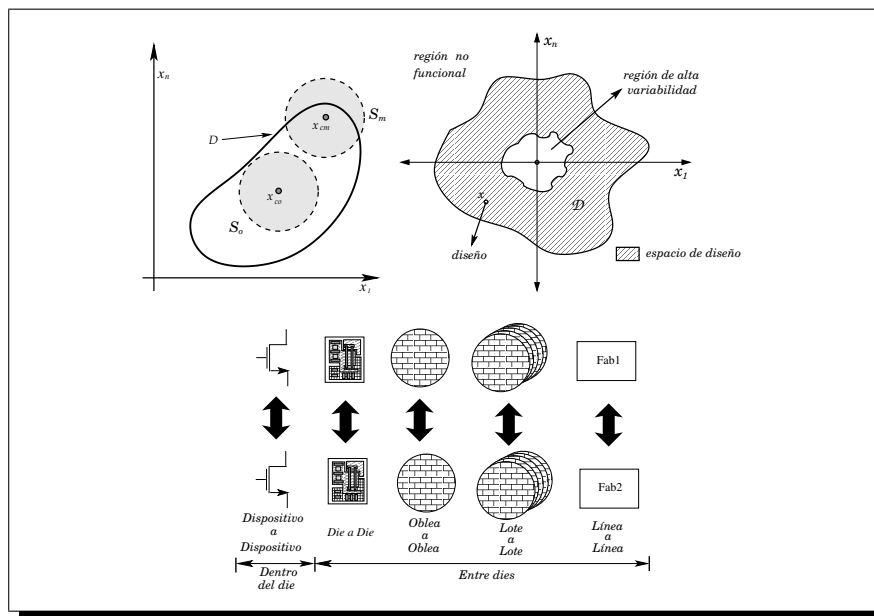


APLICACIÓN DE LA PROGRAMACIÓN GEOMÉTRICA EN LA OPTIMIZACIÓN DEL DISEÑO DE UNA FUENTE DE VOLTAJE DE REFERENCIA



Jorge Johanny Sáenz Noval

ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y
TELECOMUNICACIONES

UNIVERSIDAD INDUSTRIAL SANTANDER

Bucaramanga – 2009



UNIVERSIDAD INDUSTRIAL DE SANTANDER
Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones



APLICACIÓN DE LA PROGRAMACIÓN GEOMÉTRICA EN LA OPTIMIZACIÓN DEL DISEÑO DE UNA FUENTE DE VOLTAJE DE REFERENCIA

Jorge Johanny Sáenz Noval

Trabajo de grado para optar por el título de Ingeniero Electrónico

Director

MSc. Élkim Felipe Roa Fuentes

**ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y
TELECOMUNICACIONES**

UNIVERSIDAD INDUSTRIAL SANTANDER

Bucaramanga–Mayo de 2009

*“ A mi padre, mis hermanas
y mi mamá por
su vital apoyo en este largo
y ardúo camino. ”*

AGRADECIMIENTOS

Es amplio el número de personas las cuales directa o indirectamente apoyaron en la consecución de cada uno de mis logros durante mi carrera y estuvieron siempre acompañándome en mis éxitos y fracasos. A todos ellos un agradecimiento muy especial.

Agradezco ampliamente al profesor Élkim por su aportación a mi conocimiento como persona y como profesional, y a su apoyo en aquellos momentos donde su ayuda y comprensión fue vital. Siempre considero que su trabajo hecho con el grupo CIDIC representa un grandioso aporte al desarrollo de nuestro país.

A mi familia, por ser el eje central e incondicional sobre el cual se apoyan cada una de mis motivaciones y deseos.

A los integrantes del grupo CIDIC por su bondad y absoluta disposición en cada uno de los procesos de edición del libro y en su aporte de ideas nuevas.

A la Universidad Industrial de Santander, por ser el lugar propicio donde se construyeron gran parte de las experiencias y conocimientos adquiridos.

Jorge Johanny Sáenz Noval

RESUMEN

TÍTULO:

APLICACIÓN DE LA PROGRAMACIÓN GEOMÉTRICA EN LA OPTIMIZACIÓN DEL DISEÑO DE UNA FUENTE DE VOLTAJE DE REFERENCIA ¹

AUTOR:

JORGE JOHANNY SÁENZ NOVAL ²

PALABRAS CLAVE: Programación Geométrica, CAD, Fuente de Referencia, Optimización, DFM, Metodologías de Diseño.

DESCRIPCIÓN: En este proyecto se propone una metodología de diseño para manufactura usando programación geométrica para automatizar el diseño. Luego, esta metodología es validada en la optimización del diseño de una fuente de voltaje de referencia.

Durante los últimos años, el aumento en la densidad de transistores pronosticado por la Ley de Moore ha convertido el diseño de circuitos integrados en una fase altamente sensible a las variaciones del proceso de fabricación. Las diferencias entre lo proyectado y lo fabricado generan pérdidas de rendimiento que en un gran porcentaje son atribuidas al proceso que involucra al diseñador. Al ser alta la responsabilidad del diseñador en este problema, el diseño analógico debe concentrarse en proponer nuevas metodologías de diseño que permitan optimizar conjuntamente desempeño y rendimiento.

Las fuentes de referencia corresponden al tipo de circuitos cuyo funcionamiento y especificaciones de diseño se ven ampliamente alteradas por el escalamiento tecnológico. Pretender obtener un diseño con altas especificaciones de desempeño implica trabajar directamente sobre el rendimiento del circuito, pues un óptimo de desempeño obtenido por el esquema tradicional de optimización geométrica, no garantiza la consecución de diseños fabricables.

El presente proyecto evalúa los alcances y beneficios que tiene la implementación de la metodología para manufactura por medio del rediseño optimizado de una fuente de voltaje de referencia. Además, busca establecer los factores limitantes en el diseño para manufactura, relacionados comúnmente con la caracterización estadística y la frontera tecnológica del proceso. De esta forma, se concentra en proponer una metodología que condense la bondad en la consecución de óptimos globales de la programación geométrica, con la optimización del rendimiento funcional y paramétrico de la fuente de referencia.

¹Proyecto de Grado

²Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director MSc. Élkim Felipe Roa Fuentes.

SUMMARY

TITLE:

APPLICATION OF GEOMETRIC PROGRAMMING IN THE OPTIMIZATION OF VOLTAGE REFERENCE DESIGN. ³

AUTHOR: JORGE JOHANNY SÁENZ NOVAL ⁴

KEY WORDS: Geometric Programming, CAD, Voltage Reference, Optimization, DFM, Design Methodologies.

DESCRIPTION: This project proposes a design methodology for manufacturing using geometric programming to automate the design. This methodology is validated to optimize the design of a reference voltage.

During recent years, the increase in transistor density predicted by Moore's Law has become the design of integrated circuits highly sensitive to variations due to manufacturing process. The differences between projected and manufactured devices produce yield loss, in which a large proportion are attributed to the process involving the designer. Because of the designer has a high responsibility on this problem, the analog design must focus on proposing new design methodologies for optimizing jointly performance and yield.

The voltage reference are those type of circuits whose operation and design specifications are largely altered by the technology scaling. Attempting to obtain a design with high performance specifications involves working directly on the yield of the circuit, because an optimal performance obtained by the traditional framework of geometric optimization, does not obtain a marketable and competitive design.

This project evaluates the scope and benefits of implementing the methodology for manufacturing by means of an optimized redesign of the voltage reference. It also seeks to establish the limiting factors in design for manufacturing, commonly associated with the statistical characterization and the technological limit of the process. In this way this work focus in proposing a methodology to combine the goodness of obtaining overall optimum in programming geometric with functional and parametric yield of the voltage reference.

³Degree Project

⁴Physics Mechanical Engineering Faculty, Electric, Electronic and Telecommunications School. Director MSc. Élkim Felipe Roa Fuentes.

Contenido

1. Introducción	1
1.1. Motivación	2
1.2. Diseño para Manufactura <i>DFM</i>	3
1.3. Compromisos y retos en la implementación <i>DFM</i>	6
1.4. Algunas fuentes de referencia optimizadas	9
1.5. Organización del Documento	10
2. Modelado estadístico del proceso	13
2.1. Modelos de <i>mismatch</i>	15
2.1.1. Modelo de Pelgrom	15
2.1.2. Modelo de Drennan	17
2.1.3. Modelo Galup-Montoro	18
2.2. Discusión	20
2.2.1. Significado físico	20
2.2.2. Continuidad	22
2.2.3. Medibilidad	23
2.2.4. Precisión	24
2.2.5. Utilidad	26
2.2.6. Escalabilidad	27
2.3. Conclusión y selección.	29
3. Metodología de diseño <i>DFM</i>	31
3.1. Aproximación del espacio de diseño.	33
3.2. Mínima área que garantiza aceptabilidad.	36
3.3. Programación Geométrica	39
3.3.1. Análisis de Convexidad y modelo sub-umbral	40
3.3.2. Leyes de Kirchhoff en PG	42
3.3.3. La centralización del diseño como un programa geométrico.	43

4. Análisis de rendimiento y desempeño	47
4.1. Fuente de referencia, [1]	47
4.2. Optimización vía programación geométrica.	50
4.3. Bloque <i>PTAT</i>	51
4.4. Bloque <i>CTAT</i>	53
4.5. Bloque <i>PWL</i>	56
4.6. Rama final	58
4.7. Discusión	58
4.8. Resultados	60
4.9. Conclusiones	63
4.10. Observaciones	64
4.11. Recomendaciones para trabajos futuros	65

Lista de figuras

1.1.	(d) Curva de aprendizaje, tomada de [2]. (e) <i>DFM</i> y rendimiento, tomada de [3].	4
1.2.	(a) Sensitividad y espacio de diseño. (b) Compromiso rendimiento y desempeño.	7
1.3.	Varios niveles de variación de procesos. [4]	8
2.1.	(a) <i>Mismatch</i> entre dos transistores. (b) Distribución del <i>die</i> en la oblea.	16
2.2.	(a) Modelo del canal en ACM. (b) Perfil de la región de deflexión.	19
2.3.	(a) Circuito de medida del <i>mismatch</i> . (b) <i>Mismatch</i> de I_D vs. $V_{GS} - V_t$	21
2.4.	(a) Efecto de N_{oi} y de B_{ISQ} en el <i>mismatch</i> . (b) Comparación de los tres modelos.	25
3.1.	Flujo de diseño <i>DFM</i> adoptado para este trabajo.	32
3.2.	Gráfica de una función convexa y su respectiva prueba de convexidad.	40
3.3.	(a) Análisis de convexidad para la temperatura. (b) Análisis de convexidad para la tensión V_{SB} .	41
3.4.	Etapas fuente común con carga activa.	42
3.5.	(a) Espacio de diseño para <i>DFM</i> . (b) Centralización del diseño compatible con PG.	44
4.1.	Fuente de voltaje de referencia a optimizar, tomada de [1]	48
4.2.	Comportamiento de V_{ptat} en función de la temperatura de trabajo.	53
4.3.	Comportamiento de V_{ctat} para el rango de temperaturas de trabajo.	54
4.4.	Tensión V_{DS16} vs. temperatura	56
4.5.	Sensibilidad de los parámetros de desempeño de la fuente de referencia.	61
4.6.	Voltaje de referencia vs. temperatura.	62
4.7.	(a) Histograma para la tensión V_{ref} . (b) Histograma del coeficiente térmico.	63
4.8.	<i>Layout</i> de la fuente de referencia diseñada.	64

Lista de tablas

1.1. Compendio de trabajos de fuentes de referencia optimizadas.	10
2.1. Parámetros eléctricos y de proceso más relevantes en el <i>mismatch</i>	17
2.2. Resumen de los aspectos más relevantes para los tres modelos tratados.	28
4.1. Requerimientos de aceptabilidad para la fuente de referencia.	48
4.2. Mínima área y análisis de <i>mismatch</i>	50
4.3. Programa geométrico para el bloque <i>PTAT</i>	52
4.4. Dimensiones de los elementos del bloque <i>PTAT</i>	53
4.5. Comparación de resultados <i>PTAT</i>	53
4.6. Programa Geométrico para el bloque <i>CTAT</i>	55
4.7. Dimensiones de los elementos del bloque <i>CTAT</i>	55
4.8. Comparación de resultados <i>CTAT</i>	55
4.9. Programa geométrico para el bloque <i>PWL</i>	57
4.10. Dimensiones de los elementos del bloque <i>PWL</i>	57
4.11. Comparación de resultados <i>PWL</i>	57
4.12. Programa geométrico para la rama final.	59
4.13. Dimensiones de los elementos de la rama final.	59
4.14. Comparación de consumos de área	60
4.15. Resultados de desempeño y rendimiento obtenidos.	61

Capítulo 1

Introducción

Cuando Gordon Moore, en 1965, formula su muy conocida Ley,¹ los límites de la escalabilidad de componentes en un Circuito Integrado (CI) y su impacto en la confiabilidad del dispositivo no representaban una preocupación inmediata. A medida que se alcanzaron mayores escalas de integración y los valores nominales de dimensiones físicas, tensiones y corrientes disminuyeron, el diseño y la fabricación se tornaron altamente sensitivos a las variaciones del proceso de fabricación. Eventos físicos y eléctricos indeseados y de carácter aleatorio aumentaron la complejidad del diseño. Por lo tanto, topologías y diseños que pasaban los niveles de validación y confiabilidad tradicionales, podrían fácilmente fallar en tecnologías menores a $130nm$ [3].

En efecto, el proceso de manufactura en un CI es inherentemente imperfecto. Ligeras variaciones en la duración, la temperatura y las concentraciones químicas en cada paso, en conjunto con las tolerancias y defectos locales posteriores a la fabricación, generan desviaciones en el comportamiento eléctrico del sistema que compone el CI. Tales eventos son comunes a todos los circuitos analógicos, pero su impacto es más evidente en circuitos como las fuentes de voltaje de referencia, en donde se acentúa el costo de fabricación debido a la degradación del desempeño. Y más aún, la varianza de los parámetros del proceso y el *mismatch*² en los CIs es proporcional a la escalabilidad propuesta en la ley de Moore.

En consecuencia, las variaciones de proceso al tener un alto impacto en el rendimiento del circuito, obligan al diseñador a modificar la metodología de diseño. Los problemas de confiabilidad y robustez del circuito deben ser manejados en etapas mucho más tempranas en el flujo de diseño. De esta manera, el costo global de producción y el tiempo de ingreso al mercado del dispositivo no se ven incrementados debido a la alta probabilidad de obtención de un producto final confiable. Es así como surge una metodología complementaria de diseño denominada *Design For Manufacturing (DFM)*. Esta incluye un conjunto de técnicas para modificar el diseño de los circuitos integrados para hacerlos más confiables mejorando su

¹Determina que la densidad de transistores en un CI se duplica aproximadamente cada 18 meses.

²Término en inglés usado en CIs para indicar el descasamiento entre dos pares idénticos de elementos dispuestos en un *layout*.

desempeño funcional y paramétrico [2]. El reto básico del diseñador es establecer un diseño que asocie altas especificaciones de desempeño con entornos fiables ante variaciones del proceso, fabricación y tiempo de vida del dispositivo.

Para obtener un diseño analógico de altas especificaciones de desempeño, algoritmos de optimización matemática vía Programación Geométrica (PG) han sido aplicados [5,6,7]. Estos determinan el óptimo global en los problemas considerados factibles, pero están acotados a un subconjunto de funciones especiales. Sin embargo, un óptimo conjunto de especificaciones de desempeño comúnmente no garantiza un alto rendimiento funcional y paramétrico. El compromiso de diseño actual de CIs se resume en encontrar un punto perteneciente al espacio de diseño donde conjuntamente desempeño y rendimiento sean maximizados.

El presente proyecto evalúa los alcances y beneficios que tiene la implementación de la metodología *DFM* por medio del rediseño optimizado de la fuente de voltaje de referencia propuesta en [1]. Además, busca establecer los factores limitantes en la implementación *DFM*, relacionados comúnmente con la caracterización estadística y la frontera tecnológica del proceso. De esta forma, se concentra en proponer una metodología que condense la bondad en la consecución de óptimos globales de la programación geométrica, con la optimización del rendimiento funcional y paramétrico del dispositivo.

1.1. Motivación

El auge de los sistemas portátiles en su consigna de reducción de potencia, han impulsado la migración a regiones de polarización como la sub-umbral. Pero la variabilidad de la región de sub-umbral, la cual es la máxima en el dominio de polarización del transistor MOS [8], en conjunción con un deficiente modelado, han limitado contundentemente esta tendencia. De esta manera, las nuevas metodologías de diseño deben ser validadas usando circuitos como la fuente de voltaje de referencia propuesta en [1], la cual está polarizada en sub-umbral y cumple esencialmente con tres criterios de selección: alta variabilidad debida a *mismatch*, dependencia de factores del entorno como la temperatura y la fuente de alimentación, y presenta una topología acorde al estado del arte. Es de esta forma como la metodología del presente trabajo hace parte activa de los conocimientos y estrategias actuales que buscan contrarrestar las limitantes de la frontera tecnológica, para así contribuir a establecer soluciones viables al problema en el interior del grupo CIDIC.

Paralelamente, en los últimos años se ha evidenciado una clara tendencia en los avances tecnológicos de los dispositivos que contienen CIs; mientras se ha incrementado el desempeño promedio y decrementado el costo global del dispositivo, la complejidad del diseño y el proceso de fabricación ha deteriorado su rendimiento. Además, es importante notar que el costo de fabricación de dispositivos semiconductores continuará incrementando, en contraste con la relación costo/beneficio para el usuario, la cual naturalmente disminuirá.

El problema anterior lleva a una determinación: analizar los factores predominantes en la caída de rendimiento del dispositivo y su dependencia en los métodos y medios de diseño. En la figura 1.1(d) se observa la curva de Rendimiento vs. tiempo de dos procesos de diseño de un mismo dispositivo D_1 y D_2 . El área bajo esta curva indica el grado de adaptación del diseñador a las variaciones de proceso. El proceso D_2 sigue el flujo de diseño estándar con una validación de especificaciones al final del proceso. En el de D_1 , cada proceso es realimentado con los resultados de las etapas previas, en cada una de las cuales se evalúa en conjunto el desempeño y rendimiento del dispositivo siguiendo la metodología *DFM*. El diseño D_1 alcanza en un tiempo T_1 el mismo rendimiento aceptable o *Yield acceptable* (Y_{acc}) que D_2 obtiene en un tiempo mucho mayor, T_2 . Además D_1 obtiene un rendimiento final mucho más alto que D_2 [2]. En el caso de la figura 1.1(e) los resultados son más prometedores. Como era de esperarse, el rendimiento decrementa como lo hace el dimensionamiento característico del proceso. Además, es apreciable como más del 50 % de la caída total de desempeño esta relacionada con el diseño y un 20 % al proceso. La cantidad restante está vinculada con la impresión óptica de las capas debida a *RET*³ [3]. Estos resultados ligan de una increíble responsabilidad en el problema al diseñador de circuitos, compartida en menor proporción por los encarados del control del proceso.

Dada la amplia brecha entre la complejidad proyectada del dispositivo y las capacidades de diseño, las herramientas que el diseñador tiene a su disposición no le permiten afrontar este reto. El mejoramiento de la productividad de las herramientas de diseño ha estado en crecimiento a una tasa aproximada de 20 %/año, en comparación con la tasa de crecimiento exponencial cercana al 60 %/año en la complejidad de los dispositivos semiconductores [9]. Este déficit del 40 %, conduce a una inminente crisis en el diseño y fabricación de semiconductores. Además, las herramientas en conjunto con el correcto modelado de los dispositivos y las celdas, deben permitirle al diseñador estimar con intervalos de confianza muy estrechos el desempeño y rendimiento del CI antes de la fabricación.

El curso lógico a tomar, es pues, dedicar más esfuerzo al desarrollo de las herramientas y metodologías de diseño. Aunque *DFM* requiere además el estudio de complejas interacciones entre muchas actividades que incluyen análisis de mercadeo y la selección del producto y sus componentes, el reto más alto, cercano al 70 % del problema central esta radicado en el proceso que involucra al diseñador.

1.2. Diseño para Manufactura *DFM*

La metodología *DFM*, en el ámbito de los dispositivos semiconductores, se puede definir como una correcta interacción entre todas las fases de diseño y del proceso de fabricación, de manera tal que el desempeño y el rendimiento previstos en etapas previas a la fabricación se

³*RET*(Resolution Enhancement Techniques): Son técnicas para llevar el proceso litográfico de inscripción de máscaras a valores más pequeños y con mejor resolución.

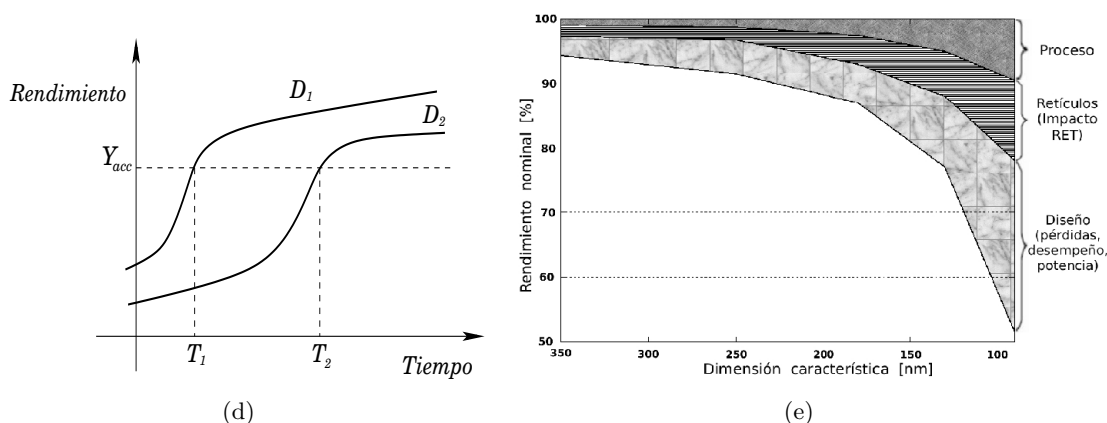


Figura 1.1: (d) Curva de aprendizaje, tomada de [2]. (e) *DFM* y rendimiento, tomada de [3].

vea reflejado en el CI fabricado. La filosofía del diseño para manufactura y sus prácticas son usadas en muchas industrias debido a que se reconoce que entre el 70 % al 90 % del costo total del producto es determinado antes de que el diseño sea lanzado a manufactura en serie.

