão ser aplicados com fatores muito elevados.

2.3 – Guia generalizada para miniaturização

Este guia foi proposto por Brews *et al.* em 1980 [7]. Baseados em dados experimentais e de simulação bidimensional de dispositivos, os autores encontraram uma relação empírica (equação 13) entre L_{min} e as espessuras do óxido de porta (t_{ox}), da profundidade de junção (x_j) e das larguras de depleção das junções de fonte e dreno (w_s e w_d , respectivamente). L_{min} é definido como o comprimento de canal para o qual o efeito de canal curto é menor que 10%, sendo que a ocorrência de efeito de canal curto é caracterizada pela variação da corrente de sublimar com $1/L$ e com V_{DD} .

$$L_{min} = A \left[x_j t_{ox} (w_s + w_d)^2 \right]^{1/3} \quad (13)$$

onde A é uma constante empírica.

Desta forma, desejando-se um processo adequado para um dado L_{min} , os parâmetros x_j , t_{ox} , N e V_{DD} são ajustados de forma a satisfazer a equação (13).

2.4 – Teoria generalizada para escalamento

Com base nas considerações apresentadas nas leis VC e VQC e na otimização das características do transistor, Baccarani *et al.* [8] propuseram uma teoria generalizada para escalamento.

De acordo com esta teoria, todas as dimensões são escaladas por um fator $1/k_d$ e as polarizações são escaladas por um fator independente $1/k_v$. A concentração de dopantes por sua vez é aumentada pela relação k_d^2/k_v . A tabela 4 resume esta lei de escalamento.

Tabela 4 – Resumo da teoria generalizada para escalamento

Parâmetro	Fator de Escala
Dimensões W, L, t_{ox} e x_j	$1/k_d$
Polarização	$1/k_v$
Concentração de dopantes, N	k_d^2/k_v

Nota-se que no caso em que $k_d=k_v$ esta lei coincide com a lei de campo elétrico constante. Assim, a lei generalizada engloba a lei EC, mas não as leis VC e VQC.

As distribuições de potencial elétrico, campo elétrico e das concentrações de elétrons e lacunas resultantes do escalamento apresentam intensidades escaladas, mas mantém-se idênticas. Por exemplo, o campo elétrico é escalado com k_d/k_v .

Como consequência, os efeitos dependentes das formas de distribuição do campo elétrico e do potencial elétrico continuam inalterados. Como exemplos destes efeitos temos o efeito de canal curto/DIBL (*Drain Induced Barrier Lowering*) e perfuração MOS (*Punchthrough*).

Pode-se concluir, a partir desta análise, que os dispositivos escalados segundo as leis VC e VQC apresentam alterações nas distribuições do campo elétrico e do potencial elétrico, com degradações nos efeitos mencionados.

O desempenho elétrico dos dispositivos escalados segundo esta lei generalizada, apresenta-se similar ao das leis EC, VC e VQC, dependendo da relação entre k_d e k_v , como indicado na Tabela 5.

Tabela 5 – Desempenho dos dispositivos escalados segundo a lei generalizada

Parâmetro	Fator
E	k_d/k_v
I_{DS}	k_d/k_v^2
P	k_d/k_v^3
P/A	k_d^3/k_v^3
t_a	k_d/k_v^2
$P \cdot t_a$	$\frac{1}{k_d^2 \cdot k_v}$

2.5 – Procedimento prático para o escalamento

As leis até então apresentadas servem como guia para o projeto de novas gerações de processos. Na prática, além do uso destas leis, faz-se o uso intensivo de simuladores (uni, bi e tridimensionais) de processo e de dispositivos. Como exemplos dos simuladores temos SUPREM-IV[9], que é um simulador bidimensional de processos, MEDICI[10] e PISCES[11], que são simuladores bidimensionais de dispositivos e DAVINCI[12], que é um simulador tridimensional de dispositivos.

Em geral, as estruturas geradas pelo simulador de processos, o qual possui modelos para as diversas etapas individuais de processos, são alimentadas aos simuladores de dispositivos, que resolvem bidimensional ou tridimensionalmente as equações da continuidade e de Poisson. Desta forma, incorporando-se as alterações decorrentes do processo de fabricação, nas características elétricas dos dispositivos e uma melhor correlação entre os valores experimentais e os resultados das simulações é obtida.

Por meio destas simulações pode-se otimizar a estrutura do dispositivo, por meio da análise dos seguintes parâmetros e limitações:

- ✓ tensão de limiar
- ✓ efeito de canal curto e DIBL
- ✓ perfuração MOS (*Punchthrough*)
- ✓ corrente de corte (I_0)
- ✓ tempo de atraso
- ✓ potência

✓ corrente de porta e de substrato/confiabilidade

O diagrama de blocos da figura 4 apresenta um procedimento para o projeto de novas gerações de processo escaladas.

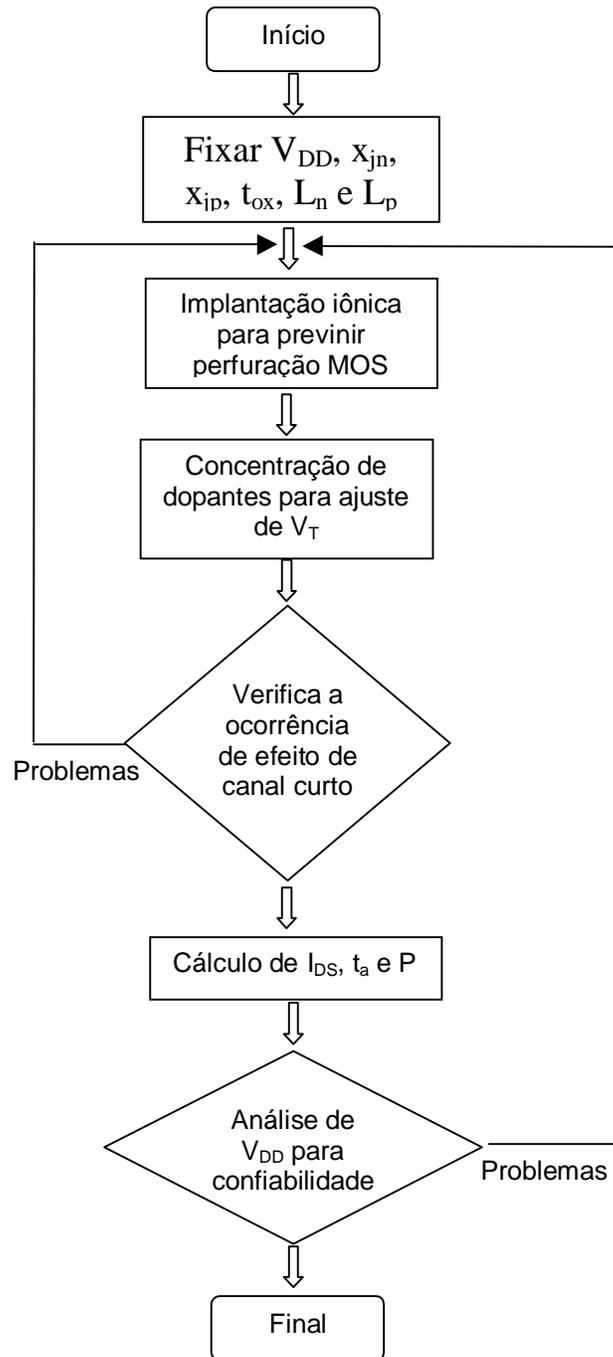


Figura 4 – Exemplo de procedimento para escalamento de uma tecnologia.

3 – Limitações em transistores MOS de pequenas dimensões

A redução das dimensões dos dispositivos faz com que uma série de efeitos secundários tornem-se mais intensos. Como principais efeitos secundários que influenciam o desempenho de transistores de pequenas dimensões temos:

- ✓ efeito de canal curto/DIBL ou $\Delta V_T \times L$ e $\Delta V_T \times V_{DS}$
- ✓ perfuração MOS
- ✓ resistência parasitária de fonte e dreno
- ✓ Capacitância da camada de inversão
- ✓ redução da mobilidade
- ✓ injeção de portadores quentes
- ✓ rupturas
- ✓ efeitos de canal estreito.

Em seguida estes efeitos são apresentados resumidamente.

3.1 – Redução de V_T com L e com V_{DS} (DIBL)

Com a redução das dimensões, a quantidade de carga espacial da região do canal consumida pelas regiões de depleção de fonte e dreno torna-se apreciável em relação a quantidade total de cargas controladas pela porta do transistor. Desta forma, com valores menores de potencial aplicado à porta ocorre a inversão da superfície da região de canal. Como consequência, uma redução no valor da tensão de limiar com comprimentos de canal menores ocorre, como ilustra a figura 5. A velocidade da redução de V_T com $1/L$ depende de t_{ox} , N e x_j . Existem alguns modelos apresentados na literatura para a descrição deste fenômeno, dentre os quais destacamos o de Yau [13], para substrato com concentração uniforme e Nataraj [14], para substrato com perfil de dopagem típico de tecnologias CMOS.