Para evaluar la eficiencia de la metodología de diseño implementada, una compilación de aspectos relevantes en el marco del diseño para manufactura es realizada, destacándose así los siguientes:

Rendimiento: Se define rendimiento paramétrico de un circuito, Y , como la función dada por:

$$Y = \int_{\mathcal{F}} \Phi_y(\mathbf{x}) d\mathbf{x} \quad (1.1)$$

donde \mathbf{y} es el vector de características del dispositivo, \mathbf{x} es el vector de valores nominales de las variables de diseño, $\Phi_y(\mathbf{x})$ es la función de densidad de probabilidad conjunta y \mathcal{F} es la región de aceptabilidad en el espacio de las variables de diseño \mathbf{x} definida por las restricciones de aceptabilidad [10]. La región de aceptabilidad es un conjunto \mathcal{A}_y de valores \mathbf{y} definido como: $\mathcal{A}_y = \{\mathbf{y} | \mathbf{y}^L \leq \mathbf{y} \leq \mathbf{y}^U\}$. Los límites superior \mathbf{y}^U e inferior \mathbf{y}^L corresponden a los límites del peor caso de las características del dispositivo aceptables por el usuario y determinadas por el mercado. En términos comunes, el rendimiento se define como la proporción de circuitos fabricados que son funcionales y satisfacen los requerimientos de desempeño y aceptabilidad. En *DFM* generalmente se establece un intervalo de 6σ para el rendimiento del dispositivo, es decir, que por lo menos el 99,9997 % de los dispositivos fabricados se consideran a nivel funcional aptos para ser comercializados.

Desempeño: Además de garantizar una alta confiabilidad de los circuitos, el diseñador debe propender al mercadeo del dispositivo; para el usuario común esto significa tener un dispositivo con buenas prestaciones, a diferencia del diseñador, al cual le representa altas

especificaciones de desempeño en su diseño. Es frecuente que rendimiento y desempeño sean un reto para el diseñador, dado a que ambos se relacionan seriamente.

Funcionalidad: La funcionalidad corresponde a plasmar en los resultados del silicio el comportamiento global del circuito. Se dice que un circuito es funcional cuando cumple las restricciones estructurales y de diseño que determinan su región de polarización y su comportamiento respectivamente, sin necesidad de que las condiciones de aceptabilidad sean cumplidas.

Aceptabilidad: Se define aceptabilidad al intervalo en el dominio de las variables de desempeño en el cual, basados ciertas premisas de calidad, deberán estar los dispositivos aptos para la comercialización. La estimación del rendimiento final del circuito dependerá altamente de la confianza y estrechez de los intervalos de aceptabilidad adoptados.

Se deben categorizar dos tipos de aceptabilidad: la aceptabilidad a nivel de desempeño y la aceptabilidad a nivel de rendimiento. La aceptabilidad de desempeño fue definida anteriormente como el conjunto de valores permisibles de las variables de desempeño. Consecuentemente, la aceptabilidad de rendimiento esta definida como el conjunto:

$$\mathcal{F} = \{\mathbf{x} \in \mathcal{R}^n \mid \mathcal{P}(\mathbf{y} \pm \Delta y) \leq \mathbf{F}(n\sigma) - \mathbf{F}(-n\sigma)\} \quad (1.2)$$

donde $\mathcal{P}(z)$ es la probabilidad del suceso z bajo una función de distribución determinada, Δy es la tolerancia de cada una de las funciones de desempeño, \mathbf{F} es la función de distribución acumulativa y n es el número de σ determinados como marco del control de calidad.

Design For Testability (DFT): Está relacionado con la inclusión de ciertos elementos característicos en el diseño, de manera tal que faciliten la medición automatizada de varias funciones de prueba en pasos predeterminados de la línea de fabricación [9].

Tiempo de ingreso al mercado (*Time-to market*): Establece el intervalo de tiempo en que un circuito es proyectado como una necesidad palpable, hasta el tiempo en que este es lanzado al mercado. El tiempo de ingreso al mercado esta íntimamente ligado al conocimiento del proceso, del circuito y la metodología de diseño que se adopte.

Costo: El costo del producto a nivel del diseñador se determina por el rendimiento y el área del CI. La metodología *DFM* a este propósito propende por la reducción de costos por medio del aumento del rendimiento, la obtención de un área óptima para un mínimo *mismatch*, en conjunto con la puesta en marcha de una sola línea de fabricación para prototipaje.

1.3. Compromisos y retos en la implementación *DFM*

Durante la última década, la industria semiconductor ha apostado por la ejecución de prácticas relacionadas con la metodología *DFM*. Sin embargo, la implantación de una nueva filosofía de diseño complementaria al esquema tradicional involucra tiempo, recursos y riesgos. A nivel del diseñador, los problemas y retos emergen comúnmente dada la inclusión de nuevas variables de diseño y desempeño. Además, la optimización en el dominio del rendimiento compromete altamente el desempeño global del circuito.

Frontera tecnológica y espacio de diseño: ¿Como precisar las especificaciones de diseño máximas que otorga determinada topología en una tecnología? es tal vez la pregunta de mayor relevancia a la hora del mercadeo de un dispositivo. Para obtener una aproximación matemática del problema, se define espacio de diseño como la solución en el dominio de las variables de diseño $\mathbf{x} = (x_1, x_2, \dots, x_n)$ de la siguiente inecuación:

$$\mathbf{y}_i^L \leq f_i(\mathbf{x}) \leq \mathbf{y}_i^U \quad \text{para} \quad i = 1, 2, \dots, m \quad (1.3)$$

donde cada una de las funciones características $f_i(\mathbf{x})$ describen ya sea un parámetro eléctrico o de desempeño, en conjunto con \mathbf{y}_i^L y \mathbf{y}_i^U que representan los límites superior e inferior de la i -ésima función característica respectivamente.

El espacio de diseño involucra todos los puntos de las variables de diseño \mathbf{x} que garanticen un diseño funcional. Comúnmente, el espacio de diseño es altamente no lineal e inconvexo, razón por la cual debe ser aproximado a una región multiplanar convexa [10]. Derivando y operando sobre (1.3) se obtiene:

$$\frac{\Delta \mathbf{y}_i^L}{\partial f_i(\mathbf{x})} \leq \Delta \mathbf{x} \leq \frac{\Delta \mathbf{y}_i^U}{\partial f_i(\mathbf{x})} \quad \text{para} \quad i = 1, 2, \dots, m \quad (1.4)$$

La ecuación (1.4) establece una dependencia del desplazamiento del espacio de diseño con la sensibilidad de las restricciones estructurales y de diseño. Como se aprecia en la figura 1.2(a), cada una de las sensibilidades determina la distancia entre el diseño inicial x_c y la frontera en el plano x_i del espacio de diseño. A la vez, la correspondencia entre varias sensibilidades, determina las fronteras entre planos del espacio de diseño. Conforme la tecnología se escala, muchas de las sensibilidades suelen aumentar, produciendo un efecto adverso al diseñador: la reducción del espacio de diseño.

Máximo global y rendimiento: La figura 1.2(b) es una representación gráfica del espacio de diseño de un dispositivo determinado. Por simplicidad, se establecen variaciones aleatorias simétricas alrededor del punto de diseño \mathbf{x}_c . Se supone ahora la existencia de un máximo global de desempeño dada una función de mérito en \mathbf{x}_{cm} y se establece el centro de

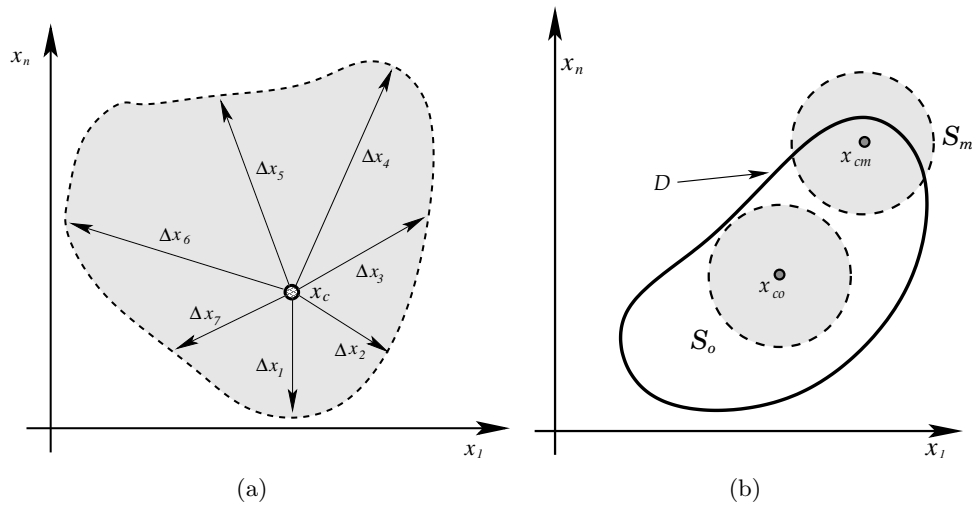


Figura 1.2: (a) Sensibilidad y espacio de diseño. (b) Compromiso rendimiento y desempeño.

la región D en x_{co} . Tanto x_{cm} y x_{co} son el centro de una región esférica de dimensión n , en la cual están inscritos el 99,9997 % 6σ de los dispositivos fabricados con estas variables de diseño. Para el caso de la región S_m , aproximadamente el 40 % de circuitos están ubicados por fuera del espacio de diseño, es decir, su funcionalidad es incorrecta. En el caso de la región S_o , por lo menos el 99,9 % de diseños están incluidos en la región de diseño, pero a diferencia de x_{cm} , el punto x_{co} no es un máximo global y posiblemente no cumpla con las restricciones de desempeño previstas por el diseñador.

Para contrarrestar este efecto, una solución práctica sería encontrar el punto más cercano a x_{co} bajo el cual se cumpla con las especificaciones mínimas de desempeño. La consecución de este punto representa un compromiso alto ligado al diseño. Es importante aclarar que en ciertos casos la región de diseño es tan estrecha, que ningún punto perteneciente a esta logra inscribir el círculo de 6σ ; para tales casos la solución correspondería a un cambio o modificación en la topología del circuito, un cambio del proceso tecnológico o una reducción de las condiciones de aceptabilidad.

Herramientas CAD: La conjunción de la metodología *DFM* con las herramientas *CAD*⁴ es considerada como un área de promisoria investigación, dado el alto déficit en la productividad del *Software* para el diseño de CIs. Con el fin de mejorar los procesos de diseño a través de las herramientas *CAD*, y así facilitar la fabricación de circuitos integrados en [9] se propone:

1. Profundizar en la física del dispositivo, para determinar mejor el desempeño de los elementos del circuito en un modelo que considere las variaciones de proceso.

⁴Siglas en inglés para *Computer Aided Design* (Diseño asistido por computador).

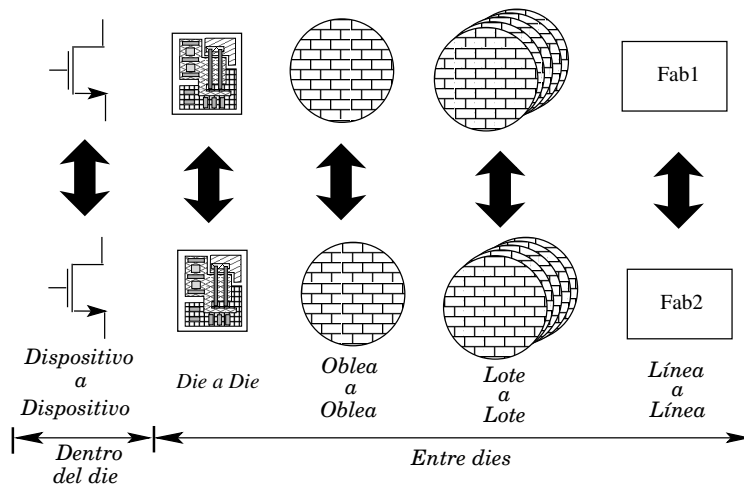


Figura 1.3: Varios niveles de variación de procesos. [4]

2. Relacionar los modelos de rendimiento, con las distribuciones estadísticas de los efectos observados en cada uno de los pasos de fabricación.
3. Desarrollar estructuras de *software* a través de las cuales puedan ser integrados varios modelos y simulaciones al proceso de diseño y fabricación.
4. Desarrollar el modelado de celdas y líneas de fabricación en conjunto con la herramientas de simulación, basados en una mejor comprensión de las interacciones entre los componentes del proceso y el rendimiento del sistema.

Modelado: Con el propósito de estimar acertadamente el comportamiento eléctrico y los factores relevantes en el desempeño y rendimiento de un CI, el proceso de fabricación debe ser correctamente caracterizado. La caracterización de la tecnología debe ser llevada a cabo en dos ámbitos: el modelado eléctrico y el modelado estadístico [4]. En el primero cada parámetro representa ya sea una cantidad física o un factor de ajuste matemático. Su caracterización comúnmente involucra un conjunto pequeño de dispositivos muestra. En el caso del segundo, la caracterización requiere de un conjunto de muestras lo suficientemente alto para una validación post-fabricación y una posible estandarización a nivel de diseño.

Las variaciones de proceso deben ser analizadas en varios niveles como se muestra en la figura 1.3. Dentro del mismo *die*⁵ y entre distintos *dies*. Las que tienen lugar en el mismo *die* afectan a los dispositivos y obedecen a defectos locales. Las que tienen lugar entre *dies*, son considerados globales al *chip* y pueden llegar a tener múltiples orígenes. Por ejemplo, las variaciones de *die a die* son debidas esencialmente a gradientes de concentración en la oblea, y las presentes entre obleas atribuidas ampliamente al desalineamiento de las máscaras [4].

⁵Un *die* es un complejo conjunto de componentes electrónicos y sus interacciones impresos y enclavados en una pequeña pieza de material semiconductor, constituyendo así un circuito discreto.

El conjunto de datos recolectados debe ser analizado estadísticamente y correlacionado con ciertas variables como la polarización y el dimensionamiento del circuito. Es frecuente que los fabricantes otorguen únicamente la caracterización a nivel de *dies*, siguiendo el modelo propuesto por Pelgrom [11]. Sin embargo, para tecnologías por debajo de los $130nm$, las alteraciones eléctricas causadas por *mismatch* no se ajustan al modelo propuesto en [11]. Además, este modelo no es válido para cualquier región de operación como inversión débil, por lo que las expresiones de *mismatch* de los transistores deben ser redefinidas en función de parámetros de modelos continuos en toda región de operación como lo es el modelo *ACM* (*Advanced Compact Model*) [12].

En este sentido, *DFM* requiere el planteamiento de nuevos modelos tendientes a incluir fenómenos físicos relevantes en tecnologías submicrón y predecir con una alta correlación el comportamiento eléctrico en cada uno de los niveles de variación de proceso.

De esta forma, establecidas las pautas del diseño para manufactura esenciales a este trabajo, se procederá con una indagación bibliográfica de trabajos relacionados al contexto de la optimización de fuentes de referencia, tanto a nivel de desempeño como de rendimiento.

1.4. Algunas fuentes de referencia optimizadas

A nivel académico existen pocos trabajos que traten acerca de la optimización de fuentes de voltaje de referencia. En la tabla 1.1 se pueden observar los trabajos mas relevantes acerca de la optimización del diseño de fuentes de referencia. Bernardo en [13] describe una estrategia secuencial para diseñar circuitos comercializables entre los cuales se encuentra una fuente de corriente de referencia, usando herramientas *CAD* disponibles. Bernardo adopta el método de Taguchi [14] para minimizar la máxima perdida de rendimiento de esta fuente. Sobre esta formulación define un rango de valores permisibles para cada una de las variables de diseño y las variaciones del entorno como la temperatura y la tensión de polarización. Ignora la variación del proceso de fabricación en el diseño, aduciendo una baja influencia en las condiciones de operación del circuito. Nam en [15] presenta una fuente de voltaje de referencia optimizada con programación evolucionaria.⁶ La función costo a minimizar, la construye asignando a cada pareja de diferencia de voltajes, a temperaturas específicas, un peso específico. Nam no propone ninguna estrategia para controlar las variaciones de proceso; además, la programación evolucionaria no garantiza la obtención de un óptimo global ni una fácil inclusión de las restricciones estructurales dentro de la estrategia de diseño.

Tajalli en [16] propone una función de mérito dependiente del área, la potencia y la relación de dos resistencias para una fuente de referencia *Bandgap*. Esta metodología realiza un barrido de la función mérito a través de ciertos valores permisibles que garanticen la disposición de dos transistores BJT en centroide común. Tajalli no establece una metodología clara para contrarrestar las variaciones de proceso, que incluyen principalmente la tensión de *offset* del

⁶Tipo de optimización estocástica, que usa el concepto de la evolución y la selección natural.

Año	Autor	Ref.	Optimización	Estrategia <i>DFM</i>
1992	Maria Bernardo	[13]	Experimentación secuencial	Método de Taguchi [14] ^(a) ^(b)
2001	Dongkyung Nam	[15]	Programación evolucionaria	Ninguna
2004	Armin Tajalli	[16]	Secuencial semiautomática	Centroide común
2004	Robert Gregoire	[17]	Algebraica (Área)	Mínima área vs. <i>mismatch</i>
2005	Carlo Roma	[18]	Mínimos cuadrados y distancia de parámetros	<i>WiCkeD</i> ^(c)

(a) Fuentes de corriente.

(b) Optimización únicamente aplicada a variación de procesos.

(c) *Software* de la empresa *MunEDA* para el análisis *DFM/DFY*.

Tabla 1.1: Compendio de trabajos de fuentes de referencia optimizadas.

amplificador operacional, el cual no se incluye en el proceso de optimización. Gregoire en [17] plantea algebraicamente la relación entre el área que debe tener cada dispositivo con los parámetros del proceso, de manera tal que se garantice la mínima área posible. Basado en lo anterior, Gregoire diseña una fuente Bandgap reduciendo el área activa con respecto al diseño original (sin consideraciones) en 44% sin degradar la desviación estándar del voltaje. Este método en varias ocasiones genera valores numéricos de dimensiones que no son fabricables, requiriendo un posterior ajuste; además esta estrategia no evalúa ni considera el desempeño global del circuito. Finalmente, Roma en [18] hace uso de la herramienta *WiCkeD* con su respectivo flujo de diseño para optimizar el rendimiento y las especificaciones de una fuente de referencia *Bandgap* en $0,13\mu\text{m}$, usada como referencia para la generación de un voltaje interno en una memoria *flash*. En este trabajo se llevaron a cabo dos diseños: El primero para realizar un ajuste entre los modelos de simulación con los resultados de silicio; el segundo se enfocó en el mejoramiento del rendimiento. Lo anterior indica una deficiencia de la herramienta *WiCkeD* en la estimación del rendimiento y el desempeño, que se deriva de un incorrecto modelado.

Apoyados en los anteriores trabajos, y dejando de lado las limitantes correspondientes a cada una de las metodologías de diseño aplicadas, se evidencia una deficiente estimación del desempeño y el rendimiento; el efecto de variación de parámetros o bien fue despreciado como en el caso de Bernardo [13], o fue mal estimado como en el caso de Roma [18]. Intentar proponer una metodología sin una evaluación previa de los modelos estadísticos, probablemente producirá considerables desviaciones entre los resultados esperados y los medidos. Dado a que se espera contrarrestar los resultados de este trabajo con una posterior fabricación, es ineludible un estudio detallado de los modelos estadísticos, el cual incluye una selección enfocada al contexto de las necesidades de este trabajo.

1.5. Organización del Documento

En el presente trabajo la propuesta de una metodología de diseño para manufactura (*DFM*) será abordada en el siguiente orden:

En el capítulo 2 serán discutidos los modelos del *mismatch* más relevantes y útiles al contexto del presente trabajo. De esta manera, cada modelo será analizado y puesto en consideración bajo 5 premisas esenciales, cada una de las cuales evalúan su desempeño en la implementación del flujo de diseño a establecer por este trabajo. De esta manera serán seleccionados dos de los modelos aquí abordados: uno para la fase previa al dimensionamiento y el otro con el propósito de validación final del diseño.

En el capítulo 3 y basados en los resultados de la selección de los modelos de *mismatch*, se propone una metodología de diseño para manufactura apoyada del uso de la programación geométrica. Tanto rendimiento como desempeño son optimizados en un procedimiento que involucra la centralización del diseño.

Finalmente, en el capítulo 4 son condensados de manera comparativa los resultados de una fuente de referencia diseñada bajo la metodología *DFM* y otra bajo el esquema clásico de diseño. Posteriormente, un *layout* del circuito propuesto es diseñado, de manera que puedan ser evidenciados los efectos de una fabricación. Por último, son planteadas algunas conclusiones, observaciones y recomendaciones útiles para futuros trabajos.

Capítulo 2

Modelado estadístico del proceso

El *mismatch* es considerado como uno de los factores altamente limitantes en el desempeño de los circuitos analógicos de alta precisión e incluso de los circuitos digitales. Se dice que las variaciones de proceso marcaran el último límite en la escalabilidad de los CIs. Consecuentemente, las variaciones aleatorias consideradas despreciables en escalas de integración anteriores, serán relevantes en las futuras generaciones de procesos. Pero el problema principal no radica en la alta variabilidad exhibida en el proceso, sino en la incertidumbre para predecir el comportamiento estocástico del conjunto de parámetros de un transistor MOS. El modelo simplista del *mismatch* de la forma $1/\sqrt{WL}$, ha demostrado deficiencias para escalas de integración sub-micrón; además no es consistente con todos los parámetros físicos del proceso debido a que cada uno de estos exhibe una distinta dependencia con el dimensionamiento.

Un conjunto de nuevos modelos de *mismatch* han sido propuestos pero muchos de estos presentan deficiencias en su implementación. Frecuentemente, la aplicabilidad de modelos estadísticos de alta correlación con los resultados de silicio representa un compromiso en tiempo de computo, dada la inminente inclusión del amplio conjunto de parámetros variables en complejas ecuaciones comportamentales. Además lo hacen prácticamente inutilizable para los cálculos manuales en el diseño. Al no tener un estimativo correcto del *mismatch*, el diseñador se ve obligado a determinar a tientas el rendimiento del circuito. La desestimación de las geometrías del transistor produce amplias desviaciones en sus parámetros eléctricos. En un orden contrario, el uso de geometrías de transistor conservadoras tiene como consecuencia el desaprovechamiento del área que además conlleva al aumento de las capacitancias [19]. Este tipo de compromisos entre el desempeño y el rendimiento de los circuitos debe ser manejado a través de modelos estadísticos de alta exactitud en etapas previas a la fabricación.

Otro problema, en apariencia preocupante, surge de la derivación misma del modelo comportamental del transistor MOS y de su inevitable dependencia en la deducción de un modelo de *mismatch*. Por ejemplo, el modelo BSIM3v3, apoya un importante número de parámetros en un ajuste matemático carente de algún sentido físico y fundamentado en la lógica del número mágico. Pero sigue siendo aún más preocupante, que modelos de *mismatch* sean

derivados de las ecuaciones cuadráticas de primer nivel, que cada vez están más lejos de ajustarse al comportamiento de un canal de dimensiones nanométricas y peor aún de un transistor en sub-umbral. Lo anterior puede indicar una profunda estandarización en la implementación de modelos del *mismatch* fundamentada en el modelo de Pelgrom [11], que probablemente se justifica en el esquema del aumento desmesurado del área para el control de la variabilidad, pero que de lleno no abarca las expectativas de procesos futuros ni brinda soluciones efectivas y precisas al problema.

A este trabajo le resulta indispensable el tratamiento y la discusión de los más importantes modelos disponibles a nivel académico, ya que fundamentados en una elección argumentativa, se procederá a aplicarla tanto en la metodología de optimización de rendimiento y desempeño, como en la etapa más importante de todo el proceso: la validación a nivel de simulación y de fabricación. Los resultados de rendimiento y desempeño que se expondrán al final de este trabajo, los cuales evaluarán su precisión y contundencia, serán dependientes de la elección del modelo. Esto debe motivar a la industria a la escogencia o el planteamiento de un modelo que se encamine como estándar en el diseño analógico y digital, y que incluya adecuadamente los retos de la variabilidad futura.

Para la correcta inclusión del fenómeno del *mismatch* en el proceso de diseño, para este trabajo un modelo estadístico de proceso debe satisfacer por lo menos los siguientes aspectos:

Significado físico: Dado a que la formulación del *mismatch* en función de parámetros eléctricos origina una sobre-estimación de este, un modelo acertado debe ser planteado desde la variación de los parámetros físicos esenciales.

Continuidad: La continuidad inherente a los procesos físicos y naturales, los posibilita a ser modelables matemáticamente. Además, dada la consigna de baja potencia y bajo voltaje en el diseño, el modelo debe incluir regiones como la sub-umbral y la inversión moderada. Para esto, el *mismatch* debe ser replanteado como una función definida por partes o su modelo debe estar sustentado en un modelo eléctrico continuo en función de la polarización.

Medibilidad: Muchos de los parámetros físicos de un transistor no pueden ser directamente medibles, por lo que muchas veces se requiere o bien un ajuste matemático en el proceso de modelado o una medición indirecta del mismo. Un modelo altamente medible debe establecer un conjunto de pruebas que garantice resultados repetibles y una medición precisa de los parámetros físicos que involucra.

Precisión: Capacidad de predecir con bajas tasas de error el comportamiento tanto determinístico como aleatorio de cada uno de los parámetros eléctricos. El comportamiento determinístico será relevante en la estimación del desempeño del circuito, en contraste con el comportamiento estocástico, el cual será útil para la estimación del rendimiento.

Utilidad: Un modelo estadístico preciso no garantiza una fácil aplicabilidad en el flujo de diseño, ni mucho menos una utilización en cálculos manuales. Los modelos que precisen de pocos parámetros, expresiones matemáticas simples y tiempos de convergencia bajos son los más aptos para el diseño. Además, este debe establecer los componentes que serán incluidos en simulaciones *pre-layout* y en la extracción *post-layout*.

Escalabilidad: Se debe garantizar que el modelo sea independiente del tipo de proceso utilizado y de la longitud mínima de canal. Consecuentemente, debe permitir la inclusión de nuevos fenómenos físicos sin un drástico cambio en su formulación.

Debido a que la optimización del rendimiento requiere el correcto conocimiento de la variabilidad de cada uno de los parámetros, el modelado del *mismatch* debe ser correctamente aplicado al diseño. Para ello, una indagación bibliográfica de los modelos más relevantes al contexto del diseño para manufactura de una fuente de referencia polarizada en sub-umbral es requerida.

2.1. Modelos de *mismatch*

El tema del *mismatch* ha empezado a ser recurrente durante los últimos años, llegando a publicar más de una docena de modelos al respecto. El modelo que se considera pionero fue propuesto por Lakshmikumar en [20] y brindó un punto de partida para el tratamiento y derivación matemática del *mismatch* aplicado al diseño analógico de alta precisión. Basados en este trabajo surgieron modelos como el de Pelgrom [11], el cual sería ampliamente aceptado por la industria. Modelos posteriores fueron considerados como ampliaciones de [11], en donde se incluían los fenómenos físicos que comenzaban a ser relevantes en la estimación del *mismatch*. Posterior a esto surgen una serie de trabajos [19, 8], que tienen la ventaja de modelar continuamente el *mismatch* del transistor MOS. Hasta el momento el único modelo que se ha mantenido como estándar en la industria es Pelgrom, pero las exigencias actuales están guiando a un inminente cambio.

La selección de los modelos conduce a tres trabajos que son considerados fundamentales, relevantes e innovadores en el análisis y determinación del *mismatch*.

2.1.1. Modelo de Pelgrom

Propuesto en [11] durante el año de 1989, se convirtió en el primer modelo de *mismatch* estandarizado por la industria, debido a su precisión y aplicabilidad adecuadas a las escalas de integración de su momento. Por primera vez son distinguidas matemáticamente las variaciones globales y locales. Además, se reconocen algunas de las causas físicas que originan el *mismatch* en la corriente como lo son: efectos de borde, cargas del implante y de superficie, efectos de óxido y efectos en la movilidad. Sobre estas consideraciones, se establece una discusión del *mismatch* basada en el análisis y las mediciones del voltaje umbral, el factor de corriente y el

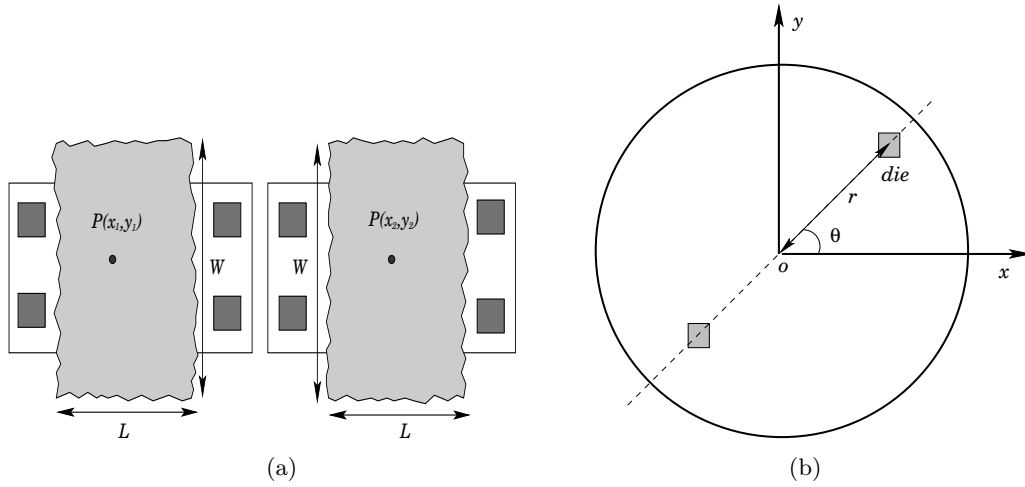


Figura 2.1: (a) *Mismatch* entre dos transistores. (b) Distribución del *die* en la oblea.

factor de sustrato del transistor MOS como función del área, la distancia y la orientación. Las mediciones son usadas para validar la teoría y derivar constantes desconocidas de las expresiones matemáticas. Su correcto análisis, medición y extrapolación lo convirtieron en el modelo más representativo del *mismatch* y uno de los trabajos más indexados a nivel de la microelectrónica.

Para Pelgrom, el valor de un parámetro P está compuesto de un término constante y otro variable aleatoriamente. Para el ámbito estocástico, se establecen variaciones locales como ruido blanco espacial caracterizado por una baja correlación con la distancia. Conjuntamente, dado a que la ubicación del *die* en la oblea después del empaquetado es desconocido, el efecto de la distribución circular en el *mismatch* puede ser modelado como un proceso estocástico adicional de una amplia correlación con la distancia. En la figura 2.1(b), se define la posición de un *die* en una oblea en términos del ángulo θ y su distancia al centro r . La función de densidad de probabilidad para el ángulo θ se caracteriza por ser constante en el intervalo $[-\pi, \pi]$, mientras que la función de densidad de probabilidad de la distancia r es razonablemente aproximada por medio de una distribución normal.

De esta manera, Pelgrom define la varianza en la desviación del parámetro P de los transistores mostrados en la figura 2.1(a), con ancho W , longitud L y espaciamiento D_x a lo largo de x , como:

$$\sigma^2(\Delta P) = \frac{A_P^2}{WL} + S_P^2 D_x^2 \quad (2.1)$$

donde A_P es la constante de proporcionalidad con el área para el parámetro P , mientras que S_P describe la variación del parámetro P con el espaciamiento. Las constantes de proporcionalidad son definidas por medio de la varianza de los parámetros. Cualitativamente, las variaciones locales decrecientan como el tamaño del dispositivo incrementa, puesto que los

Parámetros de proceso	Parámetros eléctricos
Voltaje de banda plana (V_{fb})	Corriente de dreno (I_D)
Movilidad (μ)	Voltage V_{GS}
Concentración de dopantes del sustrato (N_{SUB})	Transconductancia (g_m)
Desviación de longitud (ΔL)	Transconductancia de salida g_o
Desviación del ancho (ΔW)	
Efecto de canal corto (V_{tl})	
Efecto de ancho estrecho (V_{tw})	
Espesor del óxido de puerta (t_{ox})	
Resistencia de dreno-fuente (p_{sh})	

Tabla 2.1: Parámetros eléctricos y de proceso más relevantes en el *mismatch*.

parámetros se promedian sobre una gran área o distancia.

Pelgrom establece la expresión del mismatch en la corriente de dreno de los transistores M_i y M_{i+1} del circuito de la figura 2.3(a) como:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{4\sigma_{V_t}^2}{(V_{GS} - V_t)^2} + \frac{\sigma_{\beta}^2}{\beta^2} \quad (2.2)$$

2.1.2. Modelo de Drennan

Usado exclusivamente durante varios años por Motorola, este modelo surge como respuesta a una serie de falencias en la implementación y estimación del *mismatch* bajo el modelo de Pelgrom. Intenta describir en un sentido más físico el fenómeno del *mismatch*, obteniéndose de esta manera un modelo aplicable a un amplio rango en la polarización y las condiciones de geometría. Un aspecto clave de este modelo es que el *mismatch* es caracterizado bajo las mismas herramientas y el modelo comportamental usados en el diseño. Ratifica la dependencia del dimensionamiento con la varianza de los parámetros establecida por Pelgrom, pero a pesar de esto difiere considerablemente de los resultados obtenidas por este. Además discrepa que el aumento del área en todos los casos genere una inminente disminución del *mismatch*. Sus análisis y resultados son acertados y pertinentes, pero tiene la desventaja de ser lo suficientemente complejo para no ser implementado en ciertas fases del diseño.

Para el modelado del *mismatch*, Drennan considera dos tipos de parámetros: parámetros de proceso y parámetros eléctricos. Los parámetros de proceso son aquellos físicamente independientes de los parámetros que controlan el comportamiento eléctrico del dispositivo. Por su parte, los parámetros eléctricos son aquellos parámetros que son de interés al diseñador. La tabla 2.1 contiene ejemplos de parámetros de proceso incluidos en este modelo y los parámetros eléctricos relevantes. Por ejemplo, V_t no es un parámetro de proceso ya que depende de V_{fb} , t_{ox} , N_{SUB} , L y W . El parámetro V_t exhibe una dependencia directa con V_{fb} y t_{ox} , inversa con W y L y logarítmica con N_{SUB} . Lo anterior significa que la estimación de la varianza de V_t vía relación Pelgrom es física y matemáticamente incorrecta.

Fundamentado en la dependencia de la varianza de cada parámetro con el dimensionamiento, en conjunto con la teoría de propagación de la varianza (POV), Drennan establece:

$$\begin{bmatrix} \sigma_{\tilde{I}_{d1}}^2 \\ \sigma_{\tilde{I}_{d2}}^2 \\ \sigma_{\tilde{I}_{d3}}^2 \\ \vdots \\ \sigma_{\tilde{I}_{dn}}^2 \end{bmatrix} = \begin{bmatrix} \left(\frac{d\tilde{I}_{d1}}{d\Delta W}\right)^2 & \left(\frac{d\tilde{I}_{d1}}{dt_{ox}}\right)^2 & \left(\frac{d\tilde{I}_{d1}}{dV_{fb}}\right)^2 & \left(\frac{d\tilde{I}_{d1}}{d\tilde{\mu}_o}\right)^2 & \left(\frac{d\tilde{I}_{d1}}{d\Delta L}\right)^2 & \left(\frac{d\tilde{I}_{d1}}{dV_{tL}}\right)^2 & \left(\frac{d\tilde{I}_{d1}}{d\tilde{\rho}_{sh}}\right)^2 & \left(\frac{d\tilde{I}_{d1}}{dN_{sub}}\right)^2 \\ \left(\frac{d\tilde{I}_{d2}}{d\Delta W}\right)^2 & \left(\frac{d\tilde{I}_{d2}}{dt_{ox}}\right)^2 & \left(\frac{d\tilde{I}_{d2}}{dV_{fb}}\right)^2 & \left(\frac{d\tilde{I}_{d2}}{d\tilde{\mu}_o}\right)^2 & \left(\frac{d\tilde{I}_{d2}}{d\Delta L}\right)^2 & \left(\frac{d\tilde{I}_{d2}}{dV_{tL}}\right)^2 & \left(\frac{d\tilde{I}_{d2}}{d\tilde{\rho}_{sh}}\right)^2 & \left(\frac{d\tilde{I}_{d2}}{dN_{sub}}\right)^2 \\ \left(\frac{d\tilde{I}_{d3}}{d\Delta W}\right)^2 & \left(\frac{d\tilde{I}_{d3}}{dt_{ox}}\right)^2 & \left(\frac{d\tilde{I}_{d3}}{dV_{fb}}\right)^2 & \left(\frac{d\tilde{I}_{d3}}{d\tilde{\mu}_o}\right)^2 & \left(\frac{d\tilde{I}_{d3}}{d\Delta L}\right)^2 & \left(\frac{d\tilde{I}_{d3}}{dV_{tL}}\right)^2 & \left(\frac{d\tilde{I}_{d3}}{d\tilde{\rho}_{sh}}\right)^2 & \left(\frac{d\tilde{I}_{d3}}{dN_{sub}}\right)^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \left(\frac{d\tilde{I}_{dn}}{d\Delta W}\right)^2 & \left(\frac{d\tilde{I}_{dn}}{dt_{ox}}\right)^2 & \left(\frac{d\tilde{I}_{dn}}{dV_{fb}}\right)^2 & \left(\frac{d\tilde{I}_{dn}}{d\tilde{\mu}_o}\right)^2 & \left(\frac{d\tilde{I}_{dn}}{d\Delta L}\right)^2 & \left(\frac{d\tilde{I}_{dn}}{dV_{tL}}\right)^2 & \left(\frac{d\tilde{I}_{dn}}{d\tilde{\rho}_{sh}}\right)^2 & \left(\frac{d\tilde{I}_{dn}}{dN_{sub}}\right)^2 \end{bmatrix} \begin{bmatrix} \sigma_{\Delta W/L}^2 \\ \sigma_{t_{ox}/(LW)}^2 \\ \sigma_{V_{fb}/(LW)}^2 \\ \sigma_{\tilde{\mu}_o/(LW)}^2 \\ \sigma_{\Delta L/W}^2 \\ \sigma_{V_{tL}/W}^2 \\ \sigma_{\tilde{\rho}_{sh}/(LW)}^2 \\ \sigma_{N_{sub}/(LW)}^2 \end{bmatrix} \quad (2.3)$$

El signo (\sim) arriba indica una variable normalizada. El vector del lado izquierdo de (2.3) es un conjunto de n desviaciones estándar recolectadas de muchos *dies*, en varias condiciones de polarización y geometría. La combinación de condiciones es establecida de manera que cada una de las varianzas sea medida en sus casos más significativos. Por ejemplo, ρ_{sh} sólo afecta considerablemente a la corriente en los dispositivos de bajo W , polarizados en la región de triodo para altos valores de V_{GS} , dado a que al no haber estrangulamiento, el efecto de la resistencia de difusión ρ_{sh} empieza a prevalecer con respecto al efecto de la resistencia de canal r_{ds} . De esta manera, las medidas son tomadas teniendo en cuenta este tipo de consideraciones.

La matriz en (2.3), contiene el cuadrado de las sensibilidades de I_d con respecto a cada uno de los parámetros de proceso. Estas sensibilidades son numéricamente evaluadas usando SPICE en las condiciones de geometría y polarización correspondientes a la medida.

Determinados los valores de las dos matrices en (2.3), el vector del lado derecho puede ser calculado usando una regresión lineal. Básicamente, cada parámetro se asume como independiente, pero si el método determina una correlación significativa entre estos, implicará que un conjunto de parámetros incorrecto ha sido tomado.

2.1.3. Modelo Galup-Montoro

Este modelo fue propuesto como parte integral de una nueva generación de modelos compactos, que tienen la ventaja de modelar continuamente cualquier punto de operación del transistor desde inversión débil hasta inversión fuerte. La derivación de sus expresiones es hecha a partir del modelo eléctrico planteado igualmente por Galup-Montoro en [12] denomi-

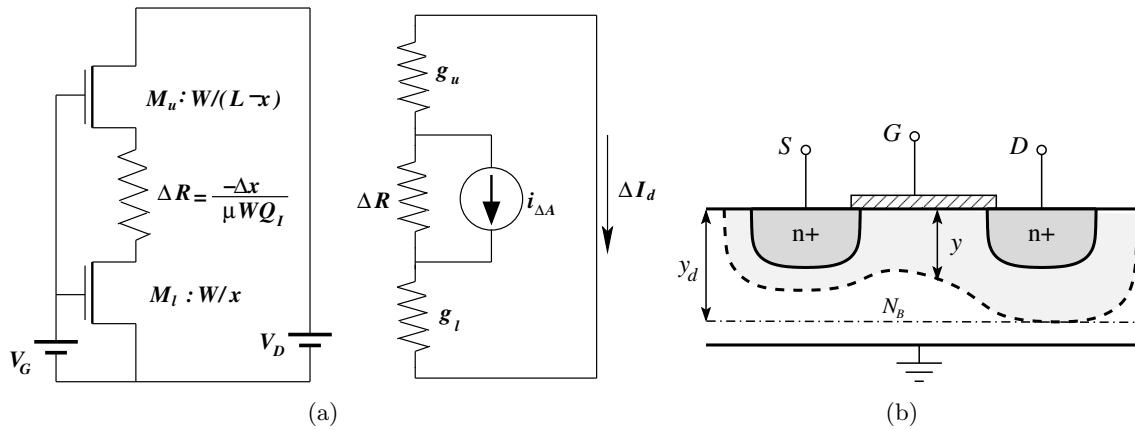


Figura 2.2: (a) Modelo del canal en ACM. (b) Perfil de la región de deflexión.

nado *ACM* (*Advanced Compact Model*). Dado a que el impacto de las fluctuaciones locales de dopantes en el *mismatch* son significativas [21], este modelo deriva sus expresiones por medio de la teoría de la fluctuación en el número de portadores, usada también para determinar el ruido $1/f$ del transistor en el modelo *ACM*.

Este modelo no incluye los efectos de la degradación en la movilidad, la velocidad de saturación y la resistencia en serie. Estos efectos llevan a expresiones complejas que no permiten abordar una discusión clara, además, las mediciones hechas por modelos anteriores [22, 19] han mostrado una alta dependencia de la corriente con las variaciones en V_t , las cuales son ampliamente atribuidas a las fluctuaciones en los dopantes.

Las fluctuaciones en la corriente de drenaje resultan de la suma de todas las fluctuaciones locales a lo largo del canal, cualquiera que sea su origen. Dado a la no uniformidad en los efectos del canal, el transistor es modelado como tres dispositivos dispuestos en serie como se muestra en la figura 2.2(a): un transistor superior, un transistor inferior y un pequeño elemento de canal de longitud Δx , representado por una resistencia. El propósito de esta fragmentación es el promediar de una manera mas acertada los efectos individuales subyacentes a nivel del canal.

Existen básicamente dos tipos de fluctuaciones de los dopantes: la globales y las locales. La primera se refiere al valor promedio a lo largo del canal, y es la responsable de establecer el denominado perfil de dopaje como se muestra en la figura 2.2(b), el cual es altamente dependiente de la tecnología. Mientras tanto, las fluctuaciones de dopantes locales ocurren en pequeñas secciones del canal, sumándose en forma de ruido blanco a lo largo del perfil de dopaje [21].

Para modelar matemáticamente el efecto del perfil de dopaje, se define la constante N_{oi} como el valor que pondera verticalmente la deflexión de carga, matemáticamente sería:

$$N_{ci} = \int_0^{y_d} N_a \left(1 - \frac{y}{y_d}\right)^2 dy \quad (2.4)$$

donde N_A es la concentración neta de dopantes tanto aceptores como donadores en un volumen elemental de la capa de deflexión.

Galup-Montoro establece que la varianza de la corriente de un transistor MOS en sub-umbral, saturación o triodo en términos de parámetros del modelo *ACM* esta determinada por:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{ci}}{WLN^{*2}} \frac{1}{i_f - i_r} \ln \left(\frac{1 + i_f}{1 + i_r} \right) \quad (2.5)$$

donde se define a N^* como:

$$N^* = \frac{nC_{ox}\phi_t}{q} \quad (2.6)$$

Podría pensarse en una reducción de la varianza de la corriente por medio de un control del proceso. Por ejemplo una reducción de impurezas o en una polarización del cuerpo para disminuir el valor de la distancia y_d .

La ecuación (2.5) puede ser simplificada para unas condiciones específicas. En el caso de inversión débil, $i_f \cong i_r$ y $i_f \ll 1$, así (2.5) se convierte para inversión débil en:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{N_{ci}}{WLN^{*2}} \quad (2.7)$$

Incluyendo los efectos de primer orden en la variación de la movilidad, el grosor del oxido de puerta y el factor de pendiente se define B_{ISQ} , incluyendo este factor en (2.6) para toda región se tiene:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{1}{WL} \left[\frac{N_{oi}}{N^{*2}} \frac{1}{i_f - i_r} \ln \left(\frac{1 + i_f}{1 - i_f} \right) + B_{ISQ} \right] \quad (2.8)$$

2.2. Discusión

Bajo cada uno de los fundamentos establecidos de los modelos elegidos, se realiza un análisis argumentativo enfocado a cada uno de los ámbitos abordados al inicio del capítulo. La idea básica es establecer las deficiencias y cualidades de cada uno de los modelos, para de esta manera efectuar la elección.