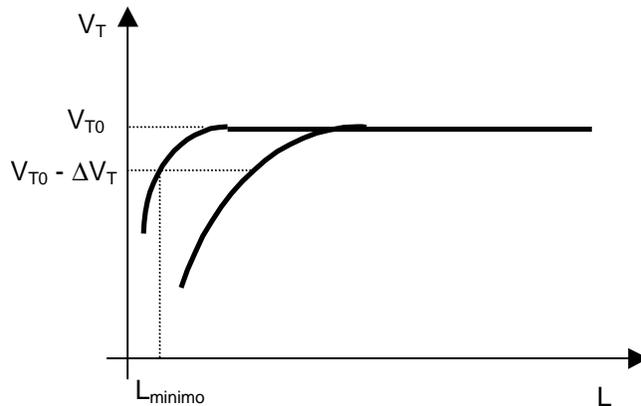


Figura 5 – Ilustração da redução de V_T com a diminuição de L .

No projeto de um processo deve-se ajustar os parâmetros t_{ox} , $N(x)$ e x_j de tal forma a obter uma variação máxima tolerada em V_T (ΔV_T), a partir de uma tensão de limiar inicial V_{T0} . O pior caso inclui V_{DD} máximo e L_{min} , considerando ainda possíveis variações nestes parâmetros.

3.2 – Perfuração MOS (Punchthrough)

A redução do comprimento de canal do transistor pode ocasionar a redução da barreira de potencial entre fonte e canal, induzida pela polarização aplicada ao dreno. Em outras palavras, o aumento da polarização reversa na junção dreno-canal, provoca um aumento da largura de depleção desta junção para o interior do canal. Caso o comprimento de canal seja pequeno, inicia-se uma interação entre as regiões de depleção das junções fonte-canal e dreno-canal, provocando a diminuição da barreira mencionada, na junção fonte-canal. Esta redução causa a injeção de portadores da fonte para o substrato, dando origem a uma corrente parasitária de I_{DS} através do substrato, não controlada pela porta.

Um dispositivo que apresenta a ocorrência de perfuração MOS apresenta um ponto de sela na sua distribuição interna de potencial elétrico e de concentração de portadores, como ilustrado nas figuras 6 A e B, respectivamente. Diferentemente, dispositivos que não sofrem a ocorrência deste efeito apresentam uma distribuição monotônica de potencial elétrico e de concentração de portadores na direção perpendicular à superfície, como ilustrado nas figuras 7 A e B, respectivamente.

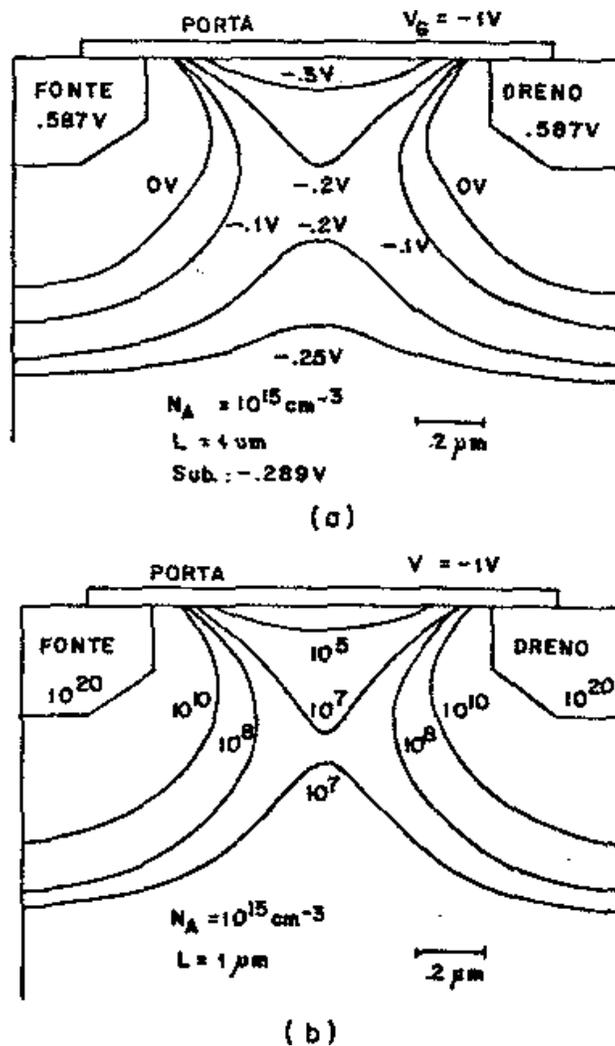


Figura 6 – Perfil do potencial elétrico (A) e da concentração de portadores (B) no equilíbrio, para um dispositivo sofrendo de perfuração MOS.

Uma forma de observar a ocorrência ou não de perfuração MOS é analisar a corrente na região de sublimiar do transistor. Como a corrente de perfuração MOS passa pelo corpo do dispositivo e, portanto, não sofre influência do potencial de porta, a corrente total na região de sublimiar não mais apresenta um comportamento exponencial com V_{GS} , como previsto teoricamente para um dispositivo sem perfuração MOS.

A perfuração MOS pode ser controlada com os seguintes parâmetros do transistor: L , $N(x)$, t_{ox} , x_j e V_{BS} . Estes parâmetros podem ser ajustados até a supressão completa da corrente de perfuração MOS. Em geral, uma implantação iônica com alta energia é realizada com o intuito de elevar a concentração do corpo do transistor, de forma a evitar que a região de depleção do dreno caminhe para o interior do canal.

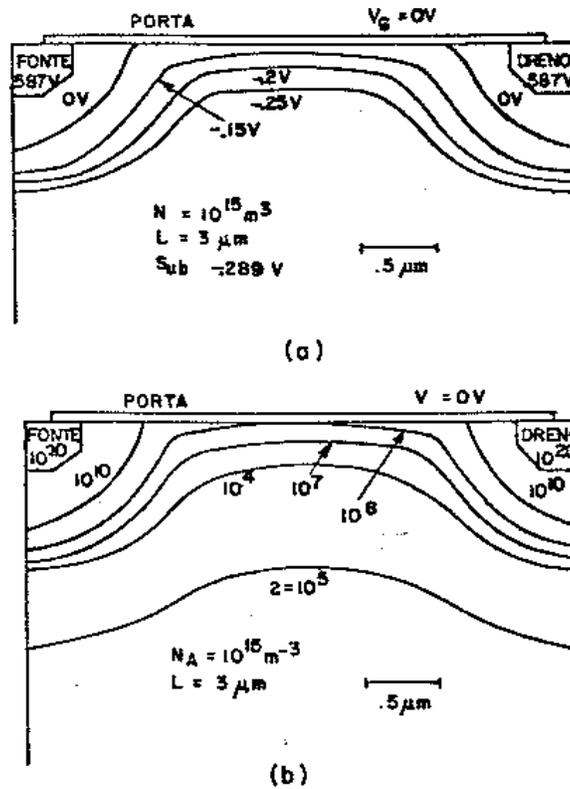


Figura 7 – Perfil do potencial elétrico (A) e da concentração de portadores (B) no equilíbrio, para um dispositivo sem a ocorrência de perfuração MOS.

3.3 – Resistência parasitária de fonte e dreno (R_{DS})

Como demonstrado anteriormente, a impedância de saída dos transistores não varia com o escalamento, no caso da lei por campo elétrico constante. No entanto, com o escalamento seguindo as demais leis, onde a tensão não é escalada na mesma proporção do escalamento das dimensões, a impedância de saída diminui com o escalamento. Desta forma, a razão R_{DS}/R_{on} aumenta, tornando a resistência parasitária mais relevante. Assim, a resistência parasitária tem uma degradação crescente sobre o ganho dos transistores (g_m) [7, 15, 16, 17], sobre a corrente de dreno e atraso das portas [18]. É recomendado que a soma das resistências parasitárias de fonte e dreno não excedam a 10% da resistência intrínseca do canal.

A resistência parasitária de fonte e dreno possui as seguintes componentes, como indicado na figura 8:

- R_{co} – Resistência de contato entre metal e difusão;
- R_d – Resistência da região de difusão;
- R_{sp} – Resistência de espalhamento próximo ao canal;
- R_{ac} – Resistência da região de acumulação, entre região de espalhamento e canal.

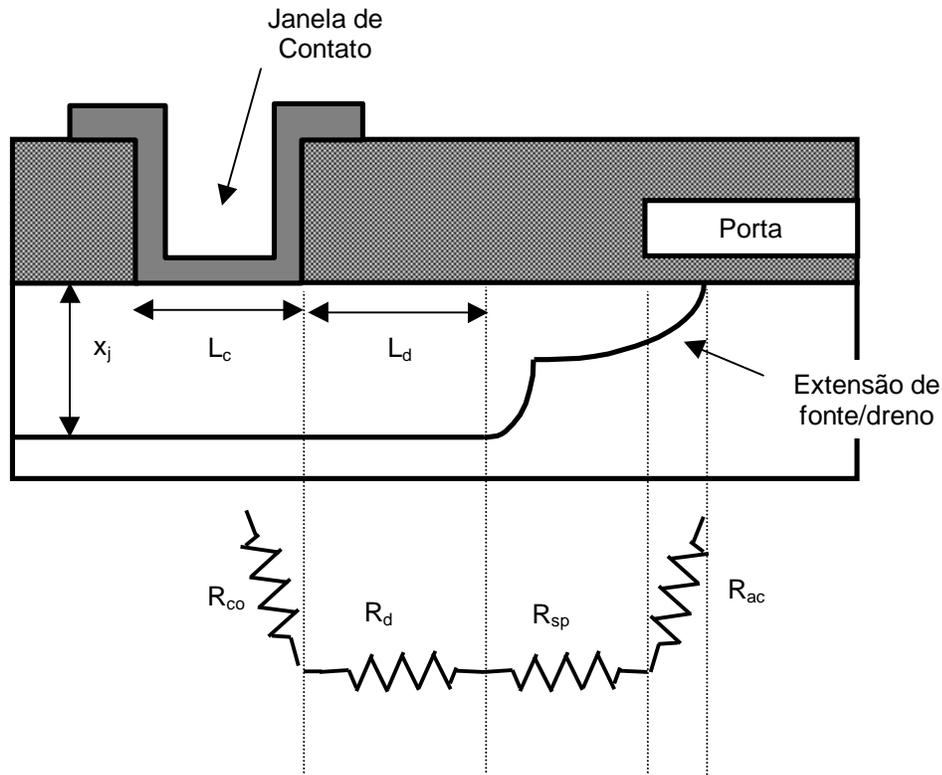


Figura 8 – Representação esquemática das componentes da resistência parasitária entre fonte e dreno.

Analisando-se separadamente cada uma destas componentes:

a) Resistência série da difusão (R_d):

Esta resistência é dada pela seguinte relação:

$$R_d = R_0 \frac{L_d}{W} \quad (14)$$

onde R_0 é a resistência de folha da difusão e L_d é o comprimento da região de difusão.

A resistência de folha era esperada seguir uma dependência de $1/x_j$ com o escalamento. Isto seria correto se a resistividade da difusão fosse constante com o escalamento. Porém, devido a dificuldades práticas para se obter junções rasas, sobretudo tipo p+ (canalização durante a implantação iônica e alto coeficiente de difusão), estas eram obtidas pela redução da dose da implantação iônica e, portanto, com o aumento da resistividade da difusão. Como consequência, a resistência de folha seguia

uma relação do tipo $\left(\frac{1}{x_j}\right)^n$ com $n \approx 6$ para junções rasas do tipo p+ [19].

Atualmente, novas técnicas para a obtenção de junções rasas foram reportadas, baseadas em recozimento térmico rápido (RTP) e implantação em silício pré-amorizados. Segundo estudos recentes, a componente R_d representa a parcela parasitária menos importante, tendo em vista também o uso de siliceto sobre toda região de fonte/dreno mais dopada. Desta forma, a resistência, associada à região mais dopada de fonte e dreno, fica restrita ou embutida na resistência de contato. No entanto permanece a componente de resistência série de difusão associada à região de extensão de fonte e dreno, também chamada de região LDD (Lightly Doped Drain). Esta região normalmente tem nível de dopagem menor e

profundidade de junção também menor, resultando esta sim numa resistência crítica (para $L \cong 100$ nm) [20].

b) Resistência de contato entre metal e difusão (R_{co}):

Esta componente refere-se à resistência entre o metal e uma secção da difusão localizada na borda do contato, e normal à corrente elétrica. Esta resistência depende da resistividade de contato (ρ_c) entre o metal e o semiconductor e da resistência de difusão, além de parâmetros geométricos. A resistência R_{co} é bem representada pela equação [21, 22]:

$$R_{co} = \frac{\sqrt{R_0 \rho_c}}{W} \coth \left(L_c \sqrt{\frac{R_0}{\rho_c}} \right) \quad (15)$$

Para transistores de dimensões pequenas, ($L_c \ll \sqrt{(\rho_c/R_D)}$), obtém-se:

$$R_{co} \approx \frac{\rho_c}{W \cdot L_c}$$

Observa-se um incremento em R_{co} para valores menores de L (dimensão mínima), o que é explicado pelo incremento de R_0 , bem como pela redução da área do contato, com o escalamento das dimensões. Atualmente, o uso de estruturas de siliceto de titânio ($TiSi_2$), cobalto ($CoSi_2$) ou níquel ($NiSi$) na porta, fonte e dreno (estrutura SALICIDE) promove uma redução eficaz na componente R_{co} , porém, mesmo assim, no futuro esta resistência representará uma limitação (para $L \cong 100$ nm) [20].

c) Resistência de espalhamento e de acumulação próximo ao canal ($R_{sp} + R_{ac}$) [17, 20, 23].

Estas duas componentes são dependentes da polarização e do perfil de dopagem, próximo à junção metalúrgica. Quanto mais abrupta o perfil, menor esta região e a sua resistência associada. Nesta região o perfil pode ser aproximado pela equação:

$$N(x) = N_A e^{k_e x} \quad (16)$$

onde $x=0$ na junção.

Observa-se uma redução de R_{sp} e R_{ac} com a redução de L , decorrente do incremento do campo elétrico entre porta e as regiões de fonte e dreno, o que aumenta a carga na região de acumulação dentro da difusão de fonte e dreno. Mesmo assim ela torna-se um valor crítico para tecnologias com $L \cong 70$ nm [20].

3.4 – Capacitância da camada de inversão e da camada de depleção na porta (Si-poli)

A carga no canal do transistor (Q_c) é expressa classicamente por [15]:

$$Q_c = C_{ox} (V_{GS} - V_T) \quad (17)$$

onde $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$ (capacitor de placas paralelas).

Como a camada de inversão (canal) tem uma certa espessura, de 1 a 3 nm tipicamente [20], a estrutura MOS não pode ser tratada como um capacitor de placas paralelas, quando a espessura do isolante de porta for da mesma ordem de grandeza. Adicionalmente, o material de porta de Si-poli, mesmo altamente dopada, apresenta uma camada de depleção de superfície.

Neste caso, a porta MOS apresenta uma capacitância efetiva por unidade de área composta por:

$$\frac{1}{C_{ef}} = \frac{1}{C_{ox}} + \frac{1}{C_c} + \frac{1}{C_{poli}} \quad (18)$$

onde $C_c = \frac{\epsilon_{Si}}{t_c}$, t_c é a espessura média do canal, $C_{poli} = \frac{\epsilon_{Si}}{t_{depl}}$, t_{depl} é a espessura da camada de depleção na porta de Si-poli.

Cálculos mais exatos mostram que o efeito da capacitância da camada de inversão é desprezível para espessuras de óxido de porta até 6 nm [24]. A Fig. 9 mostra a distribuição de portadores no canal e na porta de Si-poli, nas condições de inversão e de acumulação, obtidos por cálculos de mecânica quântica. Estas capacitâncias sérias poderiam ser incorporadas numa capacitância efetiva de óxido, pela adição das espessuras equivalentes (levar em conta diferença das constantes dielétricas dos materiais) das camadas de inversão (~0.3 nm) e de depleção do Si-poli (~0.5 nm) à espessura do óxido [20].

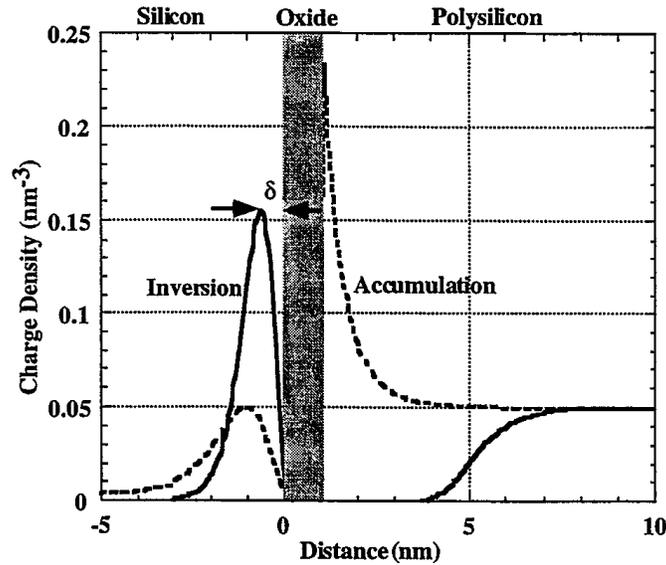


Figura 9 – Distribuição de cargas em capacitor MOS com espessura de óxido de 1 nm e porta de Si-poli, com polarização de 2 V nas condições de inversão (linha cheia) e acumulação (linha tracejada), obtidos por cálculo de mecânica quântica. Os picos das concentrações de portadores ficam distante da interface por efeito de confinamento quântico.

3.5 – Corrente de Tunelamento de Porta

Corrente de tunelamento ocorre quando uma barreira de potencial torna-se muito estreita, como ilustra a expressão:

$$J_{tun} = A \cdot \exp\left(-2\sqrt{\frac{2 \cdot m^* \cdot q \cdot \phi_B}{\hbar^2}} \cdot t_{ox}\right) \quad (19)$$

onde, A é uma constante de proporcionalidade, m^* é a massa efetiva do portador, ϕ_B é a altura da barreira vista pelo portador.