2.2.1. Significado físico

Pelgrom establece los componentes físicos mutuamente independientes que influyen en el factor de corriente β , pero tal consideración no fue tomada para la deducción de la varianza de V_t . El voltaje umbral es dependiente de V_{fb} , W , L , t_{ox} y N_{SUB} , por su parte, el factor de corriente depende de W , L , t_{ox} y μ_n , en consecuencia existiría una correlación entre la varianza de β y V_t principalmente por medio de W , L y t_{ox} . La ecuación (2.2) al despreciar los términos comunes entre las dos varianzas, esta sobrestimando el *mismatch* de la corriente,

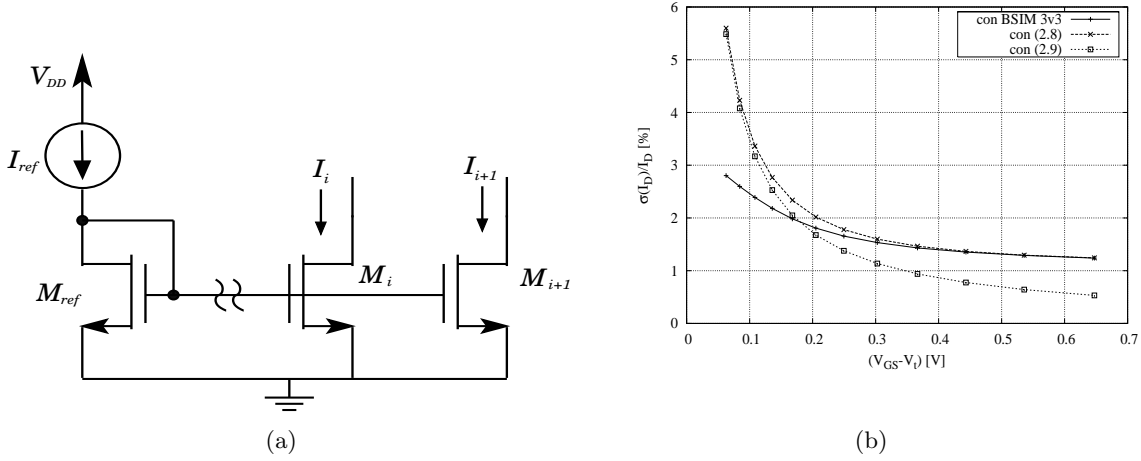


Figura 2.3: (a) Circuito de medida del *mismatch* . (b) *Mismatch* de I_D vs. $V_{GS} - V_t$

en conjunto con su derivación del modelo cuadrático, lo cual la convierten en un estimativo poco confiable. Por su parte, Drennan teniendo en cuenta esta sobre-estimación, redefine (2.2) como:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{\sigma_\beta^2 + 4\sigma_{V_{to}}^2}{(V_{GS} - V_t)^2} \quad (2.9)$$

La anterior expresión presenta una inconsistencia de unidades, para que sea coherente a este respecto σ_β^2 debería estar en $[V^2]$. En la figura 2.3(b) son contrastadas las expresiones (2.2) y (2.9) en conjunto con la estimación de la varianza de la corriente bajo el modelo BSIM3v3. Para este propósito fueron variadas UO , TOX y $VTH0$ fundamentados en la información extraída de los modelos *worst-case*¹. Luego el resultado fue contrastado con el calculo de las expresiones teóricas de primer nivel. Es interesante destacar la sobre-estimación de las dos expresiones teóricas para valores V_{OV} bajos debido a las diferencias con BSIM3v3, pero (2.9) comienza a ser un buen estimativo de la varianza de la corriente para tensiones V_{OV} mayores a 0,3[V]. Por su parte, la expresión (2.2) de Pelgrom es un mejor estimativo de la varianza de la corriente del modelo BSIM, que la ecuación de Drennan. La derivación de σ_{I_d} siguiendo el esquema de propagación de la varianza establecido por Drennan produce los mismos resultados que (2.2), por lo que la expresión (2.9) carece de contexto.

En el caso de Drennan se distinguen correctamente los parámetros de proceso mutuamente independientes, pero es el modelo de simulación quien determina su correlación por medio de un análisis de sensibilidad, siendo usado para esto comúnmente el modelo BSIM3v3. A pesar de que los resultados de las sensibilidades sean medianamente dependientes de la tolerancia al error y el paso de simulación, trabajar directamente sobre el modelo comporta-

¹Entre los modelos *worst-case* están: la media típica, el peor caso en velocidad, el peor caso en potencia y a nivel digital el peor caso del cero y el uno

mental trae beneficios como la reducción del error de estimación. Otro aspecto importante de este modelo consiste en el carácter más físico que adquiere la dependencia de los parámetros a la contribución total del *mismatch* gracias a la teoría de la propagación de la varianza, la cual indaga directamente la dependencia en la ecuación comportamental. Para Pelgrom esta se establecía a ser un termino constante dividido por el área, sin distinción alguna de la ecuación. Por ejemplo el factor de corriente β tiene una dependencia inversa con respecto al ancho W , pero esta característica no es evidenciada en el calculo de σ_β , donde se observa una dependencia proporcional.

En el trabajo de Galup-Montoro, las ecuaciones son fundamentadas en un modelo totalmente físico y continuo: el modelo *ACM*. Las derivaciones a pesar de que no incluyen los efectos en la movilidad, el oxido de puerta, o la pendiente de sub-umbral, entre otros, son coherentes y detallistas. Una amplia deducción matemática es llevada a cabo para tomar en cuenta tan sólo un factor de *mismatch*: el efecto de fluctuación de los portadores locales, lo cual mediría lo complejo que podrían resultar futuras inclusiones. Se evidencia un problema que surge de la suposición del parámetro N_{oi} como una cantidad constante. Al parámetro N_{oi} están ligados fenómenos no tenidos en cuenta en la derivación del modelo los cuales están relacionados con el efecto cuerpo y el efecto de la tensión de drenosurtidor sobre el perfil de dopaje o *DIBL* (*Drain induced barrier lowering*), entre otros. Un mejor modelado de N_{oi} podría llevar a expresiones más complejas pero que tornarían el modelo de mayor significado físico.

2.2.2. Continuidad

La deducción del modelo Pelgrom fue hecha exclusivamente para transistores con canal invertido, por lo que sub-umbral esta excluido. Debido a esto se intentará ajustar la expresión (2.2) de la manera más general, con el propósito de lograr incluir sub-umbral en esta. Otra forma de expresar la varianza de la corriente despreciando los efectos de β sería:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{A_{V_t}^2}{WL} \left(\frac{g_m}{I_D} \right)^2 \quad (2.10)$$

En principio, (2.10) definiría la varianza de la corriente de un transistor MOS en un amplio rango de polarización que incluye inversión débil, por lo tanto Pelgrom sería continuo. ¿Pero (2.10) es una expresión adecuada para σ_{I_D} ? La respuesta es si. Utilizando el concepto de propagación de la varianza usado en [23] se tiene:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \frac{1}{I_D^2} \frac{\partial I_D^2}{\partial V_t} \sigma_{V_t}^2 \quad (2.11)$$

Pero la derivada de la corriente con respecto al V_t tanto en sub-umbral como en canal invertido

es igual a $-g_m$. Por lo tanto (2.11) se convierte en:

$$\frac{\sigma_{I_D}^2}{I_D^2} = \left(\frac{g_m}{I_D}\right)^2 \sigma_{V_t}^2 \quad (2.12)$$

la cual corresponde a la misma expresión en (2.10). De esta manera se establece una continuidad del modelo Pelgrom, válida únicamente para un σ_{I_D} dependiente tan sólo del *mismatch* en V_t . Cabe aclarar que esta continuidad esta limitada a la derivación de la varianza de la corriente por medio de la ecuación cuadrática básica. Para otras ecuaciones y parámetros, es posible que el modelo Pelgrom no sea continuo.

El modelo de Drennan es tan continuo como el modelo de simulación usado para calcular las sensibilidades. Para el caso de BSIM3v3, el modelo esta definido por partes; primeramente las ecuaciones garantizan una continuidad en el rango de la polarización para un dispositivo con dimensiones estáticas, que luego es verificada calculando el error relativo entre dos iteraciones. Igualmente el cálculo de las sensibilidades involucra este principio. Pero la continuidad de los parámetros de este modelo, como la pendiente de sub-umbral n y el coeficiente DIBL en sub-umbral, estarían ampliamente limitada, dado a que estos tan solo aparecen o son relevantes para ciertas regiones de operación. Pero la continuidad de este modelo no es tan sólo dependiente del modelo de simulación, sino del conjunto de datos recolectado, pues el cálculo de cada una de las varianzas de los parámetros debe converger a un valor constante para distintas condiciones de dimensionamiento y polarización. En conclusión se podría afirmar que el modelo de Drennan es medianamente continuo.

El modelo de Galup-Montoro es continuo, pues apoya toda su deducción sobre la base de un modelo continuo como el modelo *ACM*, del cual propone una expresión para el *mismatch* válida para un amplio intervalo de polarizaciones. Esta es una gran ventaja en la medida en que el *mismatch* pasa a ser una expresión determinista como lo es la corriente. Posiblemente la inclusión de fenómenos despreciables en la deducción del modelo sigan garantizando su continuidad.

2.2.3. Medibilidad

Una cualidad importante del modelo de Pelgrom es su procedimiento original de medida, ya que son calculados tanto los parámetros eléctricos (V_t, β, K, θ) como sus varianzas en un proceso donde se involucran varios lotes durante varios años. Por lo tanto la credibilidad de los datos es alta. El calculo de los parámetros eléctricos fue realizado por medio del ajuste matemático, y partiendo de la expresión cuadrática de la corriente, válida solo para canales largos. En el contexto de Pelgrom, el error inherente a la aplicación de la ecuación cuadrática es bajo, dado que se trata de un proceso de $1,6\mu m$, pero para escalas menores se hace imprescindible la utilización de las ecuaciones comportamentales que incluyen un conjunto más amplio de parámetros. De esta manera, los procesos de ajuste de parámetros se van haciendo

más tediosos, por lo cual comúnmente se ha optado por el análisis y la medición de un único parámetro eléctrico como la corriente. Consecuentemente, el valor de cada una de las constantes del modelo Pelgrom para cada parámetro es fácilmente calculable dada la simplicidad de la ecuación (2.1). La adopción de la corriente como parámetro esencial en la medición del *mismatch* es obvia, ya que el transistor MOS es un transconductor, además la corriente es fácilmente medible por medio de la estructura de prueba de la figura 2.3(a).

El proceso de medida en el modelo Drennan parte de una correcta inclusión de parámetros de proceso altamente dependientes, de manera que estos no se vean correlacionados entre si. La etapa más criticable consiste en la selección de las dimensiones y condiciones de polarización que resultan relevantes para cada parámetro, y que permitirían calcular la varianza relacionada a ese parámetro. Esto representa un inconveniente alto para aquellos conjuntos de valores donde dos o más parámetros son significativos al *mismatch* total, pues se requerirían al menos un número de muestras iguales al número de parámetros relevantes, en una determinada condición de polarización. El diseñador está lejano de percibir estos casos excepcionales, por lo que el cálculo de la regresión matemática determinaría erróneamente las varianzas, en el mejor caso donde no se encuentre una singularidad en la matriz (2.3). Por lo tanto se requiere un alto grado de conocimiento del proceso para efectuar una correcta medición del modelo de Drennan.

La caracterización de una tecnología bajo el modelo de *mismatch* de Galup-Montoro implica la estimación de dos valores característicos: N_{oi} y B_{ISQ} . Por lo tanto el proceso de caracterización basado en una regresión matemática es sencillo y preciso. Galup-Montoro en [8] calcula N_{oi} de la curva de la varianza de la corriente en sub-umbral, luego este valor es usado junto con la ecuación (2.8) para calcular el parámetro B_{ISQ} de las medidas hechas en inversión fuerte tanto para saturación como para triodo. Pero el perfil de dopaje, que determina N_{oi} es altamente dependiente de la polarización tanto de puerta como de dren-surtidor [21]. Lo anterior indica que el procedimiento usado por Galup-Montoro para el cálculo de B_{ISQ} es incorrecto. Las medidas en [8] ratifican las suposiciones hechas acerca del carácter físico de N_{oi} , donde este exhibe un comportamiento impredecible. Además, son usadas medidas de transistores con efecto cuerpo, las cuales también modifican el perfil de dopaje.

Cabe destacar el método de medición llevado a cabo, el cual usa la estructura de prueba de la figura 2.3(a). Este proceso incluye la toma de medidas por medio de diferentes unidades de prueba, para luego promediarlas y reducir así el error. Ni Drennan ni Galup-Montoro tienen en cuenta la medición de las variaciones globales.

2.2.4. Precisión

La cualidad de precisión de un modelo depende altamente del proceso de fabricación bajo el cual este fue caracterizado. Por lo tanto, intentar comparar modelos medidos bajo distintas condiciones es algo contraproducente. Para analizar la precisión del modelo Pelgrom en escalas de integración actuales, se comparan las expresiones (2.7) y (2.12) que describen la varianza de

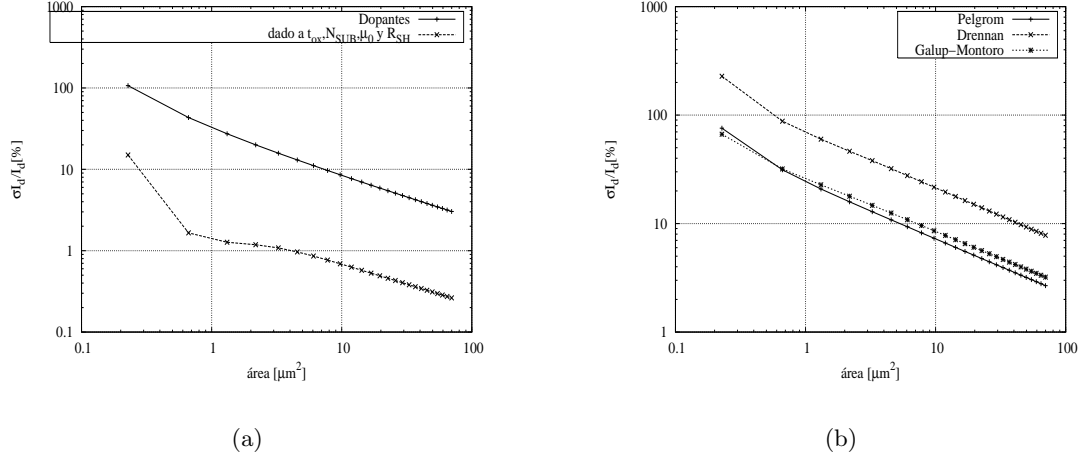


Figura 2.4: (a) Efecto de N_{oi} y de B_{ISQ} en el *mismatch*. (b) Comparación de los tres modelos.

la corriente en sub-umbral del modelo de Pelgrom y Galup-Montoro respectivamente. Ambas tienen dos cosas en común: incluyen básicamente el efecto de dopantes locales en el *mismatch* que influyen σ_{V_t} y poseen la misma estructura matemática, ya que g_m/I_D en sub-umbral es un termino constate igual a $1/nU_T$. Sentencia que puede ser reconfirmada de la figura 2.4(b), donde se comparan en simulación los tres modelos discutidos. En este sentido, la precisión de ambos modelos es la misma para el dominio de polarización sub-umbral. Los resultados mostrados en [8] muestran una adecuada aproximación en la estimación del *mismatch* por medio del modelo Galup-Montoro para transistores polarizados en inversión débil y de tamaños medianos. Estas mismas tasas de precisión no son obtenidas para transistores con dimensiones pequeñas y en aquellos cuyo cuerpo esta polarizado , presentándose así valores impredecibles para el parámetro N_{oi} . Realmente, las características eléctricas de los dispositivos de canal corto son muy dependientes a los efectos de borde en el perfil de dopaje. Además, para las mínimas dimensiones, las regiones de dopaje del dren y la fuente están muy cercanas, tornando altamente sensitiva la capa de deflexión debajo del canal. Para analizar la pertinencia de la inclusión de B_{ISQ} , se simula el efecto que tienen las variaciones despreciadas en 2.5, que para este caso corresponden a t_{ox} , N_{SUB} , μ_o y r_{sh} basadas en los modelos *worst-case*. En la figura 2.4(a) se observa una tendencia altamente lineal para áreas algo mayores a $1\mu\text{m}^2$. Por lo tanto es posible que la inclusión de B_{ISQ} modele acertadamente estos efectos para dispositivos medianos y grandes. En general, tanto para Pelgrom como para Galup-Montoro la precisión disminuye conforme el dimensionamiento del transistor se reduce. El modelo Drennan al haber sido ya usado en la industria, como el modelo Pelgrom, indicaría lo adecuado y por ende preciso que resultaría ser para la práctica. Pero la evaluación del modelo de Drennan involucra inmediatamente dos factores analizados anteriormente: el modelo de simulación usado y el conjunto de pruebas de caracterización llevado a cabo. Por lo tanto, la

precisión del modelo de Drennan es altamente dependiente del contexto bajo el cual este halla sido derivado.

2.2.5. Utilidad

La sencillez y adecuada precisión de su tiempo convirtieron a Pelgrom en un modelo estándar en la industria e indudablemente en el modelo más útil hasta el momento de la microelectrónica. La utilidad del modelo Pelgrom se fundamenta en una premisa esencial: el aumento de área del dispositivo disminuye proporcionalmente su variabilidad. A pesar de que esta consigna haya sido ampliamente usada en el diseño, no resuelve preguntas que buscan controlar más a nivel físico y de circuito el fenómeno del *mismatch*. Además el diseño ha migrado a distintas regiones de polarización y los procesos de fabricación han incluido nuevas estrategias de implantes de impurezas que no pueden ser modelados por las mismas ecuaciones que derivaron este modelo. Otro aporte esencial del modelo Pelgrom consiste en la inclusión de las variaciones globales en los procesos que involucran el diseño del *layout*, y que fundamentaron técnicas como el centroide común. La utilidad del modelo Pelgrom es indudable: las varianzas de los parámetros en relación con el dimensionamiento del dispositivo pueden ser utilizadas para estimar y diezmar el efecto del *mismatch* como también para proponer nuevos modelos. Ejemplos de estos modelos se encuentran en.

La utilidad del modelo Drennan se enfoca en la etapa de verificación del diseño, mas no en el diseño mismo, pues las expresiones en (2.3) solo pueden ser evaluadas por medio de simulación, partiendo de un dimensionamiento previo del circuito. Para efectuar estimaciones manuales de *mismatch*, se debe recurrir a modelos como el de Pelgrom. Además bajo el modelo de Drennan no pueden efectuarse fácilmente compromisos como mínima área y variabilidad, dado a que a cada parámetro esta ligada una sensibilidad dependiente del dimensionamiento. La evaluación de este tipo de compromisos implicaría la caracterización de la sensibilidad para un amplio rango de dimensiones, en conjunto con técnicas computacionales. El modelo de Drennan no propone una estrategia distinta que pueda ser llevada a nivel de *layout*, por lo que el circuito es transparente a los efectos de la disposición de los elementos en el *layout*.

El modelo de Galup-Montoro es útil en la medida en que el modelo *ACM* se asocie a las herramientas de simulación y diseño. El modelo *ACM* ya ha sido implementado en simulación [24] y sus parámetros han sido extraídos del modelo BSIM3v2 por dos métodos: confrontación de ecuaciones y simulación [25]. A pesar de eso, no existen herramientas comerciales que tengan a su disposición la implementación del *ACM*, por lo que contrastar resultados en el proceso de diseño bajo dos modelos diferentes se convierte en un proceso tedioso. La principal utilidad de Galup-Montoro es la obtención de un modelo físico, continuo y relativamente sencillo que describe acertadamente el fenómeno del *mismatch*. Por último, es bueno establecer que en este modelo no son consideradas las variaciones globales, por lo tanto resulta poco útil, para la simulación de *layouts*.

2.2.6. Escalabilidad

En el modelo Pelgrom se ha presentado una disminución en el valor de A_{V_T} con la reducción del tamaño de los transistores MOS [26]. Por lo tanto un dispositivo que ocupe un área constante mejorara su precisión en futuros procesos sub-micrón. Pero esta tendencia de mejoramiento se ve contrastada con el aumento de la densidad de integración, ligada a una menor área por dispositivo lo que conduce a un aumento en la variabilidad. La disminución del tamaño característico es cuadrática comparada con la disminución lineal de A_{V_T} [26]. Además las nuevas tendencias en circuitos promueven un bajo $(V_{GS} - V_t)$, por lo tanto una proporción alta de (g_m/I_D) , tal que los errores en V_t debido a los fenómenos inherentes al canal comienzan a ser considerables. Consecuentemente, la variabilidad de muchos parámetros se torna más dependiente a la longitud de canal que al ancho del transistor, razón por la cual el exponente de $1/L$ debe ser aumentado con respecto al exponente de $1/W$.

La insuficiencia del modelado del *mismatch* en Pelgrom es notoria en tecnologías de escalas sub-micrón. Para estimar el *mismatch* en la corriente de procesos futuros, una nueva derivación del modelo de Pelgrom basada en ecuaciones de canal corto debe ser llevada a cabo. A pesar de esto, la esencia de la ecuación (2.1) se seguirá manteniendo para muchos de los parámetros durante los años futuros. En general, se puede decir que la escalabilidad en el modelo Pelgrom es limitada.

En Drennan se siguen manteniendo la dependencia del dimensionamiento con la varianza de los parámetros dispuesta por Pelgrom, pero se le agrega un nuevo componente: las sensibilidades. En análisis entre las sensibilidades y el escalamiento realizado en la sección 1.3, muestra un aumento de estas conforme se reduce el dimensionamiento. Esta deducción es consecuente con el recurrente concepto del *mismatch*, por lo que el escalamiento produce resultados coherentes en el modelo Drennan. Se podría afirmar de esta manera que el modelo de Drennan es altamente escalable.

Finalmente se analiza la escalabilidad del modelo de Galup-Montoro. Inicialmente, la ecuación (2.7) evidencia una tendencia similar a la establecida por Pelgrom, esto puede ser corroborado de los resultados comparativos de los tres modelos presentados en la figura 2.4(b). El efecto de escalamiento sobre las corrientes i_f e i_r es mínimo, dado a que son parámetros normalizados. La gran diferencia se establece en N_{oi} , el cual es altamente sensible al escalamiento como confirmo el análisis de medibilidad de este modelo. Para adecuarlo a futuras generaciones de proceso, el modelo se torna fácilmente escalable si e incluye la dependencia del voltaje V_{DS} y V_{SB} sobre el parámetro N_{oi} . Además si se reconfirma la pertinencia de la inclusión del factor B_{ISQ} para modelar los efectos de otros parámetros.

MODELOS	PELGROM	DRENNAN	GALUP-MONTORO
Significado físico	Sobreestimación por correlación de ΔV_t y $\Delta\beta$. Derivación de varianzas del modelo cuadrático.	Distinción de parámetros mutuamente dependientes. Dependencia del carácter físico del modelo de simulación. Teoría de la propagación de la varianza.	Fundamentado en un modelo totalmente físico. No modela dependencia del perfil de dopaje con las tensiones V_{DS} y V_{BS} .
Continuidad	Continuidad limitada a ciertos parámetros y modelos.	Altamente dependiente del modelo de simulación.	Continuo en el rango de polarización
Medibilidad	Altamente medible dado la sencillez de ajuste de ecuaciones. Caracteriza también los parámetros determinísticos.	Resultados no repetibles, dependientes de dimensiones y tensiones tomadas. Para casos excepcionales la matriz (2.3) podría ser singular	Medición sencilla y precisa. Requiere extraer valores de N_{oi} a distintas condiciones
Precisión	Disminuye proporcional al área del dispositivo	Depende de las mediciones y el modelo de simulación usado.	Para sub-umbral es igual de acertado que Pelgrom. Comportamiento impredecible del parámetro N_{oi} para ciertas condiciones.
Utilidad	Variaciones globales para el análisis en <i>layout</i> Usado para estimar y reducir el <i>mismatch</i> , como para proponer nuevos modelos.	Útil para la verificación del diseño, más no para el diseño mismo. No permite compromisos relacionados al <i>mismatch</i>	Útil si ACM es asociado al diseño. Continuo sencillo y de carácter físico
Escalabilidad	Insuficiente en tecnologías sub-micrón. Describe una tendencia más no un comportamiento.	Determinado por la sensibilidad de los parámetros. Coherente con la tendencia de escalamiento.	Escalable en la medida en que N_{oi} incluya fenómenos despreciables como el <i>DIBL</i> y el efecto cuerpo.