Na verdade, não vem muito ao caso o valor absoluto da corrente de tunelamento pelo dielétrico de porta, mas sim o seu valor relativo à corrente de canal, I_{DS} , devendo ficar limitado a menos de 1% desta. Este critério impõe um limite mínimo para espessuras de SiO_2 da ordem de 1.5 nm, para polarização de 1 V. Uma solução para este problema é substituir o tradicional SiO_2 por outro dielétrico de maior constante dielétrica (teremos uma capacitância equivalente com uma espessura de dielétrico

maior e portanto menor corrente de tunelamento). Esta solução torna-se obrigatória para tecnologias com dimensões mínimas a partir de 100 nm.

3.6 - Redução da Mobilidade

Sabe-se que a mobilidade reduz-se com o aumento do campo elétrico [5, 25, 26]. A seguir, será apresentada a redução da mobilidade com o campo elétrico, separadamente em relação ao campo elétrico transversal e longitudinal.

a) campo elétrico transversal:

Conforme indicado anteriormente, em escalamento realista de dispositivos, a tensão elétrica é reduzida com um fator de escala menor do que o utilizado para as dimensões. Isto faz com que o campo elétrico aumente com o escalamento. Mesmo no caso do escalamento ideal de campo elétrico constante, o campo elétrico transversal aumenta devido ao não escalamento de ϕ_{ms} .

A redução da mobilidade provoca uma diminuição, na mesma proporção, na transcondutância do transistor. Na figura 10 é apresentada a redução da mobilidade em função do campo elétrico transversal, para diversas espessuras de óxido de porta. Na figura 11 observa-se o desvio da transcondutância em relação ao teoricamente previsto, devido à variação do campo elétrico transversal com a redução da espessura do óxido de porta.

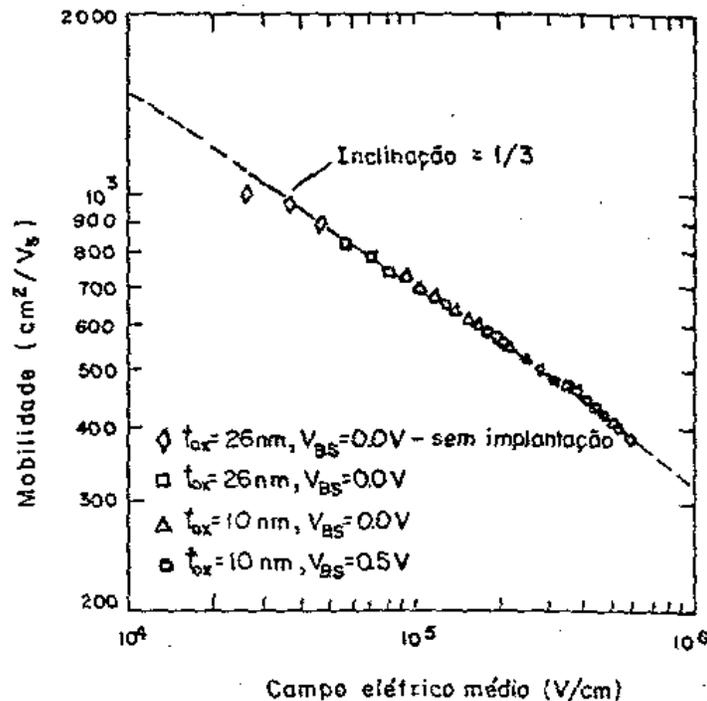


Figura 10 – Redução da mobilidade em função do aumento do campo elétrico transversal, para diversas espessuras de óxido de porta.

b) campo elétrico longitudinal

Similarmente ao campo elétrico transversal, o campo elétrico longitudinal também se eleva com o escalamento dos dispositivos, resultando também em uma redução da mobilidade. Além desta redução da mobilidade, para um campo elétrico maior que um certo campo crítico ($E > E_c$), a velocidade dos portadores satura em uma velocidade máxima de aproximadamente 10^7 cm/s. Este campo crítico vale aproximadamente 2×10^4 V/cm para elétrons e 1×10^{15} V/cm para lacunas, como apresentado na figura 12, na qual tem-se o valor da velocidade dos portadores em função do campo elétrico.

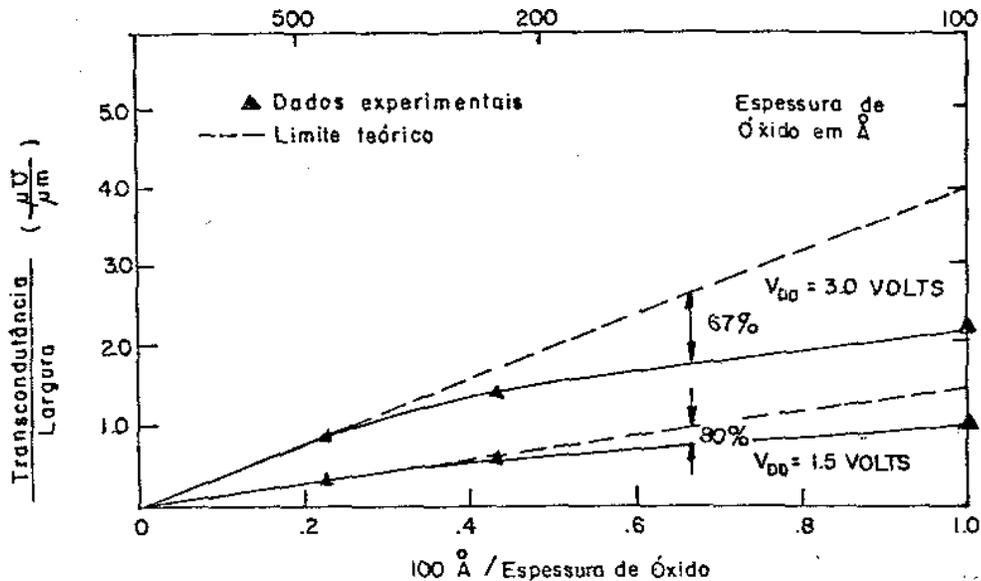


Figura 11 – Redução da transcondutância decorrente da elevação do campo elétrico transversal com a redução da espessura do óxido de porta.

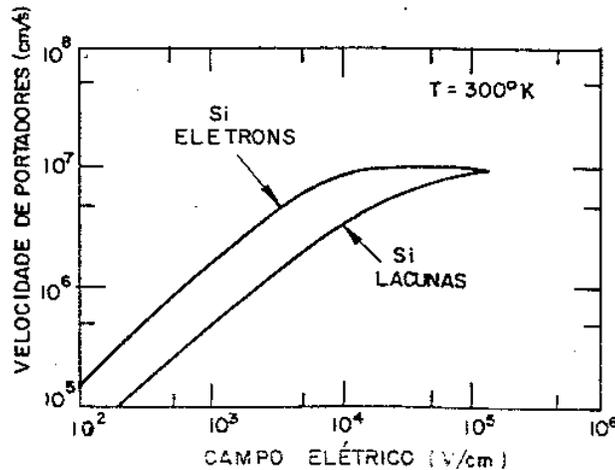


Figura 12 – Velocidade dos portadores em função do campo elétrico.

A redução da mobilidade e a saturação da velocidade dos portadores explicam um crescente desvio da transcondutância com a redução do comprimento de canal, em relação ao limite teórico fornecido pela expressão (20):

$$g_m = \mu \cdot C_{ox} \frac{W}{L} (V_{GS} - V_T) \quad (20)$$

Com comprimento de canal muito curto, o dispositivo apresenta uma saturação no valor da corrente elétrica, expresso pela equação (21), independentemente do valor de L, sendo que todos os portadores caminham com velocidade máxima (v_{max}) [15].

$$I_{DS} = f C_{ox} W v_{max} (V_{GS} - V_T) \quad (21)$$

Nesta situação limite, a transcondutância passa a ser expressa pela seguinte relação, independente de L e da tensão de porta:

$$g_m = f C_{ox} W v_{max} \quad (22)$$

Resultados experimentais mostram que em dispositivos sub-micrométricos, a transcondutância apresenta um comportamento tendendo ao expresso pela equação (22), não mais seguindo a relação quadrática teoricamente esperada em transistores de canal longo.

A mobilidade tem seu valor reduzido por espalhamento com fônons, impurezas (dopantes), cargas de interface e rugosidade da interface do Si e dielétrico. No entanto, uma espalhamento adicional começa a ocorrer quando o dielétrico tornar-se extremamente fino. Para filme de SiO_2 da ordem de 1 nm, as funções de onda de elétrons no metal e no Si começam a sobrepor-se. Nesta situação, impurezas presentes no material de porta representam um espalhamento adicional para os portadores no canal do MOSFET, reduzindo adicionalmente sua mobilidade.

3.7 – Injeção de portadores quentes

Portadores quentes são portadores que possuem alta energia cinética e que, como consequência, podem apresentar os seguintes fenômenos físicos:

- injeção de portadores no óxido de porta, transpondo a barreira de potencial entre o silício e o óxido, como mostrando na figura 13;
- ionização por impacto, criando novos portadores quentes, podendo haver multiplicação por avalanche.

Observa-se pela figura 13 que a barreira para a injeção de lacunas é muito maior que para elétrons. Adicionalmente, o coeficiente de ionização por impacto para elétrons é maior do que para lacunas. Desta forma, os efeitos de portadores quentes são mais intensos em transistores nMOS do que em pMOS.