Tabla 2.2: Resumen de los aspectos más relevantes para los tres modelos tratados.

2.3. Conclusión y selección.

En la tabla 2.2 se encuentran condensados de manera comparativa los aspectos primordiales para cada uno de los tres modelos abordados. El modelo Pelgrom al establecer de manera sencilla una tendencia más no un estimativo de variabilidad del circuito, resulta adecuado como punto inicial para el dimensionamiento del circuito. De esta manera, la ecuación (2.1) fácilmente acotaría el espacio de diseño en el contexto de la determinación de las mínimas áreas permisibles, de manera tal que etapas posteriores de dimensionamiento sean garantizadas en el marco del diseño para variabilidad. Por su parte, los bajos niveles de precisión inherentes al modelo Pelgrom en tecnologías sub-micrón, hacen de este un estimativo inadecuado para las fases que involucran la validación del diseño. La sobreestimación del *mismatch* en las etapas de validación del diseño origina un re-dimensionamiento que tenderá a exceder el área a niveles poco permisibles. Paralelamente, una sub-estimación del *mismatch* conduce a resultados de rendimiento sobreestimados que enaltecen la robustez del circuito en conjunto con las herramientas y metodologías asociadas al diseño.

Entonces, en búsqueda de un estimativo adecuado para la validación última del diseño, se encuentran ante los dos modelos restantes estudiados en el presente capítulo, los cuales deberán ser trasladados al contexto de nuestro trabajo. Por un lado, el modelo de *mismatch* fundamentado en el modelo *ACM* resulta ser un estimativo relativamente sencillo y considerablemente adecuado para la fase de verificación y re-dimensionamiento del circuito, pero presenta dos limitantes importantes. Primero, no existen las herramientas comerciales que permitan una aplicabilidad global del modelo *ACM* en el flujo de diseño y en lo cual se requiere la implementación de por lo menos dos modelos para la fase de diseño. Y segundo, la expresión que propone la varianza de la corriente se reduce a la establecida por Pelgrom, representando una transparencia en la aplicabilidad de estos dos modelos.. Al ser *ACM* un modelo relativamente reciente, requiere de desarrollos posteriores que extiendan la variabilidad exhibida por los transistores polarizados en sub-umbral.

Por otro lado Drennan propone un modelo lo suficientemente complejo como para no ser usado en la etapa de diseño, dado a que su expresión involucra de manera distinta las sensibilidades de cada uno de los parámetros físicos relevantes. El tedioso cálculo ligado a la determinación de las dimensiones en la expresión (2.3), hacen de este un modelo inadecuado para establecer compromisos entre variables de diseño ligadas al dimensionamiento con intervalos de aceptabilidad de rendimiento y desempeño establecidos. Pero la expresión en (2.3) se torna fácilmente calculable cuando las dimensiones y polarizaciones han sido previamente definidas, es decir cuando un diseño final para verificación ha sido establecido.

De esta manera y considerando que la metodología a abordar requiere de dos fases importantes relacionadas con la estimación del *mismatch*, la selección de los modelos a usar requiere el análisis bajo dos marcos de diseño distintos. Primeramente y dadas unas exigencias iniciales de circuito a nivel de variabilidad, el área debe ser restringida a valores mínimos permisibles.

Este proceso al requerir tan sólo de una tendencia de variabilidad y excluyéndose del análisis de un circuito ya dimensionado, puede ser ampliamente abordado por medio de un modelo como Pelgrom, razones suficientes que lo convierten en nuestro estimador inicial de la variabilidad. Consecuentemente, el diseño final al involucrar altas exigencias de precisión en la estimación de su desempeño y rendimiento, requiere de un modelo bastante acorde con los resultados de silicio. Al haber sido descartados tanto Pelgrom como *ACM* para este propósito, convierten a Drennan en el modelo usado para la validación del diseño. Debe ser clara la idea de que la elección de un modelo de *mismatch* esta ampliamente ligada al contexto bajo el cual se requiera determinar la variabilidad. Por ejemplo para transistores en saturación, las mismas premisas que guiaron nuestra elección no pueden ser aplicadas, pues por un lado la expresión (2.5) que propone el modelo *ACM* podría ser usada en la etapa de verificación ya que describe adecuadamente la varianza de la corriente en saturación. De otra manera, la variabilidad de inversión débil al ser más alta que la de saturación, requiere de un proceso de estimación más riguroso. En el ámbito de la metodología, la elección de los modelos del *mismatch* debe adecuarse a las intenciones y medios presentes de cada una de las etapas del flujo de diseño. De esta manera son establecidos los medios para el análisis de variabilidad de la fuente de referencia, dejando de esta manera libre el camino al planteamiento de la metodología.

Capítulo 3

Metodología de diseño *DFM*

Probablemente una de las características más importantes ligadas a los circuitos a nivel de topología consiste en su flexibilidad al diseño, la cual posibilita a los circuitos a ser ampliamente aplicables a múltiples propósitos. En este sentido, a una topología le serán ligadas un conjunto amplio de posibilidades de desempeño y varias de funcionamiento, que en el marco del diseñador serán percibidas como un conjunto de restricciones n -dimensionales comprometidas entre sí. Sin embargo, a pesar de contar con un conjunto extendido de alternativas de dimensionamiento, es común ver como cada topología comparte una misma estructura metodológica, la cual puede ser secuencialmente representada como un algoritmo de un proceso de diseño automatizado. Aún cuando las herramientas de automatización *CAD* del diseño estén muy fortalecidas en el ámbito del dimensionamiento del circuito, la injerencia humana caracterizada por su inapelable intuición es bastante requerida y necesaria. Pero el problema real de las herramientas *CAD* no se concentra allí, sino en la fase que involucra el diseño de circuito bajo procesos de fabricación altamente variables en sus parámetros, ante los cuales dos o más prototipos son requeridos para ajustar la herramienta al propósito del diseño.

En la búsqueda de una metodología de diseño altamente extrapolable a distintos circuitos y procesos tecnológicos se ha visto abocada la industria en su consigna de reducción de costos y tiempo de diseño. La flexibilidad de una metodología de diseño en el ámbito de los circuitos parte de un correcto modelado de los bloques mínimos funcionales como los transistores, de manera que el comportamiento a nivel de sistema de cualquier dispositivo pueda ser correctamente estimado, controlado y por lo tanto optimizado. Para este propósito, métodos de optimización como la programación geométrica han sido ampliamente aplicados al diseño analógico. Por otro lado, la carencia de metodologías de diseño aplicables a distintos procesos tecnológicos ha sido uno de los cuellos de botella del diseño analógico, pues además de introducir problemas relacionados con fenómenos debidos al escalamiento de los dispositivos, debe propender por encontrar diseños óptimos ante variables disímiles para cada tecnología e inherentemente aleatorias. De esta manera, el conjunto herramientas *CAD* y metodología de diseño debe ser establecido bajo la lógica de adaptabilidad, que garantice una baja dependen-

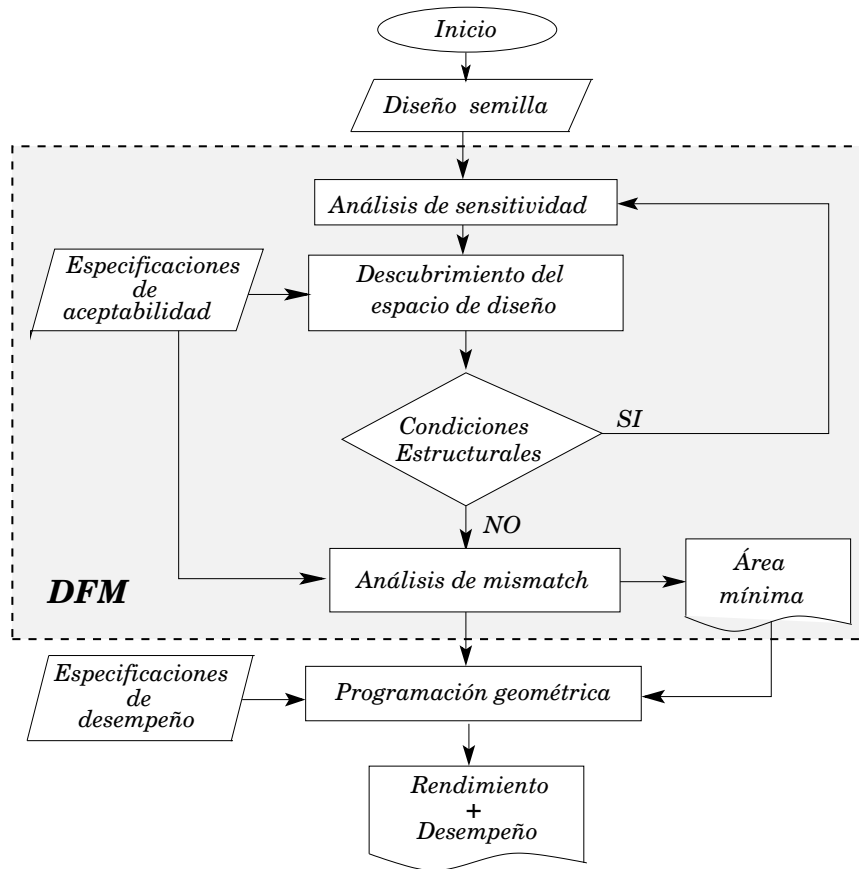


Figura 3.1: Flujo de diseño *DFM* adoptado para este trabajo.

cia con el contexto que involucra el diseño y al cual están envueltos el proceso de fabricación, la topología y la aplicación del mismo.

La metodología presentada en este trabajo y que corresponde a la esquematizada en la figura 3, tiene como principal objetivo hacer extensible las herramientas *CAD* al diseño de cualquier circuito analógico. Literalmente hablando, esta flujo de diseño debe ser capaz de aplicarse bajo distintos entornos de diseño que incluyen los procesos de escalas nanométricas. Como punto de partida es tomado un diseño semilla que le permite a la herramienta establecer un marco funcional del comportamiento del circuito deseado. El primer paso consiste en una aproximación del espacio de diseño basada en un análisis de sensibilidades del diseño semilla. Esta aproximación tiene como objetivo guiar el diseño a valores de variables de diseño factibles a nivel funcional, de manera que procesos posteriores como programación geométrica estén acotados en el marco *DFM*. Establecida la funcionalidad global del circuito, se procede a establecer un análisis de *mismatch*, el cual basado en unas condiciones de aceptabilidad de rendimiento iniciales determinadas por el diseñador, delimita los valores de las variables de dimensionamiento estableciendo restricciones de área mínima. La estimación del área mínima requerida es efectuada por medio de un análisis de la variabilidad de cada elemento del circuito

en cada una de las funciones de desempeño establecidas. El peor caso del área estimada por este análisis, permite garantizar la inclusión de los intervalos de desempeño al conjunto 6σ de dispositivos fabricados. Por último, una etapa de optimización de desempeño sobre el conjunto de diseño anteriormente acotado para *DFM* es llevada a cabo. Para este propósito se usa la programación geométrica, pues representa un tipo de optimización convexa que obtiene óptimos globales sobre conjuntos limitados y definidos. Es necesario reconocer como la metodología *DFM* de la figura 3 puede ser aplicada sin distinción a cualquier circuito analógico, garantizando una flexibilidad y aplicabilidad en el diseño.

3.1. Aproximación del espacio de diseño.

La alta no-linealidad y no-convexidad exhibida comúnmente por el espacio de diseño lo hacen inadecuado para los propósitos de centralización y optimización requeridos en este trabajo. La no linealidad limita el ajuste del espacio de diseño a algún tipo de superficie conocida, conjuntamente, la no-convexidad lo excluye de ser optimizado y en su defecto centralizado. Además, la consecución de una aproximación fiel del espacio de diseño es generalmente innecesaria y altamente prohibitiva a nivel de tiempo de computo.

Para estimar el espacio de diseño es común aproximarlo a una figura convexa de centro conocido o fácilmente calculable, con el fin de guiar el diseño a una región menos vulnerable a fallas funcionales aproximadas por el contorno de la figura convexa. En uno de los trabajos pioneros publicados [27], el espacio de diseño es aproximado por medio de la exploración unidimensional de cada una de las n variables de diseño, partiendo del conocimiento de $m \geq n+1$ ¹ considerados factibles y por lo tanto contenidos en la región de diseño. Una forma de encontrar $n+1$ puntos requiere primero encontrar un punto contenido en la región de diseño determinado por la intuición del diseñador o por medio de un programa de optimización, para luego efectuar una búsqueda uni-dimensional a lo largo de las coordenadas positivas y negativas. De esta manera, el centro de la región de diseño es estimado por medio del centro de la hiper-esfera de mayor volumen contenida en el poliedro determinado por la exploración de los m planos. La aproximación es iterativamente hecha de manera que se garantice que el rendimiento $Y^{(k+1)}$ sea siempre mayor al rendimiento Y^k . De igual manera, aproximaciones posteriores [10,28,29], han empleado conceptos similares bajo distintas figuras geométricas entre las cuales incluyen poliedros, elipsoides y *polytopes*.

Por simplicidad y garantizando una posterior implementación en el formato PG, el espacio de diseño es aquí aproximado por $n+m$ superficies n -dimensionales, n de los cuales son planos normales a cada una de las variables de diseño establecidas inicialmente. De la misma manera, las m superficies restantes corresponden a las mínimas áreas permitidas para cada uno de los dispositivos obtenidas por medio de las condiciones iniciales de aceptabilidad a nivel de

¹Se garantiza que una región n -dimensional esta acotada cuando por lo menos $n+1$ planos linealmente independientes la describen.

rendimiento. Para los circuitos cuyas variables de diseño son siempre de dimensionamiento, la relación de n y m correspondería a $n = 2m$. Entonces y considerando que las dimensiones de los transistores son W y L , se define un vector n -dimensional para cada una de las variables de diseño en un arreglo alternado de W y L de manera que:

$$\mathbf{x} = (W_1, L_1, \dots, W_m, L_m) \quad (3.1)$$

donde W_i y L_i son el ancho y el largo del i -ésimo dispositivo respectivamente. Por su parte, se construye cada una de las superficies relacionadas al área mínima aceptada de la siguiente manera:

$$\mathbf{A} = (x_1 \cdot x_2, \dots, x_{2i} \cdot x_{2i+1}, \dots, x_{n-1} \cdot x_n) \quad (3.2)$$

donde x_i es el i -ésimo componente del vector de variables de diseño determinado por (3.1). Determinados los contornos del espacio de diseño, el espacio de diseño es aproximado a la región acotada definida por:

$$\mathcal{D} = \{\mathbf{x} \in \mathfrak{R}^n \mid \mathbf{x} \leq \mathbf{x}_U, \wedge, \mathbf{A}_{min} \leq \mathbf{A}\} \quad (3.3)$$

donde $\mathbf{x}_U \in \mathfrak{R}^n$ y $\mathbf{A}_{min} \in \mathfrak{R}^m$. Los valores de los límites superiores de las variables de diseño además de estar ligados a condiciones de polarización y áreas permisibles, también se encuentran ampliamente influenciados por el comportamiento del circuito. Por su parte, los límites inferiores asegurarían de entrada un cumplimiento de las restricciones de aceptabilidad a nivel de rendimiento.

A diferencia de otras propuestas, para este trabajo resulta apropiado la aproximación del espacio de diseño como la intersección de las regiones de factibilidad a nivel de aceptabilidad de rendimiento y la región de funcionalidad, excluyéndose de garantizar un directo cumplimiento de una aceptabilidad de desempeño. Este concepto debe ser claro, ya que es común que la aceptabilidad de desempeño sea definida como una región continua y no como un conjunto de puntos dispersos, haciendo parecer el problema algo sencillo. Además es evidente que el interés primordial hasta ahora no consiste en garantizar un buen desempeño, puesto que esta responsabilidad estará asociada a la fase de optimización vía programación geométrica. Entonces, al establecer de esta manera el espacio de diseño, etapas posteriores de esta metodología como la optimización, estarán más acotadas en el marco del diseño de manufactura *DFM* y no bajo la simple determinación de límites máximos y mínimos condicionados en una lógica de valores permisibles.

Para esta construcción de la aproximación del espacio de diseño se hace imprescindible tomar como punto de partida un diseño funcional y con límites de aceptabilidad adecuados. Ya se había visto en la sección 1.3 como el conjunto de diseño se veía restringido con el aumento de la sensibilidad, por lo tanto bajo esta misma estructura puede ser definida de manera más general el desplazamiento de las variables de diseño para una de las funciones de desempeño

características involucradas. Es así como se tiene que:

$$\Delta \mathbf{x} = \left\| \frac{\mathbf{y}_i^{k+1} - \mathbf{y}_i^k}{\frac{\partial \mathbf{y}_i}{\partial \mathbf{x}}} \right\| \quad (3.4)$$

donde \mathbf{x} es el vector de variables de diseño anteriormente definido, \mathbf{y}_i es el vector de la i -ésima función de desempeño y Δx representa el desplazamiento desde el punto \mathbf{x}_{k+1} al punto \mathbf{x}_{k+2} . Para p funciones de desempeño, la aproximación del espacio de diseño puede ser obtenida iterativamente como la aproximación n -dimensional establecida en (3.4) del punto \mathbf{x}^{k+1} . Además, dada la manera distinta en como son priorizadas las funciones de desempeño para cada una de las aplicaciones y circuitos, una ponderación de relevancia debe ser agregada al desplazamiento del espacio de diseño. De esta manera se establece que la aproximación del espacio de diseño para la iteración $k + 1$ esta definida por:

$$\mathbf{x}^{k+2} = \sum_{i=1}^p c_i \left\| \frac{\mathbf{y}_i^{k+1} - \mathbf{y}_i^k}{\frac{\partial \mathbf{y}_i}{\partial \mathbf{x}}} \right\| + \mathbf{x}^{k+1} \quad (3.5)$$

donde c_i representa la ponderación de relevancia de la i -ésima función de desempeño. La evaluación de $\Delta \mathbf{x}$ es hecha de manera iterativa hasta que cualquiera de dos condiciones sea cumplida. La primera establece que por lo menos el 90 % de los transistores están por fuera de la región determinada en el diseño semilla, y la segunda que se exceda un área máxima para cualquier elemento del circuito. La elección de 90 como porcentaje de circuitos obedeció en nuestro caso a que en el la fuente de referencia a diseñar existen dos elementos flexibles a las regiones de polarización de corte y saturación, pero en realidad una exploración absoluta del espacio de diseño funcional debe ser efectuada hasta que el 100 % de los elementos estén por fuera de su región de polarización determinada en diseño. De otra manera, la máxima área esta basada en el dimensionamiento del diseño semilla y es delimitada a valores de área adecuados para fabricación. En principio y conociendo lo fácil que es averiar un diseño, la primera condición parecería ser suficiente para este propósito, pero los resultados mostraron una tendencia en el establecimiento de la polarización del circuito para cada una de las iteraciones. Por lo tanto, para evitar una sobre-estimación del espacio de diseño que conllevaría a valores de áreas no fabricables, la restricción del área debió ser agregada. Se permite concluir lo adecuado que resulta ser el método numérico, en conjunto con la polarización estable a la cual fue sometido el diseño semilla y que correspondió a la establecida por Juan Mateus [1].

En consonancia con la búsqueda del espacio de diseño, se hace evidente la adopción de un control por parte del algoritmo, de manera tal que las dimensiones de aquellos transistores que han salido de su región de polarización no sean desplazadas, pues en esta etapa ya no se garantizaría una funcionalidad del circuito. Para esto es dispuesto un vector con arreglos de unos y ceros que es determinado por medio de los análisis en simulación del punto de operación y luego multiplicado por la variable Δx .

Otro problema aparente provendría de los casos en los cuales la sensibilidad de la ecuación

(3.5) fuese cercana a cero, porque indudablemente desbordarían la estimación de la búsqueda del conjunto de diseño. Para sobrellevar este problema, en el algoritmo es dispuesta una mínima sensibilidad permisible definida de acuerdo al establecimiento de un máximo desplazamiento $\Delta \mathbf{x}$ para cada función de desempeño. El cálculo de la sensibilidad mínima es estimado como una función de la máxima área permisible y el número de iteraciones a la cual el algoritmo convergerá en el mejor caso.

De esta manera queda definida la construcción del espacio de diseño; no obstante, para que este espacio sea más adecuado a nuestras intenciones, una metodología del establecimiento de la mínima área para el control de variabilidad debe ser abordado.

3.2. Mínima área que garantiza aceptabilidad.

Con el fin de mejorar el espacio de diseño construido para propósitos de *DFM*, se debe establecer una dependencia de los requerimientos del área para cada elemento del dispositivo con la aceptabilidad de los parámetros de desempeño. Dado a que la inclusión de la variabilidad en el programa geométrico lo tornaría complejo, es necesario definirla como una superficie restrictiva del espacio de diseño. Además resultaría bastante contraproducente establecer dependencias entre funciones de desempeño y variables de diseño de un circuito previamente no dimensionado. Por lo tanto, esta etapa de nuevo ratifica la necesidad de un diseño semilla, con las mismas condiciones definidas en la sección anterior.

Basados en el planeamiento teórico y los resultados obtenidos en [17], se establece que para un circuito de n dispositivos con varianzas $\sigma_1, \dots, \sigma_n$ asumidas a ser independientes, el área óptima global que garantiza una desviación estándar total σ_{total} de un parámetro o una función de desempeño determinada es:

$$A_{total} = \left(\frac{\rho_1 + \rho_2 + \dots + \rho_n}{\sigma_{total}} \right)^2 \quad (3.6)$$

donde ρ_1, \dots, ρ_n en su manera más general son funciones que describen la contribución del i -ésimo elemento a la varianza global total σ_{total} . Sus valores pueden ser determinados de manera analítica o por medio de simulación. De manera particular, el área para el i -ésimo elemento que garantiza un área óptima global es:

$$A_i = \rho_i \frac{A_{total}}{\rho_1 + \rho_2 + \dots + \rho_n} \quad (3.7)$$

Es necesario destacar dos cosas: el modelo usado para las derivaciones de (3.6) y (3.7) corresponde al propuesto por Pelgrom y es inminente la sobrestimación del área debido a la suposición de no correlación entre los n dispositivos. El análisis del *mismatch* en el capítulo 2 concluyó que Pelgrom resulta ser el estimativo más adecuado para esta fase del diseño, justificando de nuevo la primera sentencia. Por otro lado, la sobre-estimación del área en circuitos

de alta correlación entre las variables de diseño de sus dispositivos producirá valores de área que podrían hasta exceder el espacio de diseño. Esta circunstancia podría ser solucionada fácilmente con una reducción de área estimada en un proceso de prueba y error, o en el caso más crítico con el cálculo de sensibilidades involucrando posibles correlaciones entre distintas variables de diseño.