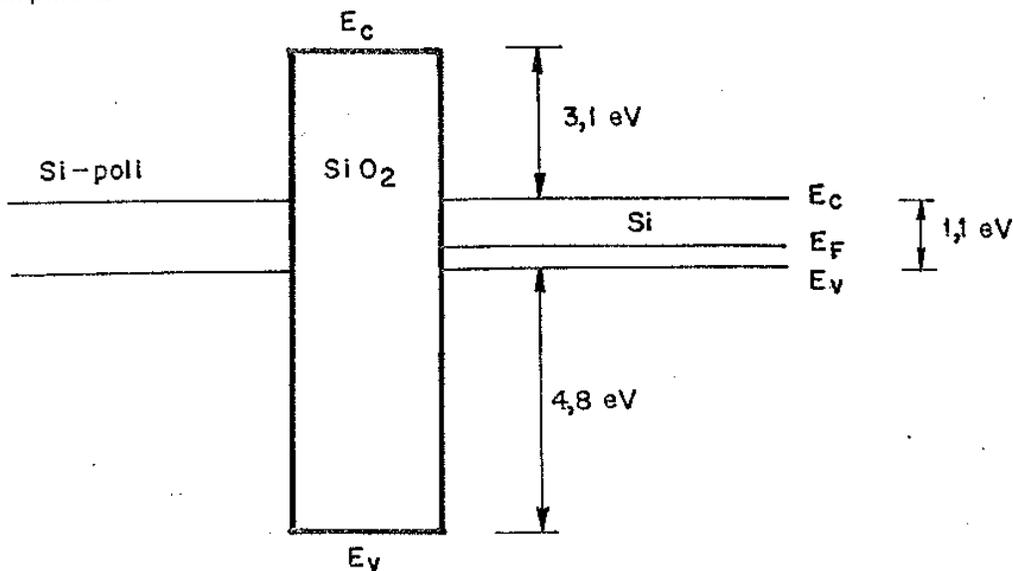


Figura 13 – Diagrama de faixas de energia, indicando as barreiras para elétrons e lacunas na estrutura MOS.

3.7.1 – Efeitos de portadores quentes em transistores MOS

Devido ao aumento do campo elétrico com o escalamento, transistores de menor dimensão são mais sujeitos à ocorrência dos efeitos de portadores quentes, uma vez que os portadores adquirem maior energia cinética.

Com relação à injeção de portadores no óxido de porta, existem 4 modos principais em transistores nMOS, os quais encontram-se apresentados na figura 14.

- elétrons quentes do canal
- elétrons quentes e lacunas quentes produzidos por avalanche;
- elétrons quentes do substrato, induzidos por ionização secundária;
- elétrons térmicos quentes.

O fenômeno de portadores quentes gera uma série de efeitos nos dispositivos MOS, tais como:

- corrente de porta e de substrato

Dependendo do campo elétrico próximo ao dreno, um significativa corrente de substrato pode ser originada, a ponto de provocar quedas ôhmicas apreciáveis no substrato. A corrente de porta, embora não produza queda ôhmica significativa, causa degradações em g_m e V_T com o tempo, reduzindo assim o tempo de vida do componente.

b) degradação da mobilidade ou transcondutância

A degradação da mobilidade e da transcondutância está associada a geração de estados de interface e no óxido, causados pela injeção de portadores quentes.

c) degradação da tensão de limiar

Parte dos portadores quentes no óxido são capturados por estados no óxido, incrementando a densidade de carga aprisionada. Esta carga, por sua vez, produz uma alteração no valor de V_T . Em se tratando da injeção e captura de elétrons, a variação induzida na tensão de limiar será positiva.

d) ruptura do transistor

A queda de potencial no substrato, produzida pela corrente de substrato, pode causar uma polarização direta na junção fonte-substrato, acionando o transistor bipolar parasitário associado à estrutura MOS.

e) efeito tiristor parasitário (*Latch-up*) em CMOS

Uma das origens do disparo da estrutura tiristor parasitário intrínseco à estrutura CMOS é a presença de corrente de substrato produzida por elétrons quentes.

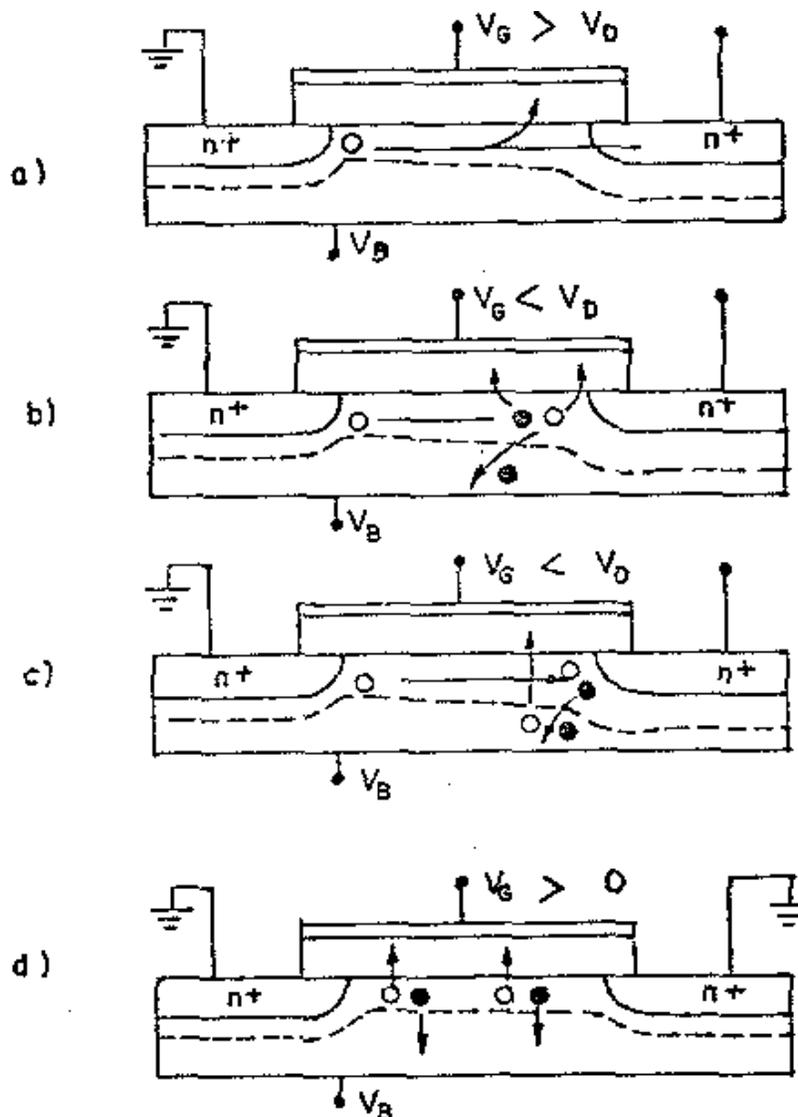


Figura 14 – Representação esquemática dos modos de injeção de portadores quentes em transistores MOS.

3.7.2 – Minimização dos efeitos de portadores quentes

A geração de portadores quentes apresenta uma forte dependência com a intensidade e a distribuição do campo elétrico. Assim, seus efeitos apresentam dependência com os seguintes parâmetros: a) polarizações V_G , V_D , V_B ; b) dimensões L , t_{ox} , X_j ; c) dopagem de substrato; d) forma do perfil do dreno próximo ao canal; e) temperatura.

Motivado pelo escalamento, a maioria destes parâmetros são alterados no sentido inverso ao desejado para a redução de portadores quentes. Resta, no entanto, a opção de alterar o perfil do dreno próximo ao canal. Estes perfis fazem com que uma maior parte da tensão aplicada tenha sua queda dentro do dreno e que o campo elétrico dentro do canal seja menor [27]. Entre estas estruturas, a estrutura LDD (*lightly doped drain*) tem sido a mais empregada [27, 28]. Como desvantagem desta solução temos um incremento na resistência parasitária R_{DS} .

3.8 – Rupturas do transistor

Limitações mais severas referem-se a condições em que o transistor apresente características totalmente diferentes, devido a algum processo de ruptura. Apresentamos a seguir algumas possibilidades de rupturas.

a) Ruptura do óxido de porta

Um óxido de alta qualidade apresenta uma ruptura intrínseca destrutiva para campos elétricos intensos, maiores que 10^7 V/cm [29]. Por motivos de confiabilidade operava-se os dispositivos com campo elétrico no óxido até na faixa de 1 a 1.5×10^6 V/cm [30]. Atualmente no entanto, campos típicos são da ordem 5×10^6 V/cm, devendo no futuro aumentar mais ainda [20]. Modelos sobre mecanismos de ruptura têm sido apresentados, incluindo ruptura por corrente de porta, [31]. Após certa carga ($I \times T$) ter passado pelo óxido, a fração de carga capturada nas armadilhas no óxido torna-se elevada, elevando o campo elétrico localmente. A partir desta evolução tem início a ruptura [32]. A ruptura do óxido é muito dependente da presença de contaminantes e da estrutura do mesmo. Assim, realiza-se hoje em dia pesquisa em processos para obtenção de isolantes mais “duros”, incluindo novos materiais como oxinitretos [33-36].

b) Ruptura bipolar parasitária

Como reporta anteriormente, a corrente de substrato pode causar a polarização direta da junção fonte-substrato, colocando em operação o transistor bipolar parasitário no transistor MOS. Quando isto ocorre, a corrente elétrica I_{DS} aumenta drasticamente, comumente observando-se nas curvas $I_{DS} \times V_{DS}$ do transistor o efeito “*snap-back*”, ou seja, uma redução de V_{DS} após a ruptura bipolar.

c) Ruptura por avalanche de dreno e/ou canal

Para um campo elétrico acima do valor crítico, entre dreno/substrato e/ou canal/substrato, pode haver a ruptura devido à multiplicação de portadores por avalanche [37, 38]. Na maioria dos transistores de pequenas dimensões, perfuração MOS e ruptura bipolar parasitária, ocorrem com tensões menores, como ilustra a Figura 15. Esta figura expressa as limitações em tensão V_{DS} quanto às diversas rupturas, válido para uma dada estrutura de transistores [38].

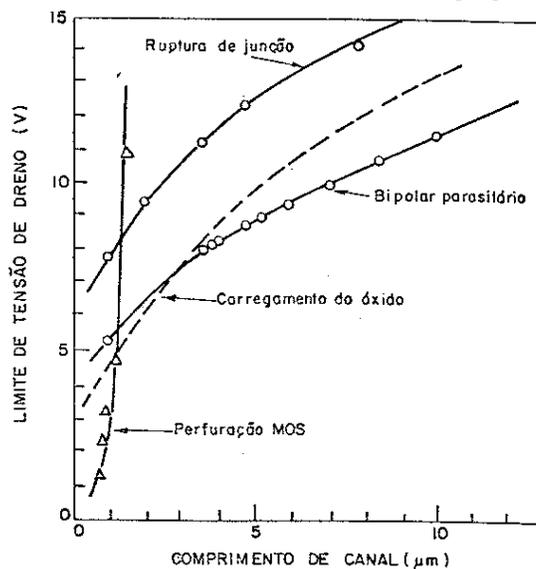


Figura 15 – Limites admissíveis para V_{DS} em função dos vários mecanismos de ruptura possíveis.

3.9 – Efeitos das limitações sobre o escalamento dos transistores e elaboração de “guias de estrada”.

As limitações discutidas nos itens anteriores devem ser muito bem estudadas para permitir o correto escalamento dos dispositivos. Isto porque elas:

- afetam o desempenho elétrico dos dispositivos
- determinam as condições limites de operação e
- determinam condições de contorno para o projeto da estrutura física dos transistores e do processo de fabricação.

A tensão de alimentação foi mantida constante na fase inicial do escalamento, como valor padrão de 5 V. Após 1990 no entanto, devido às diversas limitações apresentadas, não foi mais possível manter este valor. Após esta data observamos uma redução gradual do seu valor, como ilustra a Fig. 16. O valor de V_T de certa forma acompanha a mesma tendência, como também ilustrada na mesma figura. Isto se faz necessário para manter uma boa margem de tensão de comando (drive) para bom desempenho de velocidade de chaveamento. A redução do valor de V_T traz no entanto, uma grave dificuldade associada à alta corrente de corte I_0 , como explicado no item 2.1 acima. Como consequência, existe muita incerteza quanto ao valor mais apropriado de V_T adotar, ilustrado na figura. Por exemplo, ao se chegar a tecnologias com V_{DD} de 0.5 V, não sobra muito espaço para a escolha de V_T que atenda tanto a condição de baixo valor de corte, I_0 , ao mesmo tempo de oferecer alta corrente para rápida comutação. Este é um dos problemas mais sérios para as futuras tecnologias.

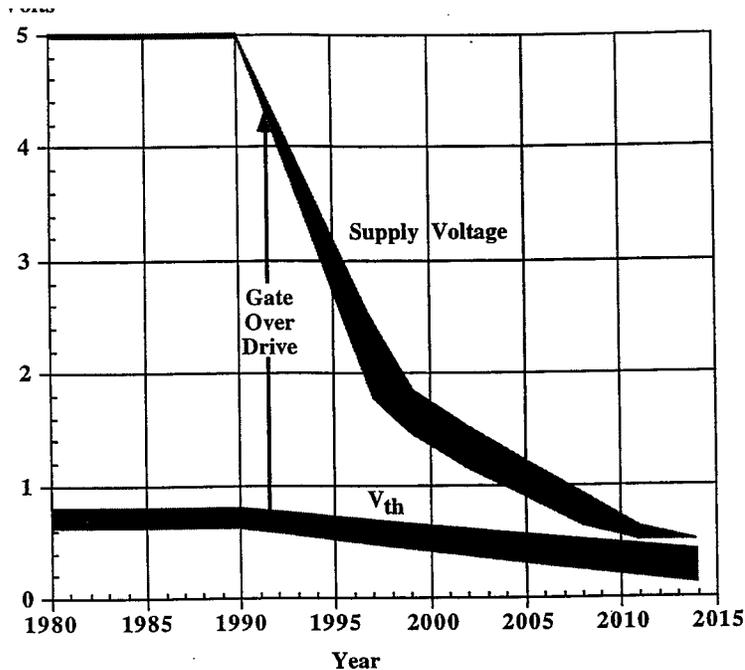


Figura 16 – Evolução dos valores de V_{DD} e V_T

O estudo das limitações dos dispositivos constitui também um dos ingredientes para a elaboração de “guias de estrada” de evolução tecnológica de semicondutores. Associações de empresas de semicondutores, juntamente com instituições públicas e universitárias, formaram grupos de trabalho para a elaboração de parâmetros das próximas gerações tecnológicas. Estes trabalhos fazem-se necessários com intuito de definir padrões e estratégias comuns para a definição e a solução dos futuros requisitos, bem como das ações necessárias. A tabela 6 apresenta parâmetros selecionados, definidos para as várias gerações tecnológicas, dos relatórios de 1997 e 1999 destes grupos de trabalho [39, 40]. Em negrito são apresentados os valores dos parâmetros para os quais ainda não existe solução tecnológica para a sua obtenção, representando temas e desafios de pesquisa atuais.

Tabela 6 – Dados selecionados dos relatórios NTRS [39] e ITRS [40] dos parâmetros tecnológicos atuais e futuros previstos [20].

Ano	1997	1999	2002	2005	2008	2011	2014
Dimensão mínima (nm)	250	180	130	100	70	50	35
DRAM (início de vendas)	256M	1G	(3G)	8G	(24G)	64G	(192G)
Área chip DRAM (mm ²)	280	400	460	530	630	710	860
Espessura equivalente de óxido (nm)	3-5	1.9-2.5	1.5-1.9	1.0-1.5	0.8-1.2	0.6-0.8	0.5-0.6
Res. máx. de material de porta ($\mu\Omega\cdot\text{cm}$)		60	43	33	23	16	11
Res máx. de contato siliceto/si ($\Omega\cdot\text{cm}^2$)		30×10^{-8}	17×10^{-8}	10×10^{-8}	5×10^{-8}	2.5×10^{-8}	1.5×10^{-8}
Resistência de folha da extensão S/D ($\Omega/$)		350-800	250-700	200-625	150-525	120-525	100-400
X _J da extensão S/D (nm)	50-100	42-70	25-43	20-33	16-26	11-19	8-13
Perfil da extensão S/D (nm/dec.)		14	8.5	6.5	4.5	3.2	2.2
V _{DD}	1.8-2.5	1.5-1.8	1.2-1.5	0.9-1.2	0.6-0.9	0.5-0.6	0.5

Uma questão de grande importância está relacionada com os problemas associados aos sistemas fotolitográficos e aos processos e dificuldades de realização da etapa para a definição de dimensões altamente submicrométricas. Grande parte do custo da fabricação de CI's está relacionada a esta etapa.

Embora, em termos teóricos, seja possível obter-se definições de padrões até níveis atômicos, tais sistemas seriam extremamente lentos e não atingiriam os requisitos de velocidade necessários para produção em larga escala.

Os sistemas mais avançados de litografia atuais utilizam impressão por projeção ótica, operando no limite de difração de Rayleigh. A imagem de um padrão principal (normalmente reduzida de 4 a 5 vezes) é projetada sobre a superfície da lâmina. Para tal, estes sistemas utilizam um complexo sistema de lentes. A resolução do sistema fotolitográfico está diretamente relacionada com o comprimento de onda da luz utilizada para sensibilizar o fotorresiste. A figura 17 apresenta os valores de comprimento de onda e de dimensão mínima utilizados em função do ano.

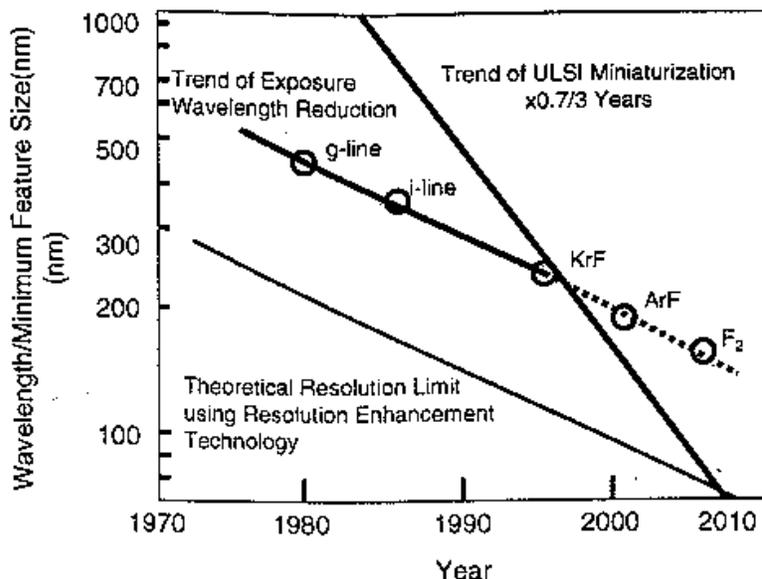


Figura 17 – Valores do comprimento de onda/mínima dimensão em função do ano.