La inclusión de las condiciones de aceptabilidad a nivel de rendimiento involucraría la extensión de las expresiones (3.6) y (3.7) a p dimensiones, en conjunto con la determinación de los valores ρ_1, \dots, ρ_n para cada una de las p funciones de desempeño. Para el caso de la fuente de referencia son tomados como variables de desempeño: el PSRR, la regulación de línea, el coeficiente térmico, la potencia, el voltaje de referencia y la temperatura media. Este último a pesar de no ser una función de desempeño conocida, es altamente variable y garantiza el buen funcionamiento de la fuente. El cálculo de cada uno de los valores ρ_1, \dots, ρ_n al ser fácilmente derivado, requiere de la determinación por medio de simulación. De esta manera, la expresión (2.3) establecida por el modelo de Drennan puede ser extrapolada a funciones de desempeño bajo el concepto de la propagación de la varianza. Por lo tanto se tendría que la varianza de cada una de las seis variables de desempeño establecidas anteriormente tendrá la forma:

$$\begin{bmatrix} \sigma_{y_1}^2 \\ \sigma_{y_2}^2 \\ \sigma_{y_3}^2 \\ \sigma_{y_4}^2 \\ \sigma_{y_5}^2 \\ \sigma_{y_6}^2 \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial y_1}{\partial i_1}\right)^2 & \left(\frac{\partial y_1}{\partial i_2}\right)^2 & \cdots & \left(\frac{\partial y_1}{\partial i_n}\right)^2 \\ \left(\frac{\partial y_2}{\partial i_1}\right)^2 & \left(\frac{\partial y_2}{\partial i_2}\right)^2 & \cdots & \left(\frac{\partial y_2}{\partial i_n}\right)^2 \\ \left(\frac{\partial y_3}{\partial i_1}\right)^2 & \left(\frac{\partial y_3}{\partial i_2}\right)^2 & \cdots & \left(\frac{\partial y_3}{\partial i_n}\right)^2 \\ \left(\frac{\partial y_4}{\partial i_1}\right)^2 & \left(\frac{\partial y_4}{\partial i_2}\right)^2 & \cdots & \left(\frac{\partial y_4}{\partial i_n}\right)^2 \\ \left(\frac{\partial y_5}{\partial i_1}\right)^2 & \left(\frac{\partial y_5}{\partial i_2}\right)^2 & \cdots & \left(\frac{\partial y_5}{\partial i_n}\right)^2 \\ \left(\frac{\partial y_6}{\partial i_1}\right)^2 & \left(\frac{\partial y_6}{\partial i_2}\right)^2 & \cdots & \left(\frac{\partial y_6}{\partial i_n}\right)^2 \end{bmatrix} \begin{bmatrix} \sigma_{I_{d1}}^2 \\ \sigma_{I_{d2}}^2 \\ \vdots \\ \sigma_{I_{dn}}^2 \end{bmatrix} \quad (3.8)$$

donde y_1, \dots, y_6 representa cada una de las funciones de desempeño nombradas anteriormente, σ_{y_i} define el rango de aceptabilidad de la función de desempeño e i_i con σ_i representa la corriente y la varianza del i -ésimo dispositivo respectivamente.

La aceptabilidad a nivel de rendimiento de cada función de desempeño estaría determinada por su valor nominal más la suma de una desviación estándar. Por ejemplo, para el caso del PSRR sería $60dB \pm 2dB$. Siguiendo con nuestro análisis, las varianzas de la corriente según Pelgrom pueden ser expresadas en sub-umbral como:

$$\sigma_{I_{d_i}}^2 = (g_{m_i})^2 \frac{A_{V_T}}{A_i} \quad (3.9)$$

donde A_{V_T} es la constante de proporcionalidad del proceso definida en (2.1) y g_{m_i} con A_i son la transconductancia y el área del i -ésimo dispositivo respectivamente. Como en la fuente de referencia se incluyen dispositivos con casos especiales como los resistores, su varianza

puede ser expresada como:

$$\sigma_{I_R} = \left(\frac{\partial I_R}{\partial R} \right)^2 \frac{A_R^2}{WL} \quad (3.10)$$

donde A_R es la constante de proporcionalidad de Pelgrom para el resistor R , I_R es la corriente que circula por este y WL es su área. Los valores de ρ para la j -ésima función de desempeño serían:

$$\rho_{ji}^2 = \begin{cases} (g_{m_i})^2 A_{VT} \left(\frac{\partial y_j}{\partial i_i} \right)^2 & \text{para el } i\text{-ésimo transistor} \\ A_R^2 = \left(\frac{\partial I_R}{\partial R} \right)^2 \left(\frac{\partial y_j}{\partial i_i} \right)^2 & \text{para el } i\text{-ésimo resistor} \end{cases}$$

Aplicando el concepto propuesto [17] y de conformidad con la formula (3.6), las p funciones de desempeño producen p valores posibles para el área total, cada uno de los cuales es dependiente de la restricción de aceptabilidad dispuesta y en consecuencia diferente a los demás. Entonces, para garantizar un cumplimiento de todo el conjunto de restricciones es necesario tomar el peor caso que corresponde al valor máximo del área obtenido:

$$A_{total} = \text{máx} \{(A_{total1}, \dots, A_{totalp})\} \quad (3.11)$$

donde $p = 6$ para nuestro caso. De la misma manera y disponiendo de la formula (3.7), las condiciones de aceptabilidad producirían p pares posibles para cada uno de los n dispositivos, que expresados en forma matricial serían:

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{12} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \dots & A_{pn} \end{bmatrix} \quad (3.12)$$

donde A_{ji} es el área del i -ésimo dispositivo para la j -ésima función de desempeño. Por lo tanto, esta evaluación establecería p probables áreas diferentes para cada dispositivo. Es claro que la suma de los peores casos de área individuales es siempre mayor o igual a el área establecida por la formula (3.11). Con el propósito de evitar una sobre-estimación del área bajo este procedimiento, es fácilmente deducible que el área del j -ésimo dispositivo deberá ser:

$$A_i = \text{máx} \{(A_{1i}, \dots, A_{pi})\} \frac{A_{total}}{\sum_{k=1}^n \text{máx} \{(A_{1i}, \dots, A_{pi})\}} \quad (3.13)$$

La expresión en (3.13) siempre garantizará que la suma de las áreas individuales del circuito sea igual al área determinada por la ecuación (3.11).

Cualitativamente cada una de las derivadas en (3.8) es calculada por medio de un análisis

de sensibilidad vía simulación. Para el caso de Hspice, el análisis de sensibilidad en función de las variables de desempeño no está disponible, por lo que debe ser implementado. Con este fin se dispone de una fuente de corriente de valor definido² en paralelo con un resistor drenador-fuente de alto valor en cada transistor. De esta manera es comparado individualmente cada dispositivo y su efecto sobre los p parámetros de desempeño. El resistor es colocado para propósitos de convergencia y reducción de error por estimación.

Propuesta la metodología para restringir el espacio de diseño a valores donde fuesen garantizadas las condiciones de aceptabilidad que definían la robustez del diseño ante la variabilidad, no es evidente aún la manera como el diseño obtendrá valores de parámetros de desempeño adecuados. Por parte del diseñador es claro que su intuición fue limitada conforme se restringió el conjunto de posibilidades de diseño, pero esto no representa una gran preocupación, pues herramientas de optimización como programación geométrica pueden obtener valores óptimos globales bajo conjuntos acotados. A pesar que no exista un marco teórico amplio sobre el uso de PG con fines de optimización de desempeño y rendimiento, las condiciones hasta ahora expuestas permiten la implementación de PG para este propósito.

3.3. Programación Geométrica

Es bien conocida la bondad en la consecución de óptimos globales bajo el esquema de programación geométrica [5,6,7], pero no ha sido ampliamente reportado su uso en los procesos que involucran de manera conjunta funciones de desempeño y rendimiento, más aún, no existe una fundamentación teórica que determine la factibilidad en la implementación de las ecuaciones comportamentales del transistor MOS en sub-umbral como funciones soportadas por un programa geométrico.

Trabajando simultáneamente con funciones de rendimiento y desempeño en una estructura de optimización que admite sólo una función objetivo, tanto desempeño como rendimiento deben ser representados como una función compuesta, la cual tiene que ser correctamente ajustada a una función de tipo posinomial. Es bien conocido que el óptimo global de la composición de dos o más funciones no garantiza un óptimo global para cada función, reafirmando de nuevo el alto compromiso al cual es sometido un problema de este tipo.

Además, al no ser reportado anteriormente un trabajo que involucre la optimización vía PG del diseño de un circuito con transistores operando en sub-umbral, se requiere un estudio previo de factibilidad, pues al ser este un tipo de optimización convexa requiere que las funciones que modela sean igualmente convexas.

²Se define un valor adecuado basados en el compromiso entre el error de discretización y el mínimo valor de corriente soportado por Hspice.

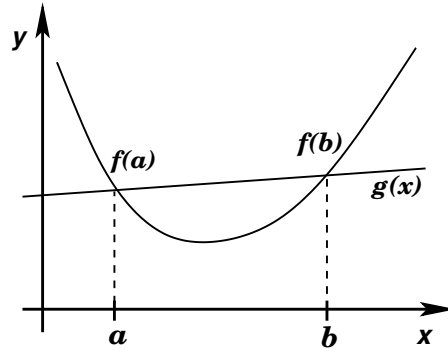


Figura 3.2: Gráfica de una función convexa y su respectiva prueba de convexidad.

3.3.1. Análisis de Convexidad y modelo sub-umbral

En realidad toda función continua en un intervalo determinado se puede modelar por medio de una función de tipo monomial o posinomial bajo el esquema de minimización de la norma basado en el método de mínimos cuadrados. No obstante, el inconveniente real no surge en el proceso de modelado, sino en aquel que involucra la determinación del conjunto de funciones ajustables al tipo de modelo usado. Al ser PG un tipo de optimización convexa, el modelado de una función cóncava por medio de una función convexa debe ser justificado, o descartado en un contexto que involucra tanto el grado de convexidad exhibido por la función como el intervalo adoptado.

Una función $f : \Re^n \rightarrow \Re$ es convexa si el **dom** f ³ es un conjunto convexo y si para todo $x, y \in \mathbf{dom} f$, y θ con $0 \leq \theta \leq 1$, se tiene que:

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y) \quad (3.14)$$

De manera más sencilla y particular la convexidad de una función $f : \Re \rightarrow \Re$ en el intervalo $[a, b]$ puede ser evaluada por medio de la traza de una recta $g(x)$ de a hasta b como se muestra en la figura 3.2. Se dice entonces que la función f es convexa cuando todo punto de la recta trazada es mayor o igual a todo valor de $f(x)$ en el interior del intervalo. Matemáticamente sería:

$$f(x) \leq \frac{f(b) - f(a)}{b - a}x + \frac{bf(a) - af(b)}{b - a} = g(x) \quad \forall x \in [a, b] \quad (3.15)$$

Debido al cambio de variable involucrado en el proceso de conversión del programa geométrico a una forma convexa, la ecuación (3.15) debe ser redefinida para la nueva variable $z = \log(x)$ como:

$$f(z) \leq \frac{f(b') - f(a')}{b' - a'}z + \frac{b'f(a') - a'f(b')}{b' - a'} = g(z) \quad \forall x \in [a', b'] \quad (3.16)$$

³**dom** f representa el dominio de la función f

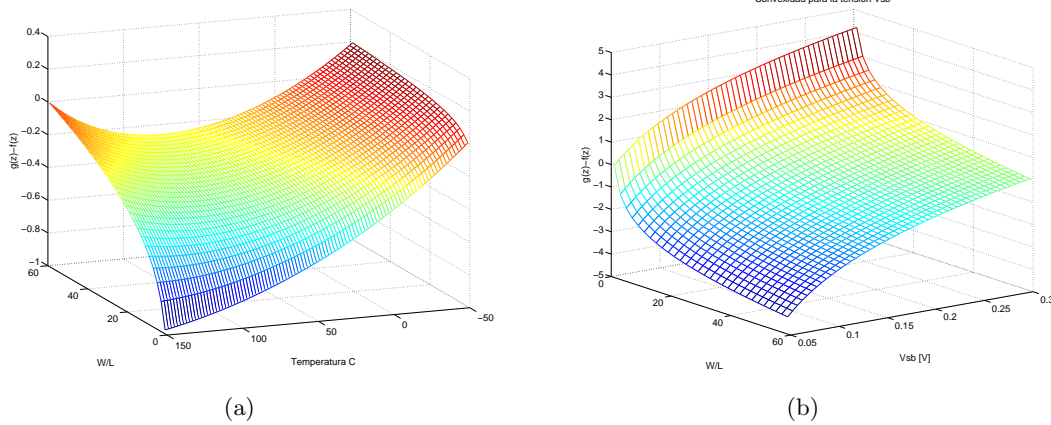


Figura 3.3: (a) Análisis de convexidad para la temperatura. (b) Análisis de convexidad para la tensión V_{SB} .

donde $a' = \log(a)$ y $b' = \log(b)$. La función $f(z)$ es equivalente a la función $f(x)$ en el intervalo $[a', b']$.

De esta manera se puede evaluar la convexidad de la tensión V_{GS} ⁴ en sub-umbral en función de los parámetros involucrados en el diseño: temperatura, V_{GS} , V_{DS} , W , L y V_{SB} . Cada uno de estos parámetros es barrido a lo largo de sus valores permisibles en conjunto con la relación de aspecto W/L que representa las variables de diseño adoptadas. El cumplimiento de la desigualdad (3.16) puede ser evaluado graficando $g(z) - f(z)$.

El análisis establece una convexidad en un conjunto aceptable de valores W/L para todos los parámetros anteriormente mencionados, exceptuando la temperatura y el voltaje surtidorpuerta V_{SB} . De esta manera, en las figuras 3.3(a) y 3.3(b) son graficados $g(z) - f(z)$.

Para el caso del análisis de la convexidad de la temperatura, en la figura 3.3(a) se evidencia una tendencia a la concavidad para los valores cercanos a 150° en contraste con una mejora de la misma para temperaturas por debajo de los 50° , justificándose de esta manera la proporción cercana a diez entre el error del modelo para 150° y el de -50° . Por otra parte, el análisis para la tensión V_{SB} establece una convexidad para valores bajos cercanos a 1 de la relación W/L , además se muestra como el grado de convexidad aumenta conforme la tensión V_{SB} se acerca a 0,3[V]. En un análisis exhaustivo del comportamiento de la fuente de referencia a optimizar mostrada en la figura 4.1, se evidenció una tendencia en el establecimiento de las tensiones V_{SB} de los transistores M_{14} y M_{12} a un intervalo estrecho y determinado. Esta particularidad permitió modelar la corriente I_{DS} a lo largo del punto de polarización establecido en el diseño semilla, presentándose así tasas de error relativamente bajas y aceptables. De otra manera, al requerirse durante el diseño de la fuente una alta precisión en la estimación de la temperatura en un amplio intervalo que comprende los $200^\circ C$, las tasas de error presentadas por el modelo

⁴Ha sido demostrado anteriormente en [6] como el modelo de V_{GS} como una función monomial genera menos error de estimación que el modelado de I_{DS} .

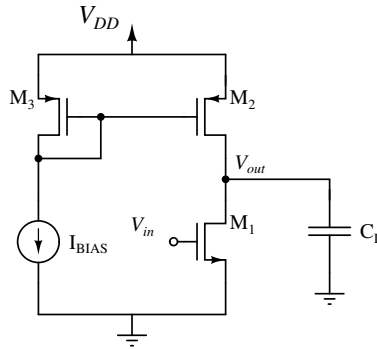


Figura 3.4: Etapa fuente común con carga activa.

no permitirían la obtención de un diseño con un coeficiente térmico adecuado para nuestros propósitos y por el contrario tenderían el programa geométrico a resultados erróneos. Es así como es descartado un modelado de la temperatura a un tipo de función soportada por PG. Esto representa un gran limitante al diseño vía PG, pero puede ser solucionada modelando sobre dos o más temperaturas determinantes en el diseño.

3.3.2. Leyes de Kirchhoff en PG

En un extenso número de casos, un programa geométrico requiere de la definición de restricciones de posinomio tipo igualdad. En el contexto del diseño de circuitos integrados, este tipo de situaciones son comúnmente evidenciadas en el establecimiento de tensiones de nodo y corrientes de rama, bajo el marco de las Leyes de Kirchhoff. Ante la imposibilidad de incluir de igualdades de tipo posinomial en el PG, las cuales comúnmente describen este tipo de situaciones, una estrategia basada en la efectividad y la precisión es requerida.

En trabajos como [6], se establece un conjunto discreto de valores de tensiones V_{DS} , para cada uno de los cuales un programa geométrico es ejecutado. La reducción del espacio de diseño en conjunto con el alto tiempo de computo requerido para el calculo de las tensiones de n ramas presentes en cualquier circuito, hacen de este un método inadecuado para muchos casos. Al no ser efectiva esta solución, se plantea un método de aproximación de las leyes de Kirchhoff a forma monomial. Para el caso de dos transistores conectados en serie en la etapa fuente común con carga activa de la figura 3.3.2 con una fuente de alimentación V_{DD} se tiene:

$$V_{DD} = V_{DS1} + V_{DS2} \quad (3.17)$$

Es posible introducir una constante de dependencia entre las dos tensiones de drenaje, de manera que: $V_{DS2} = k \cdot V_{DS1}$. Por lo tanto, la tensión de alimentación estaría representada como:

$$V_{DD} = (1 + k)V_{DS1} \quad (3.18)$$

Es evidente que (3.18) no es un tipo de función monomial, pero $(1 + k)$ puede ser razon-

ablemente ajustado por el método de mínimos cuadrados a una función monomial del tipo ak^b . El ajuste de $(1+k)$ a una función de tipo posinomial resulta ser bondadoso para valores de k mayores a la unidad, de manera tal que se se cumpla $V_{DS2} \geq V_{DS1}$. Por lo tanto, la ecuación (3.17) puede ser representada por medio de dos restricciones tipo igualdad y una desigualdad de control, expresadas como:

$$\begin{aligned} V_{DD} &= ak^b V_{DS1} \\ V_{DS2} &= kV_{DS1} \end{aligned} \quad (3.19)$$

El rango de valores permisibles de V_{DS} puede ser controlado por medio de la restricción de k a un intervalo determinado. Este método es igualmente aplicable a n elementos y en el establecimiento de la corriente en cada uno de los nodos. De esta manera se garantiza una correcta definición de la polarización de cualquier circuito en PG, sin que el espacio se vea restringido a valores discretos de V_{DS} .

3.3.3. La centralización del diseño como un programa geométrico.

Ya fue establecida la necesidad de llevar al diseño a una región menos crítica ante las variaciones de proceso, procedimiento que involucraría una inminente centralización del diseño. Dado que se pretende hacer uso de PG en la etapa de optimización conjunta de desempeño y rendimiento, una propuesta similar a la centralización del diseño compatible con PG debe ser implementada. Como la estructura de la programación geométrica permite asociar las condiciones de aceptabilidad de desempeño como restricciones de tipo posinomial, nuestra atención estará encaminada únicamente al establecimiento de una función objetivo tendiente a maximizar el rendimiento del circuito.

Dada una determinada figura geométrica convexa n -dimensional y de centro \mathbf{x}_c , el proceso de centralización de diseño a nivel general puede ser expresado como un problema de optimización convexa de la forma:

$$\text{minimizar } \|\mathbf{x} - \mathbf{x}_c\| \quad \forall \mathbf{x} \in \mathcal{D} \quad (3.20)$$

Pero la expresión (3.20) presenta dos claros inconvenientes en su implementación como programa geométrico: la función $\|\mathbf{x} - \mathbf{x}_c\|$ no puede ser fácilmente modelada como una función de tipo posinomial y el espacio de diseño establecido en (3.5) no es comprensiblemente ajustado a una forma geométrica convexa conocida.

Retomando de nuevo la definición de la aproximación del espacio de diseño establecida en (3.5) y analizando su representación gráfica esquematizada en la figura 3.5(a), se pueden categorizar dos zonas altamente importantes en el diseño para manufactura: una región de diseño no funcional que limita externamente la región de diseño \mathcal{D} y una región interna que

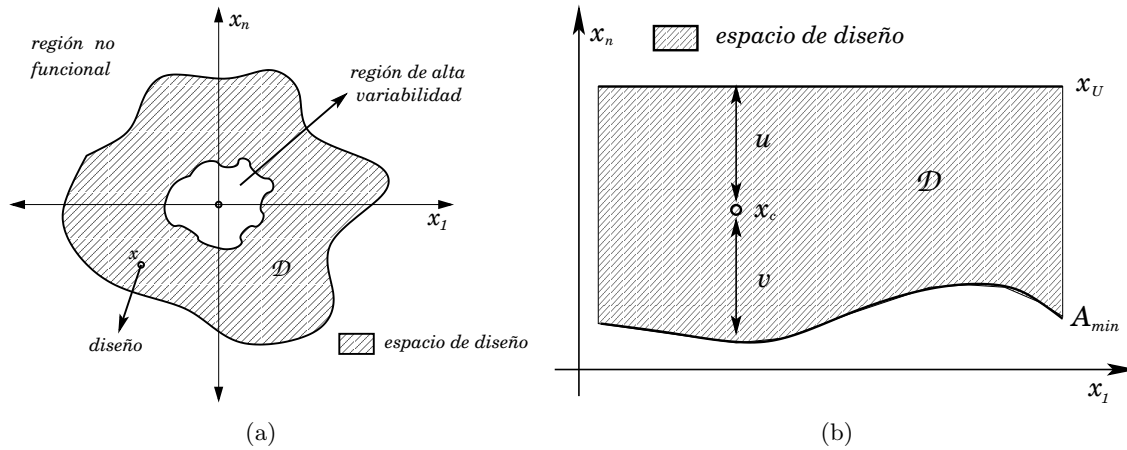


Figura 3.5: (a) Espacio de diseño para *DFM*. (b) Centralización del diseño compatible con PG.

incluye el origen y define la aceptabilidad a nivel variabilidad del circuito. La representación de la región \mathcal{D} en la figura 3.5(a) puede ser asemejada a un anillo circular n -dimensional, cuyo centro geométrico probablemente no corresponderá a la región de diseño \mathcal{D} .

Adecuándose a la definición ya dispuesta para \mathcal{D} , se establece un nuevo proceso de centralización del diseño. De esta manera, esta centralización puede ser redefinida según la figura 3.5(b) como el procedimiento de encontrar un punto \mathbf{x}_c que pertenezca a la región de diseño \mathcal{D} que maximice conjuntamente los vectores distancia \mathbf{u} y \mathbf{v} . Expresado como un problema de optimización se tendría:

$$\begin{aligned} & \text{máximizarse} \{ \|\mathbf{u}\| + \|\mathbf{v}\| \} \\ & \text{tal que } \mathbf{x} \leq \mathbf{x}_u, \wedge, \mathbf{A} \geq \mathbf{A}_{min} \end{aligned} \quad (3.21)$$

donde \mathbf{u} y \mathbf{v} corresponden a los vectores distancia de las dimensiones y el área respectivamente. Pero de nuevo la expresión (3.21) no corresponde a un problema de optimización convexa y por lo tanto no puede ser expresado como un programa geométrico, pero este puede ser redefinido en formato PG como:

$$\begin{aligned} & \text{minimizar } \{ a_1 \|\mathbf{u}\|^{-1} + a_2 \|\mathbf{v}\|^{-2} \} \\ & \text{tal que } \mathbf{A}_{min} + \mathbf{v} \leq \mathbf{A}, \wedge, \mathbf{x} + \mathbf{u} \leq \mathbf{x}_u \end{aligned} \quad (3.22)$$

donde a_1 y a_2 pueden representar simultáneamente constantes de ajuste de unidades y factores de ponderación de relevancia. La diferencia en los exponentes para \mathbf{u} y \mathbf{v} esta ligada a la relación cuadrática de las variables de diseño dimensionales \mathbf{x} con el área \mathbf{A} .