Lasers a base de fluoreto de criptônio (KrF) e fluoreto de argônio (ArF) são utilizados para as gerações tecnológicas de $0,25 \mu\text{m}$ e $0,18 \mu\text{m}$, respectivamente. Entretanto, os melhores resultados são

obtidos com laser KrF. Para as gerações tecnológicas futuras acredita-se que o uso de lasers a base de flúor (F_2) sejam necessários, como indicado na figura 17. O gráfico apresentado mostra que o processo óptico apresenta um limite para sua utilização para no máximo a dimensão um pouco abaixo de 100 nm [41]. Após este limite, outras técnicas tornam-se imprescindíveis, tais como projeção de elétrons, multi-feixes de elétrons, raio X, feixe de íons, matriz de pontas de microscopia de força atômica, etc.

4 – Limites de escalamento

Nos itens anteriores apresentamos os procedimentos utilizados para o escalamento e as limitações que surgem concomitantemente. Analisamos agora o limite do escalamento, ou seja, até onde podemos reduzir as dimensões, tensões e energia da informação. A Figura 18 mostra a evolução da energia da informação em operações lógicas, ao longo dos anos, com uma contínua redução da mesma. Até onde esta evolução poderá continuar? Iniciamos com a apresentação de limites teóricos e a seguir apresentamos alguns dados experimentais reportados a título de exemplos.

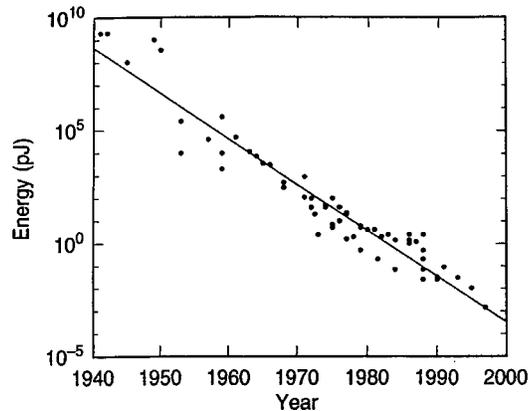


Figura 18 – Evolução da energia usada para efetuar operações lógicas, ao longo dos anos.

4.1 – Limites teóricos

De acordo com Meindl [42], os limites de escalamento de dispositivos podem ser agrupados em 5 classes:

1. Limites fundamentais
2. Limites do material
3. Limites do dispositivo
4. Limites do circuito
5. Limites do sistema

Com o intuito de analisarmos o escalamento até o nível de dispositivo, apresentaremos a seguir apenas considerações quanto aos primeiros limites [42].

1) Entre os limites fundamentais temos:

a) Devido a flutuações térmicas no material, qualquer informação (energia armazenada) com energia próxima à da flutuação térmica, terá alta probabilidade (estatística de Boltzman) de ser perdida [43]. Assim, é necessário que a informação tenha no mínimo [42]:

$$\Delta\epsilon > 4 kT \quad (23a)$$

ou ainda, para manter a probabilidade de erro menor que 10^{-19} , a energia deve ser maior ainda, ou seja [44]:

$$\Delta\epsilon > 165 kT \quad (23b)$$

b) O princípio da incerteza da mecânica quântica diz que

$$\Delta p \cdot \Delta r > h \quad (24)$$

ou que

$$\Delta \epsilon \cdot \Delta t > h \quad (25)$$

A partir desta relação, obtém-se que a mínima energia armazenada, a ser detectada num intervalo de tempo Δt deve ser maior que $h / \Delta t$ [45]

$$\Delta \epsilon > h / \Delta t \quad (26)$$

2) Os limites dos materiais, por sua vez, referem-se à propriedade dos materiais. Entre estas citamos:

a) Campo elétrico máximo que o material suporta sem ruptura por avalanche, E_c

b) Velocidade máxima ou velocidade de saturação dos portadores, v_{max}

c) Massa efetiva dos portadores, m_e

Assim, o atraso mínimo relacionado a estes limites pode ser obtido por:

$$\tau_{min} = \frac{L_{min}}{v_{max}} + \frac{\Delta V / E_c}{v_{max}} \quad (27)$$

Considerando os valores para silício ($E_c = 3 \times 10^5$ V/cm³, $v_{max} = 1 \times 10^7$ cm/s) e $\Delta V = 4 KT / q$ (limite fundamental) obtém-se $\tau_{min} = 3 \times 10^{-14}$ s. Da massa efetiva pode-se obter ainda a espessura mínima de uma barreira de potencial para a qual a corrente de tunelamento seja desprezível. Quanto menor a massa efetiva, maior esta espessura mínima. Assim, com silício pode-se fabricar dispositivos com dimensões físicas menores que no GaAs, o qual apresenta menor massa efetiva de elétrons [43].

3) Em dispositivos MOSFET, o comprimento mínimo do canal é determinado pela ocorrência da perfuração MOS (*punchthrough*), degradação das características devido à efeito de canal curto, bem como pelas demais limitações apresentadas no item 3. Estima-se que um limite prático deve estar em torno de 25 nm [46] ou mesmo de 10 nm [47]. O uso de tecnologia CMOS/SOI, de baixa temperatura (N₂ líquido) e/ou novas estruturas de dispositivos MOS irão contribuir para a extensão do tempo de vida das tecnologias CMOS de Si. Entre novas estruturas vem sendo proposto o transistor MOS de dupla porta (DG-CMOS/SOI), ilustrada na Fig. 19. A Fig. 20 ilustra como a introdução de inovação tecnológica é necessária quando uma dada evolução tecnológica chega à saturação da sua capacidade. O que virá após o limite de escalamento na tecnologia CMOS? Novos conceitos de dispositivos e circuitos terão lugar. Entre estas já são sugeridos: a) dispositivos de bloqueio Coulombiano, entre outros dispositivos de um único elétron; b) dispositivos quânticos, onde se controla o estado do elétron de um átomo (hidrogênio, por exemplo); c) estruturas de nano-tubos de carbono é outra idéia proposta. São tubos de 1.4 nm de diâmetro e de 10 μ m de comprimento que constituem canais de corrente e que permitem realizar circuitos tipo moleculares.

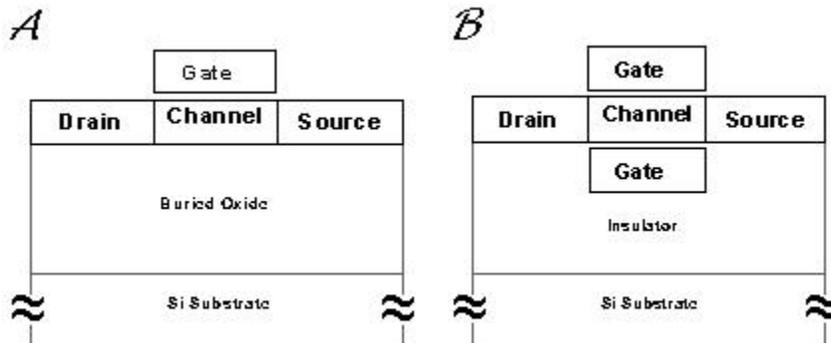


Figura 19 – Desenho em corte de estrutura de a) CMOS/SOI e b) DG-CMOS/SOI [48].

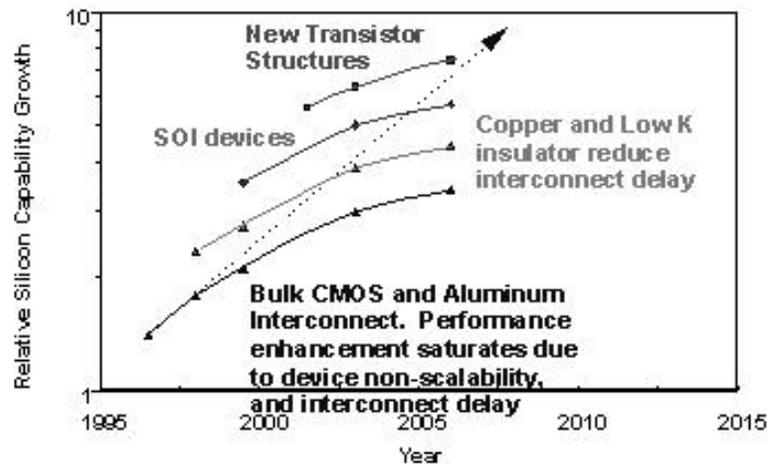


Figura 20 – Tendências de desempenho de circuitos CMOS, com desafios e soluções tecnológicas [28]

Referências:

- [1] G. E. Moore, "Progress in Digital Integrated Electronics", *IEDM Tech. Digest*, p. 11-13, 1975.
- [2] J. D. Meidl, "Ultra-Large Scale Integration", *IEEE Trans. Electron Devices*, v. 31, n. 11, p. 1555-1561, 1984.
- [3] E. J. Rymaszeroski, "Dense, Denser, Denser", *J. Electron Mater.*, v. 18, n. 2, p. 217-220, 1989.
- [4] R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. H. Rideout, E. Bassous and A. R. LeBlanc, "Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions", *IEEE Journal Solid-St. Circuits*, v. 9, p. 256-268, 1974.
- [5] S. C. Sun and J. D. Plummer, "Electron Mobility in Inversion and Accumulation Layers Thermally Oxidized Silicon Surfaces", *IEEE Trans. Electron Devices*, v. 27, p. 1497-1508, 1980.
- [6] P. K. Chatteyee, W. R. Hunter, T. C. Holloway and Y. T. Lin, "The Impact of Scaling Laws on the Choice of n-Channel or p-Channel for MOS VLSI", *IEEE Electron Device Letters*, vol. 1, p. 220-223, 1980.
- [7] J. R. Brews, W. Fichtner, E. A. Nicollian and S. M. Sze, "Generalized Guide for MOSFET Miniaturization", *IEEE Electron Device Letters*, vol. 1, p. 1-4, 1980.
- [8] G. Baccarani, M. R. Wordeman and R. H. Dennard, "Generalized Scaling Theory and Its Application to a 1/4 Micrometer MOSFET Design", *IEEE Trans. Electron Devices*, v. 31, p. 452-462, 1984.
- [9] SUPREM-4, Bidimensional Process Simulator, Stanford University, 1993.
- [10] Avant! MEDICI, Bidimensional Device Simulator, 2001.
- [11] PISCES-II, Bidimensional Device Simulator, Stanford University, 1993.
- [12] Avant! DAVINCI, Tridimensional Device Simulator, 2001.
- [13] L. D. Yau, "A Simple Theory to Predict the Threshold Voltage of Short-Channel IGFET's" *Solid-State Electronics*, v. 17, p. 1059-1063, 1974.
- [14] B. S. Nataraj and R. Kumar, "Threshold-Voltage Variations in VLSI MOSFET's Due to Short Channel Lengths", *IEEE Journal Solid-St. Circuits*, v. 22, p. 905-908, 1987.
- [15] Y. El-Mansy, "MOS Device and Technology Constraints in VLSI", *IEEE Journal Solid-St. Circuits*, v. 17, p. 197-203, 1982.
- [16] D. M. Brown, M. Ghezze and J. M. Pimbley, "Trends in Advanced Process Technology-Submicrometer CMOS Design and Process Requirements", *Proc. of the IEEE*, vol. 74, p. 1678-1702, 1986.
- [17] K. K. Ng and W. T. Lynch, "The Impact of Intrinsic Series Resistance on MOSFET Scaling", *IEEE Trans. Electron Devices*, v. 34, p. 503-511, 1987.
- [18] J. S. Kim and H. C. Lin, "Modeling Effective Source Resistance of a Short-Channel MOSFET", *Proc. of the IEEE 1984 Custom Integrated Circuits Conference*, p. 335-338, 1984.
- [19] H. Shichijo, "A Re-Examination of Practical Performance Limits of Scaled n-Channel and p-Channel MOS Devices for VLSI", *Solid-State Electronics*, v. 26, p. 969-986, 1983.
- [20] J. D. Plummer, P. B. Griffin, "Material and Process Limits in Silicon VLSI", *Proceedings of The IEEE*, vol. 89, no. 3, pp. 240-258 (2001).

- [21] H. Murrmann and D. Widmann, "Current Crowding an Metal Contacts to Planar Devices", *IEEE Trans. Electron Devices*, v. 16, p. 1022-1024, 1969.
- [22] H. Berger, "Models for Contracts to Planar Devices", *Sol.-St. Electr.*, v. 15, p. 145-158, 1972.
- [23] K. K. Ng and W. T Lynch, "Analysis of the Gate-Voltage-Dependent Series Resistance of MOSFET's", *IEEE Trans. Electron Devices*, v. 33, p. 965-972, 1986.
- [24] S.-Y. Oh, S. -G. Choi, C. G. Sodini and J. L. Moll, "Analysis of the Channel Inversion Layer Capacitance in the Very Thin-Gate IGFET", *IEEE Electron Device Letters*, v. 4, p. 236-239, 1983.
- [25] A. G. Salmis and J. T. Clemens, "Characterization of the Electron Mobility in the Inverted <100> Surface", *IEDM Tech. Digest*, p. 18-21, 1979.
- [26] F. F. Fany and A. B. Fowler, "Transport Properties of Electron in Inverted Silicon Surfaces", *Phys. Rev.*, v. 169, p. 619-631, 1968.
- [27] E. Takeda, H. Kume, T. Toyabe and S. Asai, "Submicrometer MOSFET Structure for Minimizing Hot-Carrier Generation", *IEEE Journal Solid-St. Circuits*, v. 17, p. 241-248, 1982.
- [28] J. J. Sanchez, K. K. Hsueh and T. A. DeMassa, "Drain-Engineered Hot-Electron-Resistant Device Structures: A Review", *IEEE Trans. Electron Devices*, v. 36, p. 1125-1132, 1989.
- [29] J. Schadel, "Device Failure Mechanisms in Integrated Circuits", *Proc. of ESSDERC*, p. 13-16, 1983.
- [30] J. L. Moll and E. Y. Sun, "Physical Effects in Small Geometry MOS Transistors", *Jap. J. Appl. Phys.*, vol. 19, p. 77-83, 1980.
- [31] S. Holland, I. C. Chan, T. P. Ma and C. Hu, "On Physical Models for Gate Oxide Breakdown", *IEEE Electron Device Letters*, vol. 5, p. 302-305, 1984.
- [32] I. -C. Chan and S. E. Holland, "Electrical Breakdown in Thin Gate and Tunneling Oxides", *IEEE Trans. Electron Devices*, v. 32, p. 413-422, 1985.
- [33] F. L. Terry, R. L. Aucoin, M. L. Naiman, P. W. Wyatt and S. D. Senturia, "Radiation Effects in Nitrided Oxides", *IEEE Electron Device Letters*, v. 4, p. 191-193, 1983.
- [34] M. A. Schmidt, F. L. Terry Jr., B. P. Mathur and S. D. Senturio, "Inversion Layer Mobility of MOSFET's with Nitrided Oxide Gate Dielectrics", *IEEE Trans. Electron Devices*, v. 35, p. 1627-1632, 1988.
- [35] J. A. Diniz, P. J. Tatsch, M. A. A. Pudenzi, "Oxynitride Films Formed by Low Energy NO⁺ Implanted into Silicon", *Appl. Phys. Lett.*, v. 69, n. 15, p. 2214-2215, 1996.
- [36] J. A. Diniz, J. Godoy Fo., P. J. Tatsch, J. W. Swart, "Radiation Hardening of Oxynitrides Formed by Low Nitrogen Implantation into Silicon Prior do Oxidation", *The 199th Meeting of The Electroch. Soc.*, Abstract n. 262, v.2001-1, Washington DC, USA, 25 a 29 de março de 2001.
- [37] J. Chen, T. Y. Chan, P. K. Ko and C. Hu, "Subbreakdown Drain Leakage Current in MOSFET", *IEEE Electron Device Letters*, v. 8, p. 515-517, 1987.
- [38] S. M. Sze, *Physics of Semiconductor Devices*, J. Wiley & Sons, p. 485, 1981.
- [39] Semiconductor Industry Association, "Nacional Technology Roadmap for Semiconductors", San Jose, CA, SIA, 1997.
- [40] Semiconductor Industry Association, "Internacional Technology Roadmap for Semiconductors", San Jose, CA, SIA, 1999.
- [41] L. R. Harriot, "Limits of Lithography", *Proc. of the IEEE*, v. 89, n. 3, p. 366-374, 2001.
- [42] J. D. Meindl, "Ultra Large Scale Integration", *IEEE Trans Electron Devices*, v. 31, p. 1555-1561, 1984.
- [43] C. Mead and L. Conway, *Introduction to VLSI Systems*, Addison-Wesley Public Co., 1980.
- [44] R. W. Keyes, "Fundamental Limit of Silicon Technology", *Proc. of The IEEE*, v. 89, n. 3, p. 227-239, 2001.
- [45] R. W. Keyes, "Physical Limits in Digital Electronics", *Proc. of The IEEE*, v. 63, p. 740-767, 1975.
- [46] H. Iwai, "CMOS Technology – Year 2010 and Beyond", *IEEE Journal Solid-St. Circuits*, v. 34, n. 3, p. 357-366, 1999.
- [47] D. J. Frank et al, "Device Scaling Limits of Si MOSFETs and Their Application Dependencies", *Proc. of The IEEE*, v. 89, n. 3, p. 259-288, 2001.
- [48] R. W. Guernsey, F. L. Gandour, "A Competitividade da Indústria Eletrônica", ABINEE TEC 2001, São Paulo, 21-25 de Maio 2001.