El esquema de optimización en (3.22) puede ser igualmente aplicado a la minimizar o

maximizar n funciones o variables.

Propuesta la estructura de optimización del rendimiento por medio del establecimiento de un procedimiento de centralización de diseño que involucra el uso de PG, culmina la propuesta de diseño *DFM*. Por un lado, un método práctico y sencillo para aproximar el espacio de diseño fue propuesto, el cual permite acotar las soluciones del PG en el ámbito funcional. Paralelamente, el área mínima es estimada en función de los rangos de aceptabilidad. En su conjunto, ambas estrategias constituyen la fundamentación formal de la metodología propuesta en este trabajo. El paso siguiente corresponde a certificar la funcionalidad de esta metodología en el diseño de la fuente de voltaje de referencia propuesta en [1].

Capítulo 4

Análisis de rendimiento y desempeño

Conjunto al planteamiento de una metodología de diseño, un correcto proceso de validación final debe ser llevado a cabo. El objetivo primordial de la fase de validación de una metodología de diseño puede entenderse como el ajuste empírico de la fundamentación matemática al entorno real de diseño. Es bien conocido el proceso de validación a nivel de desempeño de circuitos analógicos y de la metodología adoptada; lo contrario sucede en el proceso de verificación de metodologías de diseño *DFM*, donde el panorama de la variabilidad es aún muy ambiguo para ser completamente objetivo desde todo marco del diseño.

Por parte de este trabajo, el capítulo 2 abordó clara y precisamente los factores predominantes en la elección de un modelo de *mismatch* para validación. Pero ahora, se requiere que la verificación final aquí expuesta sea certificada en un circuito que exija de la implementación de una metodología de diseño *DFM*, como lo es la fuente de referencia de [1]. El presente capítulo está totalmente dedicado a la exposición, análisis y discusión de los resultados bajo la metodología de diseño aplicada en la fuente de referencia de Mateus, cuyo dimensionamiento es adoptado como el diseño semilla inicial.

4.1. Fuente de referencia, [1]

Como circuito a optimizar es usada la fuente de referencia de la figura 4.1 propuesta por Mateus. La necesidad de una metodología *DFM* en el diseño de este circuito está determinada esencialmente por dos razones. Por un lado, un amplio número de fuentes de referencias se encuentran ampliamente condicionadas a los factores variables tanto externos como aquellos ligados al proceso. Paralelamente, la tendencia de bajo consumo en su premisa de reducción de potencia requieren la exploración de regiones de polarización como sub-umbral, que para el caso de esta fuente conduce a amplias diferencias entre lo diseñado y lo fabricado.

Primero deben ser establecidas las variables de desempeño esencialmente importantes para

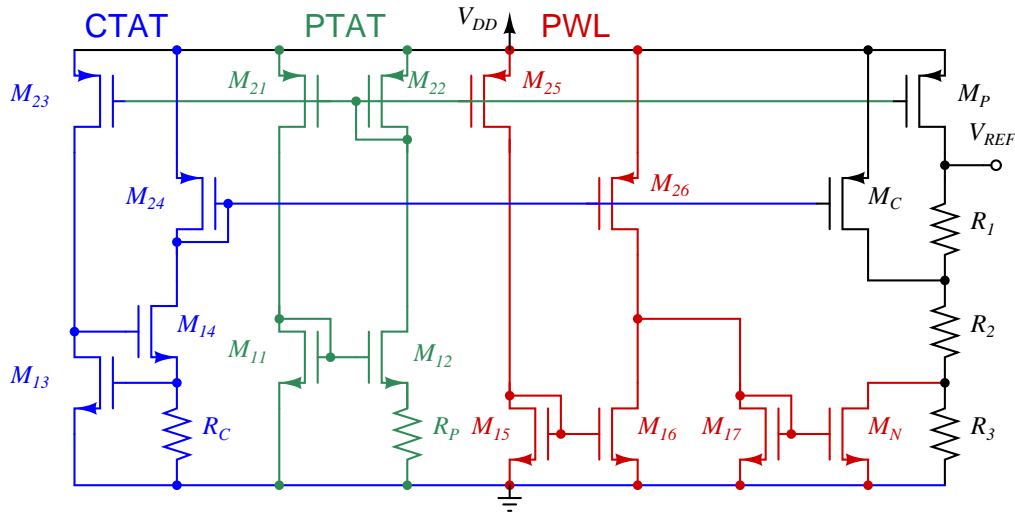


Figura 4.1: Fuente de voltaje de referencia a optimizar, tomada de [1]

Parámetro	Desempeño	Rendimiento
Potencia	$\leq 4\mu W$	$\pm 0,250\mu W$
Voltaje de referencia	$[80, 300]mV$	$\pm 2,5mV$
Coefficiente térmico	$\leq 4\mu V/^{\circ}C$	$\pm 1,2\mu V/^{\circ}C$
Temperatura media $^{\circ}C$	$[20, 60]^{\circ}C$	$\pm 8^{\circ}C$
Regulación de línea	60dB	$\pm 2dB$
<i>PSRR</i>	30dB	$\pm 2dB$

Tabla 4.1: Requerimientos de aceptabilidad para la fuente de referencia.

la fuente de referencia a optimizar, de manera que el estudio de rendimiento enfoque estas condiciones en el marco del diseño para manufactura. Como variables de desempeño son tomadas la potencia, la regulación de línea, la temperatura media, el *PSRR*, el coeficiente térmico y el voltaje de referencia. La dependencia directa entra la regulación de línea y el *PSRR*, permite abordar el proceso de optimización directamente sobre la regulación de línea, prescindiendo de esta manera del cálculo ligado *PSRR*. Lo anterior es posible dado a que la optimización de la regulación de línea implica un aumento de la ganancia en banda plana de la función de transferencia característica para el *PSRR*. Como primera medida, los resultados de desempeño obtenidos por Mateus en [1] son contrastados con las expectativas de este trabajo. Es así como se definen los rangos de aceptabilidad tanto a nivel de desempeño como a nivel de rendimiento, buscando siempre que el desempeño del circuito final sea mayor al establecido por Mateus. El proceso de elección de los rangos de aceptabilidad de rendimiento implica un conocimiento previo de la variabilidad el circuito, en conjunto con los márgenes de calidad determinados en el diseño para manufactura. La variabilidad de un circuito puede ser estimada por medio de un análisis estadístico vía Monte Carlo, el cual para propósitos comparativos debe ser equivalente al usado en la etapa de validación final.

Se garantiza aceptabilidad de desempeño por medio de la solución del PG, pero la aceptabilidad a nivel del rendimiento es pautada por la estimación de las áreas mínimas para cada uno de los elementos. Es destacable como los parámetros *PSRR* y la regulación de línea son considerados flexibles a nivel de variabilidad, pues no deterioran ampliamente el comportamiento de la tensión de referencia. Siguiendo la metodología planteada en la sección 3.2 en el capítulo 3, el análisis del área arroja las estimaciones de mínima área para aceptabilidad de la tabla 4.2. El parámetro de desempeño asociado a cada elemento y que permite evidenciar compromisos a nivel de variabilidad dentro del circuito es al cual va ligado el mayor consumo de área del elemento. Los valores máximos del área corresponden a los transistores M_{22} y M_{21} , en donde el *PSRR* muestra el peor desempeño. Los resistores de la etapa final en conjunto con el transistor M_N presentan el menor requerimiento de área. Lo anterior es debido para el caso de los resistores a la baja proporción entre su variabilidad y su valor nominal, mientras que para el caso de M_N es atribuido a su baja influencia en la determinación de parámetros relevantes como la temperatura media T_{med} . A nivel global, el área mínima total debida principalmente al voltaje de referencia es tan sólo 1/3 el área reportada por Mateus. Aparentemente, estos requerimientos de área implicarían un cumplimiento directo de las restricciones de aceptabilidad para la fuente de referencia de Mateus, pero debe ser claro que muchas de las áreas estimadas en la tabla 4.2 no cumplen con los requisitos de funcionalidad del circuito. Por ejemplo, el área mínima de las resistencias de la etapa final corresponde a $5\mu m^2$, pero dicho valor no es consecuente con la funcionalidad del circuito ya que implicaría corrientes de rama del orden de las centenas de *mA*. Otro caso relevante corresponde al consumo de área del transistor M_{24} ; en el ámbito de la variabilidad el requerimiento de área para este elemento es de tan sólo de $581\mu m^2$ comparado con un valor de $2000\mu m^2$ reportado por Mateus en [1]. A nivel de funcionalidad es claro que M_{24} requiere de grandes áreas para mantenerse en sub-umbral en su punto crítico correspondiente a $-50^\circ C$, pero esta tendencia no es evidenciada en la tabla comparativa de *mismatch*. En suma, todos los resistores en conjunto con transistores como M_{24} producen un considerable aumento del área para cumplir con las condiciones de funcionalidad del circuito.

Es importante notar como prácticamente todo el bloque *PWL* encuentra su peor desempeño en la temperatura media. Por otro lado, se resalta lo influyentes que resultan ser los transistores M_{21} y M_{22} en las variaciones del *PSRR*; esto es debido a que el punto de polarización de las tensiones V_{GS} de los transistores *PMOS* lo establece el bloque *PTAT* por medio de los transistores M_{21} y M_{22} . Además de limitar el espacio de diseño para cumplir con los límites de aceptabilidad, este método permite evidenciar los compromisos del dimensionamiento en función de los parámetros de desempeño.

Elemento	Área mínima [μm^2]	Peor desempeño
M_{11}	325,51	PSRR
M_{12}	105,86	Potencia
M_{21}	637,21	PSRR
M_{22}	635,12	PSRR
R_P	133,44	Coefficiente térmico
M_{13}	30,40	Regulación de línea
M_{14}	320,68	Regulación de línea
M_{23}	73,11	Potencia
M_{24}	581,89	Regulación de línea
R_C	166,00	Potencia
M_{15}	166,28	Temperatura media
M_{16}	168,41	Temperatura media
M_{17}	12,22	Temperatura media
M_{25}	303,32	Temperatura media
M_{26}	323,49	Temperatura media
M_P	284,28	Voltaje de referencia
M_C	423,35	Regulación de línea
M_N	5	Regulación de línea
R_1	5	Regulación de línea
R_2	5	Regulación de línea
R_3	5	Regulación de línea
Área total	4694,3	Voltaje de referencia.

Tabla 4.2: Mínima área y análisis de *mismatch*

4.2. Optimización vía programación geométrica.

Fundamentados en los análisis y planteamientos efectuados en las secciones anteriores, se procede a diseñar la fuente de voltaje de referencia vía PG. Ante la imposibilidad de incluir directamente la temperatura como variable independiente en los modelos, un proceso de modelado alterno se hace necesario. El estudio nos mostró como las tensiones puerta-fuente para transistores *NMOS* y *PMOS* exhiben un comportamiento altamente lineal con la temperatura, característica que puede ser aprovechada a favor en el proceso de modelado. De esta manera el problema de optimizar la fuente de referencia de la figura 4.1 involucraría la inclusión de n definiciones de restricciones, cada una de las cuales pertenecerían a un valor específico de temperatura en el intervalo de interés. La alta linealidad de las tensiones V_{GS} respecto a la temperatura permitirían establecer en principio la necesidad de tan sólo 2 puntos de modelado, siendo los más críticos e importantes los extremos ubicados a $-50^\circ C$ y $150^\circ C$.¹ Igualar los voltajes de referencia en $-50^\circ C$ y $150^\circ C$ bajo la suposición de corrientes de salida de los bloques *PTAT* y *CTAT* lineales, permite cancelar las componentes dependientes de la

¹Resultados posteriores muestran como el punto correspondiente a T_{med} puede ser correctamente establecido variando las dimensiones de M_{17} y M_{26} .

temperatura y obtener una salida idealmente con un coeficiente de temperatura cero.

La inclusión del compendio de restricciones para cada temperatura involucra que las variables de diseño que describen los parámetros eléctricos sean declaradas tantas veces como puntos en el intervalo fueron adoptados. Lo anterior sumado a la inclusión de variables de control para cada rama o nodo tornarían el problema de optimización algo complejo e ineficiente. Con el propósito de aumentar la comprensión del PG aplicados en los procesos de centralización del diseño, la optimización del circuito de la figura 4.1 no es llevada a cabo de manera conjunta sino para cada uno de sus bloques esenciales: *PTAT*, *CTAT*, *PWL* y la rama final. Este procedimiento tiene la desventaja de reducir el espacio de diseño pero permite evidenciar claramente las tendencias y los compromisos en el proceso de optimización.

CVX es el *software* usado para encontrar la solución de los programas geométricos, siendo este una herramienta gratuita y altamente versátil para la optimización convexa [30]. La optimización bloque por bloque sigue el orden establecido en un proceso común de diseño, y que involucra en primera medida los bloques más independientes e influyentes del circuito. Bajo este razonamiento el primer bloque a tratar correspondería a *PTAT* y el último a la rama final.

4.3. Bloque *PTAT*

Encargado de la compensación proporcional a la temperatura de la fuente de referencia, es posiblemente el bloque más importante en conjunto con la etapa de corriente lineal por partes en lo que al ámbito de la variabilidad se refiere. Es además el responsable de establecer las tensiones V_{GS} para una gran proporción de transistores *PMOS*. Sus parámetros de desempeño destacables y que pueden influir en el comportamiento global del circuito son: potencia, la regulación de línea (RL) y coeficiente térmico. La definición del PG para el bloque *PTAT* parte de un adecuado modelado de las tensiones en las cuatro ramas presentes en el circuito de la figura 4.1 por medio de la propuesta establecida en la subsección 3.3.2.

Las restricciones de desempeño definidas en el PG buscan establecer un diseño mejor o igual que el dado inicialmente, garantizando así una mejora del desempeño. Por ejemplo, los compromisos de potencia con valor del resistor del bloque *PTAT* son incluidos por medio de la definición de dos restricciones. La primera limita el resistor a valores menores o iguales a $250k\Omega$, valor correspondiente al reportado por Juan Mateus. La segunda, por su parte, limita el aumento de la corriente promovido por la disminución del resistor al valor establecido igualmente por Mateus. Para el caso de la regulación de línea, en *PTAT* esta es estimada por:

$$LR_{PTAT} \approx \frac{g_{ds11}R_p}{1 + g_{ds11}R_p} \quad (4.1)$$

Por lo tanto, esta puede ser incluida como una restricción desigualdad, la cual establece un valor de *PSRR* mínimo determinado por el diseño semilla de Mateus. De otro lado, las

Minimizar	
Distancia al centro:	$\ \mathbf{u}\ ^{-2} + \ \mathbf{v}\ ^{-1}$
Sujeto a:	
Restricciones DFM	
Área mínima	$\mathbf{A}_{min} + \mathbf{u} \leq \mathbf{A}$
Funcionalidad	$\mathbf{x} + \mathbf{v} \leq \mathbf{x}_u$
Desempeño @150°C	
Resistencia máxima	$R_p \leq 250k\Omega$
Potencia	$I_{ptat} \leq 380 \text{ nA}$
RL mínima	$RL_{min} \cdot (1 + g_{ds1}R_p)g_{ds1}^{-1}R_p \leq 1$
Condiciones estructurales @ 150°C, -50°C	
$L_{11} = L_{12}, \wedge, L_{21} = L_{22}$	$L_{11} == L_{12}, \wedge, L_{21} == L_{22}$
$W_{12}/W_{11} \geq 2$	$2W_{11} \cdot W_{12}^{-1} \leq 1$
$V_{DD} = V_X + V_{DS22}$	$V_{DD} == a_{k_1} \cdot V_{DS22} \cdot k_1^{b_{k_1}}$
	$V_X == k_1 \cdot V_{DS22}$
$V_X = V_{ptat} + V_{DS12}$	$V_{DD} == a_{k_2} \cdot V_{DS22} \cdot k_2^{b_{k_2}}$
	$V_{DS12} == k_2 \cdot V_{ptat}$
$V_{DD} = V_{DS11} + V_{DS21}$	$V_{DD} == a_{k_3} \cdot V_{DS11} \cdot k_3^{b_{k_3}}$
	$V_{DS21} == k_3 \cdot V_{DS11}$
$V_{GS11} = V_{ptat} + V_{GS12}$	$V_{GS11} == a_{k_4} \cdot V_{DS22} \cdot k_4^{b_{k_4}}$
	$V_{GS12} == k_4 \cdot V_{ptat}$
$V_{GS21} = V_{GS22}$	$V_{GS21} == V_{GS22}$
$V_{GS21} = V_{DS22}$	$V_{GS21} == V_{DS22}$
$V_{GS11} = V_{DS11}$	$V_{GS11} == V_{DS11}$

Tabla 4.3: Programa geométrico para el bloque *PTAT*.

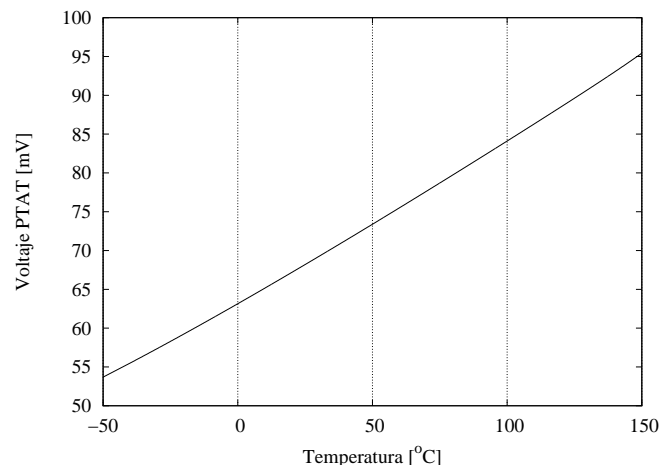
longitudes de canal son condicionadas a ser iguales en cada uno de los espejos de corriente. Junto a esto, es establecida una restricción de mínima proporción de anchos W_{12}/W_{11} que tiene como función principal la de mantener funcional el bloque *PTAT*. Luego, la función objetivo es construida siguiendo el planteamiento y los resultados obtenidos en las secciones anteriores. Es interesante destacar que los parámetros de desempeño fueron medidos a 150°C, donde la mayoría de estos presentan sus valores más críticos. El compendio de restricciones que garantizan una correcta polarización del bloque *PTAT* y valores adecuados de desempeño pueden ser vistas en la tabla 4.3.

Las variables de diseño establecidas por el programa geométrico pueden ser vistas en la tabla 4.4. Es destacable el aumento de la longitud para cada uno de los transistores y su tendencia al establecimiento en un valor cercano a $20\mu m$. Además, se evidencia una similitud en la relación W_{12}/W_{11} establecida por Juan Mateus a la dada por el PG, en conjunto con la tendencia del proceso de optimización a reducir el área por medio del establecimiento de una baja relación W_{21}/L_{21} . A pesar de esto, la estimación del área brindada por PG es algo más de tres veces a la establecida por Mateus. El valor de la tensión V_{ptat} es estimado por medio del método propuesto en la sección 3.3.2, obteniéndose así diferencias entre simulación y PG

Variable	W_{11}	L_{11}	W_{12}	W_{21}	L_{21}	W_{RP}	L_{RP}
PG	$85,99\mu m$	$19,36\mu m$	$867,90\mu m$	$157,64\mu m$	$20,648\mu m$	$3\mu m$	$611,81\mu m$

Tabla 4.4: Dimensiones de los elementos del bloque *PTAT*.

Parámetro @ 150°C	PG	Hspice	Juan Mateus
Potencia μW	0,98	1	0,85
V_{ptat} [mV]	94,79	95,42	84,80
LR [dB]	63,13	64,59	58,63
Área μm^2	26817,6		7512,5

Tabla 4.5: Comparación de resultados *PTAT*.Figura 4.2: Comportamiento de V_{ptat} en función de la temperatura de trabajo.

cercanas a $1mV$ como lo muestra la tabla 4.5. El aumento de algo más del 15% en potencia es debido a la disminución del valor del resistor R_P . La gráfica del voltaje *PTAT* en función de la temperatura puede ser vista en 4.3

4.4. Bloque *CTAT*

Una gran proporción del correcto comportamiento funcional de la fuente de referencia es debida a *CTAT*. Los transistores M_{24} y M_{14} representan dos elementos esencialmente importantes en el establecimiento de la mínima tensión de polarización V_{DD} . Además, las exigencias de área para el establecimiento en la región de sub-umbral del transistor M_{24} podrían ser tan altas que desbordarían el mismo espacio de diseño establecido al inicio del flujo de diseño.

El programa geométrico implementado para la optimización del bloque *CTAT* es descrito en la tabla 4.6. Se deben garantizar los intervalos correctos para las tensiones de dreno de las ramas de *CTAT* en su peor caso ubicado en $-50^\circ C$. Por un lado se define una restricción

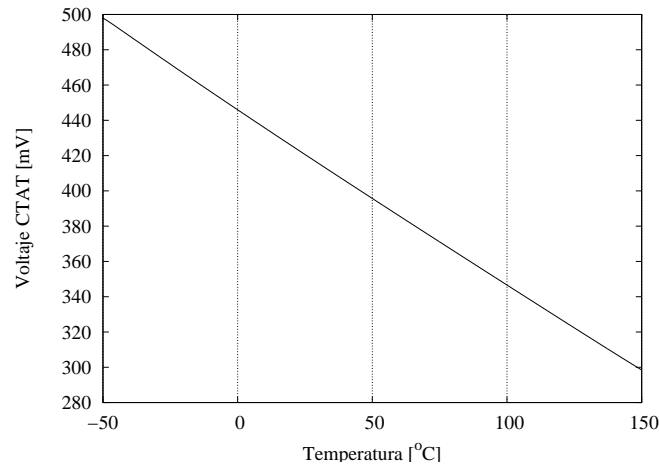


Figura 4.3: Comportamiento de V_{ctat} para el rango de temperaturas de trabajo.

superior para la constante de control k_{a5} que garantice un valor de tensión V_{GS14} que no deje por fuera de sub-umbral a M_{14} . Consecuentemente, la tensión mínima de polarización V_{DD} es regulada por medio de la correcta definición del límite superior de k_{a7} de manera tal que se garantice que la tensión dren-fuente del transistor M_{14} sea mayor al voltaje térmico U_t .

Otro punto relevante a establecer y que está íntimamente vinculado a las variaciones sistemáticas de los espejos de corriente, corresponde a la definición de las mismas longitudes para cada uno de los transistores que comparten una tensión V_{GS} común. Como función de desempeño relevante se selecciona la regulación de línea, la regulación de línea y la potencia. El PSRR y en su defecto la regulación de línea son definidos a ser condiciones del programa geométrico, con la premisa de mantener igual o mejor el desempeño de la fuente de referencia con respecto al diseño de Mateus. La función objetivo está compuesta por una componente de centralización de diseño y otra relacionada con el valor de las corrientes de rama. La optimización busca minimizar también la potencia, pues se ha visto como este bloque puede llegar a consumir hasta el 35 % de la potencia total del circuito.

El establecimiento de las condiciones estructurales por medio de las tensiones de rama debe ser cuidadosamente abordado. Se cuenta con tensiones de dren que fijan a su vez tensiones de puerta lo cual exige un bajo error de estimación para el modelo. La experiencia demostró cómo la posibilidad de encontrar una solución para estos casos era altamente dependiente del intervalo de valores de L asumidos en el modelo inicial, pues el establecimiento de una longitud del transistor a través del modelo liga a una tensión V_{DS} específica, que para el caso de no cumplir con las restricciones de rama causa que el programa geométrico sea no factible.

El dimensionamiento vía PG produjo los resultados consignados en la tabla 4.7. El incremento del área para el control de la variabilidad conserva la misma tendencia de relaciones W/L establecidas anteriormente, exceptuando para el transistor M_{14} y M_{23} . El aumento del área para M_{14} según el análisis del *mismatch* de la tabla 4.2 es debido a la variabilidad de

Minimizar	
Distancia al centro y potencia:	$\ \mathbf{u}\ ^{-2} + \ \mathbf{v}\ ^{-1} + I_p + I_c$
Sujeto a:	
Restricciones DFM	
Area mínima	$\mathbf{A}_{min} + \mathbf{u} \leq \mathbf{A}$
Funcionalidad	$\mathbf{x} + \mathbf{v} \leq \mathbf{x}_u$
@ $T = -50^\circ C$	$k_{5m\acute{a}x} = \frac{V_{DS13m\acute{a}x}}{(V_{DD} - V_{DS13m\acute{a}x})}$
@ $T = -50^\circ C$	$k_{7m\acute{a}x} = \frac{V_{ctat} _{T=150^\circ C}}{U_T}$
Desempeño @ $-50^\circ C$	
Regulación de línea	$1 \geq LR_{min}^{-1} \cdot (1 + g_{ds4} R_c) g_{ds4}^{-1} R_c^{-1}$
Potencia	$I_3 + I_4 \leq 1, 1\mu A$
Condiciones estructurales @ $150^\circ C, -50^\circ C$	
$V_{DD} = V_{DS13} + V_{DS23}$	$V_{DD} == a_{k_5} \cdot V_{DS23}^{bk_5}$
	$V_{DS13} == k_5 \cdot V_{DS23}$
$V_{DS23} = V_{GS14} + V_{ctat}$	$V_{DS13} == a_{k_6} \cdot V_{ctat}^{bk_6}$
	$V_{GS14} == k_6 \cdot V_{ctat}$
$V_Y = V_{ctat} + V_{DS14}$	$V_Y == a_{k_7} \cdot V_{ctat}^{bk_7}$
	$V_{DS14} == k_7 V_{ctat}$
$V_{DD} = V_Y + V_{DS14}$	$V_{DD} == a_{k_8} \cdot V_{DS24}^{bk_8}$
	$V_Y == k_8 \cdot V_{DS24}$
$V_{GS13} = V_{ctat}$	$V_{GS13} == V_{ctat}$
$V_{GS24} = V_{DS24}$	$V_{GS24} == V_{DS24}$
$V_{ctat} = I_4 \cdot R_c$	$V_{ctat} == I_4 \cdot R_c$
$L_{23} == L_{22}$	$L_{23} L_{22}^{-1} == 1$
$V_{GS23} = V_{GS21}$	$V_{GS23} == V_{GS21}$

Tabla 4.6: Programa Geométrico para el bloque CTAT.

Variable	W_{13}	L_{13}	W_{14}	L_{14}	W_{23}	L_{23}	W_{24}	L_{24}	L_{RC}	W_{RC}
PG [μm]	47, 115	24	272, 77	4, 60	28, 04	20, 64	930, 15	20	616, 10	2, 40

Tabla 4.7: Dimensiones de los elementos del bloque CTAT.

Parámetro @ $-50^\circ C$	PG	Hspice	Juan Mateus
Potencia μW	2,384	1,319	1,298
$V_{ctat} [mV]$	574,96	497,93	505,22
LR [dB]	37,31	35,42	44,48
Área μm^2	23045,88		3958,5

Tabla 4.8: Comparación de resultados CTAT.

la regulación de línea, en la cual este elemento juega un papel primordial. De otro lado, la disminución en la relación de M_{23} es debida a la minimización de la potencia establecida en conjunto con el rendimiento como función objetivo del PG. El consumo de área es mayor al establecido por Mateus por casi 6 veces.

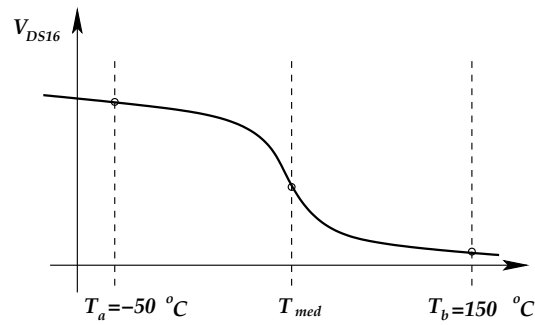


Figura 4.4: Tensión V_{DS16} vs. temperatura

En la tabla comparativa 4.8 se destaca una sobre-estimación vía PG de casi el doble de la potencia con respecto a los resultados obtenidos en simulación. La explicación más acertada a estas diferencias podría situarse en el proceso de modelado de la tensión V_{SB} , en la cual se había demostrado anteriormente un bajo grado de convexidad y por ende unas tasas altas de error por estimación.

4.5. Bloque *PWL*

El bloque *PWL* establece una corriente lineal por partes la cual sumada a las componentes no-lineales de I_{ptat} e I_{ctat} a bajas temperaturas permite la obtención de un coeficiente térmico altamente simétrico con respecto a la temperatura media T_{med} . En el trabajo original [1], Mateus había sugerido disminuir la alta variabilidad cercana a los $100^{\circ}C$ exhibida por uno de los parámetros más importantes para el correcto funcionamiento del circuito: T_{med} . En la etapa de *PWL*, este parámetro es establecido por medio de la relación de aspecto y las longitudes de cada uno de sus transistores. Se evidencia como la temperatura media puede establecerse prescindiendo de los transistores M_{17} y M_N , elementos que si tienen una gran influencia en la estimación del voltaje de referencia por lo que serán incluidos luego en el análisis del bloque final.

El comportamiento de la tensión de drenaje para M_{16} en función de la temperatura es esquematizado en la figura 4.4. Se establecen tres puntos esenciales para el modelado: T_a , T_{med} y T_b . A diferencia de otros bloques anteriores, este requiere un nuevo modelo evaluado en la temperatura media T_{med} . La temperatura media en el contexto de la fuente de voltaje de referencia, se define como la temperatura a la cual la tensión drenaje-fuente del transistor M_{17} es igual al voltaje de Fermi ϕ_F . Pero el establecimiento de una expresión que relacione la temperatura media con las dimensiones de los transistores del bloque *PWL* es un proceso bastante complejo para ser abordado aquí. Entonces, para garantizar una flexibilidad en las soluciones otorgadas por el programa geométrico, y basados en los altos errores de estimación ligados a las tensiones de drenaje, el programa geométrico es puesto a seguir la filosofía del mejor esfuerzo. De esta manera se definen para cada una de las tres temperaturas de interés

Minimizar	
Distancia al centro (mejor esfuerzo):	$\ \mathbf{u}\ ^{-2} + \ \mathbf{v}\ ^{-1} + p^2 + q^2 + r^2$
Sujeto a:	
Restricciones <i>DFM</i>	
Área mínima	$\mathbf{A}_{min} + \mathbf{u} \leq \mathbf{A}$
Funcionalidad	$\mathbf{x} + \mathbf{v} \leq \mathbf{x}_u$
Desempeño @150°C	
$T_{med} = 50^\circ C$	$W_{16} \cdot L_{16}^{-1} == 18,5 \cdot p$
Potencia	$I_6 + I_5 \leq 110nA @ T = 150^\circ C$
	$I_6 + I_5 \leq 160nA @ T = -50^\circ C$
Condiciones estructurales	
$V_{DS16} == \phi_F, @ T = T_{med}$	$V_{DS16} == 190mV$
$V_{DS16} < V_{th}, @ T = -50^\circ C$	$V_{DS16} == 480mV \cdot r$
$V_{DS16} < \phi_F, @ T = 150^\circ C$	$V_{DS16} == 20mV \cdot q$
$I_{26} = I_{16} + I_{17}$	$I_{26} == ak_x I_{17} k_x^{bk_x}$
	$I_{16} = k_x I_{17}$
$V_{DD} = V_{DS15} + V_{DS25}$	$V_{DD} == a_{k_9} V_{DS15} k_9^{bk_9}$
	$V_{DS25} = k_9 V_{DS15}$
$V_{DD} = V_{DS16} + V_{DS26}$	$V_{DD} == a_{k_{10}} V_{DS16} k_{10}^{bk_{10}}$
	$V_{DS25} = k_{10} V_{DS15}$
$V_{GS15} = V_{GS16} @ T = -50^\circ C, T_{med}, 150^\circ C$	$V_{GS15} == V_{GS16}$
$V_{DS15} = V_{GS15} @ T = -50^\circ C, T_{med}, 150^\circ C$	$V_{DS15} == V_{GS15}$
$V_{GS25} = V_{GS21} @ T = -50^\circ C, T_{med}, 150^\circ C$	$V_{GS25} == V_{GS21}$
$L_{15} = L_{16}$	$L_{15} \cdot L_{16}^{-1} == 1$

Tabla 4.9: Programa geométrico para el bloque *PWL*

Variable	W_{25}	L_{25}	W_{15}	L_{15}	W_{16}	W_{26}	L_{26}
PG [μm]	44,17	20,64	69,75	11,57	168,29	142,51	20

Tabla 4.10: Dimensiones de los elementos del bloque *PWL*.

Parámetro @ 150°C	PG	<i>Hspice</i>	Juan Mateus
Potencia nW	396,03	321,399	164,7
T_{med} °C	50	45	47,88
Área μm^2	6516,16		216

Tabla 4.11: Comparación de resultados *PWL*.

restricciones de voltajes V_{DS} , los cuales tendrán la forma: $V_{DS} = pV_{DS_{sim}}$, donde $p \geq 1$. El valor $V_{DS_{sim}}$ es obtenido por medio de simulación del circuito semilla. Luego la función objetivo es construida como la suma de las normas de los vectores de centralización de diseño más los valores cuadrados en cada temperatura p , q y r . Junto a esta definición de restricciones, se establece una relación W_{16}/L_{16} a un valor de 18,5 obtenido por medio de simulación y que permite fijar la temperatura media T_{med} a un valor de 50°C.

4.6. Rama final

La definición de un conjunto de restricciones de desempeño para cada uno de los anteriores bloques brinda garantías de correcto funcionamiento a la rama final. Sin embargo, funciones de desempeño global como el coeficiente térmico, la regulación de línea y el voltaje de referencia están altamente condicionados al diseño de la etapa final, razón por la cual deben ser correctamente incluidos dentro del PG.

El objetivo del programa geométrico de la rama final corresponde a una función compuesta que involucra la centralización del diseño y el coeficiente térmico. Para minimizar adecuadamente el coeficiente térmico, en el PG es definida la variable p , la cual representa un factor de dependencia entre los voltajes de referencia para $-50^{\circ}C$ y $150^{\circ}C$. Para obtener un coeficiente de temperatura bajo, el PG busca igualar ambas tensiones de referencia por medio de la minimización del valor p , sujeto a $p \geq 1$. La experiencia reveló como el fijar el valor del voltaje de referencia alteraba enormemente el resultado del programa geométrico. Debido a esto, el voltaje de referencia se deja a libre elección del PG. Funciones de desempeño como el $PSRR$ y la potencia son condicionadas a ser mayores y menores respectivamente que las establecidas por Mateus. En algunos casos estas condiciones tuvieron que ser relajadas, pues el PG era establecido a no tener solución.

Los valores de coeficiente térmico establecidos como límites de aceptabilidad en la tabla 4.1 requieren que la estimación del voltaje de referencia brindada por la solución del programa geométrico tenga un error aproximadamente igual a 0,1%. A nivel de modelado, estas tasas de error son tan sólo posibles si el modelado es efectuado alrededor del punto óptimo ya conocido. La solución arrojada por el PG brinda una aproximación al comportamiento de real de la fuente, razón por la cual se requiere de un posterior ajuste manual en el cual están incluidas las variables de diseño: W_N , W_C , W_P , L_{R1} , L_{R2} y L_{R3} . El aumento de $PSRR$ que se logro con respecto al diseño de Mateus fue debido al incremento de la longitud de canal de cada uno de los transistores, promoviendo de esta manera bajos valores de transconductancia drenó-fuente.

4.7. Discusión

En la tabla 4.14 son contrastados los consumos de área obtenidos en este trabajo y los establecidos por Mateus. Es evidente el aumento del área por un factor mayor a 6 en la metodología DFM sobre la clásica estrategia de diseño. Por ahora, las estimaciones de área de este trabajo parecen propias de un diseño sobre-dimensionado, por lo cual deben ser revalidadas o descartadas. Para efectuar un análisis de la influencia del área en la variabilidad de la fuente de referencia anteriormente diseñada, la dependencia de la sensibilidad de cada uno de las variables de desempeño es graficada en la figura 4.5 en función del área normalizada al diseño establecido. En esta gráfica se exhibe una tendencia exponencial en las sensibilidades

Minimizar	
Distancia al centro y TC:	$\ \mathbf{u}\ ^{-2} + \ \mathbf{v}\ ^{-1} + \ \mathbf{p}\ ^2$
Sujeto a:	
Restricciones DFM	
Área mínima	$\mathbf{A}_{min} + \mathbf{u} \leq \mathbf{A}$
Funcionalidad	$\mathbf{x} + \mathbf{v} \leq \mathbf{x}_u$
Desempeño	
RL mínima	$PSRR_{min}(1 + R_T g_{dsp})R_T^{-1} g_{mp}^{-1} \leq 1$
Resistencias	$R_1 + R_2 + R_3 \leq 205k\Omega$
Potencia	$I_C + I_P \leq 1, 3\mu A$
Coefficiente térmico	$V_{ref}(T = 150^\circ C) = p \cdot V_{ref}(T = -50^\circ C)$
Condiciones estructurales	
$V_{DD} = V_{DSP} + V_{ref}$	$V_{DD} == a_{k_{12}} V_{ref} k_{12}^{b_{k_{12}}}$ $V_{DSP} == k_{12} V_{ref}$
$V_{ref} = V_M + V_{R1}$	$V_{ref} == a_{k_{13}} V_{R1} k_{13}^{b_{k_{13}}}$ $V_M == k_{13} V_{R1}$
$V_M = V_{R2} + V_{R3}$	$V_M == a_{k_{14}} V_{R2} k_{14}^{b_{k_{14}}}$ $V_{R3} == k_{14} V_{R2}$
$V_{DSC} = V_{R1} + V_{DSP}$	$V_{DSC} == a_{k_{15}} V_{R1} k_{15}^{b_{k_{15}}}$ $V_{DSP} == k_{15} V_{R1}$
$I_M = I_C + I_P$	$I_M == a_{k_{16}} I_P k_{16}^{b_{k_{16}}}$ $I_C == k_{16} I_P$
$I_M = I_{MN} + I_{R3}$	$I_M = a_{k_{17}} I_P k_{17}^{b_{k_{17}}}$ $I_{R3} == k_{17} I_{MN}$

Tabla 4.12: Programa geométrico para la rama final.

Variable	W_P	W_C	W_N	L_N	W_{17}	W_{R1}	L_{R1}	W_{R2}	L_{R2}	W_{R3}	L_{R3}
PG	227,188	603	80,82	80,82	127,69	3	86,3	3	73,8	3	92,5

Tabla 4.13: Dimensiones de los elementos de la rama final.

de las funciones de desempeño con el área total del dispositivo. De nuevo, el parámetro que exhibe la más alta variabilidad corresponde al voltaje de referencia, donde una reducción de una décima parte del área establecida originaría un aumento de 35 veces la variabilidad exhibida en el diseño bajo *DFM*. Además, es posible constatar la bondad del aumento del área para este circuito. Con respecto al diseño de Mateus, se obtiene una mejora en el comportamiento ante el *mismatch* de aproximadamente 5 veces para la mayoría de las variables de desempeño exceptuando la potencia, cuya sensibilidad no es destacable frente a otras.

En la tabla 4.14 se muestra además la distribución del área del chip para cada uno de los elementos de la fuente de referencia. Es predominante el aumento en la proporción del área para los transistores correspondientes al bloque *PWL*, los cuales son relevantes en la determinación de la temperatura media T_{med} . Esta propensión está ampliamente motivada por los márgenes de aceptabilidad estrechos adoptados para la temperatura media T_{med} y

Elemento	Este trabajo		Juan Mateus	
	Área	% del total	Área	% del total
M_{11}	1664,76	1,8	600	4,01
M_{12}	16802	18,13	4800	32,05
M_{21}	3253,68	3,51	400	2,67
M_{22}	3253,68	3,51	400	2,67
R_P	1835,43	1,98	1312,5	8,76
M_{13}	1130,64	1,22	10	0,07
M_{14}	1227,74	1,31	56	0,37
M_{23}	578,76	0,62	80	0,53
M_{24}	18603	20,07	2000	13,35
R_C	1480,75	1,6	1812,5	12,1
M_{15}	807,22	0,87	12	0,08
M_{16}	1947,12	2,1	12	0,08
M_{17}	12309,32	13,28	12	0,08
M_{25}	912,1	0,98	96	0,64
M_{26}	2850,2	3,07	96	0,64
M_P	4691,16	5,06	800	5,34
M_C	12060	13,01	1312	8,76
M_N	6531,87	7,05	96	0,64
R_1	258,9	0,28	281,75	1,88
R_2	221,4	0,24	393,75	2,63
R_3	227,5	0,3	393,75	2,63
Total	92694,81	100	14975,75	100

Tabla 4.14: Comparación de consumos de área

detallados en la tabla 4.14. Mateus por su parte, presenta valores bajos de áreas para el bloque PWL , razón por la cual se exhibe una fluctuación del parámetro T_{med} cercana a los $100^\circ C$. El elemento que más consume área, llegando a ocupar el 20% del área total, corresponde a M_{24} . Este requerimiento de área esté motivado esencialmente en el ámbito de funcionalidad, evitando que M_{24} salga de sub-umbral.

Aunque la proporción del aumento del área requerida para el control de la variabilidad pueda ser guiada de manera intuitiva, el análisis de sensibilidad es propicio para establecer dependencias entre la variabilidad global del conjunto de parámetros de desempeño del circuito, en conjunto con las variables de diseño.

4.8. Resultados

Una vez establecidas las dimensiones de diseño obtenidas mediante la aplicación de programación geométrica, se procede a validar el circuito diseñado. El proceso de validación implica un análisis de variabilidad efectuado por medio de Monte Carlo, y bajo el modelo de *mismatch* propuesto por Drennan [23]. Los efectos del *mismatch* son estimados para cada elemento del

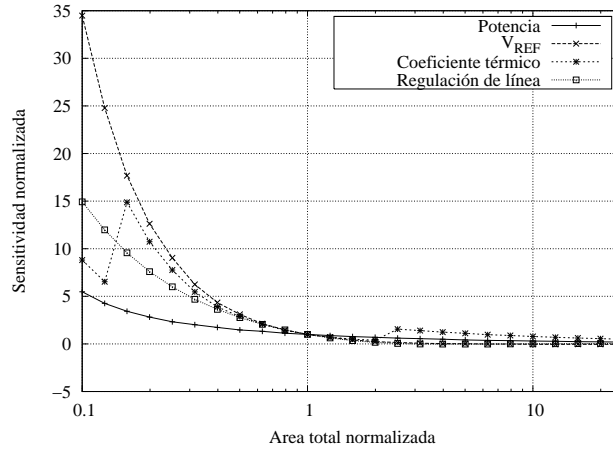


Figura 4.5: Sensibilidad de los parámetros de desempeño de la fuente de referencia.

Especificación	Potencia [μW]	V_{ref} [mV]	TC [$\mu V/^{\circ}C$]	T_{med} [$^{\circ}C$]	$PSRR$ [dB]	LR [dB]
Este trabajo (Pre-layout)	4,134 $\pm 0,0607$	98,17 $\pm 2,259$	3,64 $\pm 1,737$	44,95 $\pm 20,46$	28,78 $\pm 0,125$	63,73 $\pm 0,119$
Este trabajo (Post-layout)	4,126 $\pm 0,0432$	98,04 $\pm 2,01$	2,47 $\pm 1,447$	37,40 $\pm 21,03$	24,319 $\pm 0,115$	51,682 $\pm 0,192$
Juan Mateus	3,833 $\pm 0,327$	199,95 $\pm 16,41$	14,62 $\pm 55,86$	35,714 $\pm 87,88$	25,32 $\pm 1,239$	47,17 $\pm 0,699$

Tabla 4.15: Resultados de desempeño y rendimiento obtenidos.

circuito por medio de la inserción de una fuente de corriente entre drenos y surtidor. Se efectuaron simulaciones tipo Monte Carlo para un número de muestras igual a 800 de las cuales se extrajeron las medias y las varianzas 6σ obteniéndose así los resultados en la tabla 4.15. En el diseño pre-layout se certifica un cumplimiento de las restricciones de desempeño definidas en la tabla 4.1 para todas las variables de desempeño exceptuando el $PSRR$ y la potencia; para el caso post-layout, las capacitancias parásitas hacen además que regulación de línea no cumpla las expectativas. El consumo de potencia mayor al reportado por Mateus obedece principalmente a la disminución en el valor de los resistores. De otro lado, las condiciones de aceptabilidad de rendimiento se garantizan para las variables de desempeño: potencia, voltaje de referencia, $PSRR$ y regulación de línea. Mientras tanto, el coeficiente térmico a pesar de que no cumple las expectativas presenta valores de varianza cercanos a los establecidos en las condiciones iniciales de aceptabilidad; no se percibe la misma tendencia para la temperatura media T_{med} , cuya varianza en el análisis de Monte Carlo supera por casi 3 veces la establecida en las condiciones de aceptabilidad. Establecer un comportamiento menos variante de T_{med} trae como consecuencia un considerable aumento del área para la mayoría de los transistores que conforman el bloque PWL , cuestionando ampliamente la posibilidad de fabricación de la fuente de referencia.

En este trabajo se obtiene una mejora en desempeño del diseño inicial de Mateus por

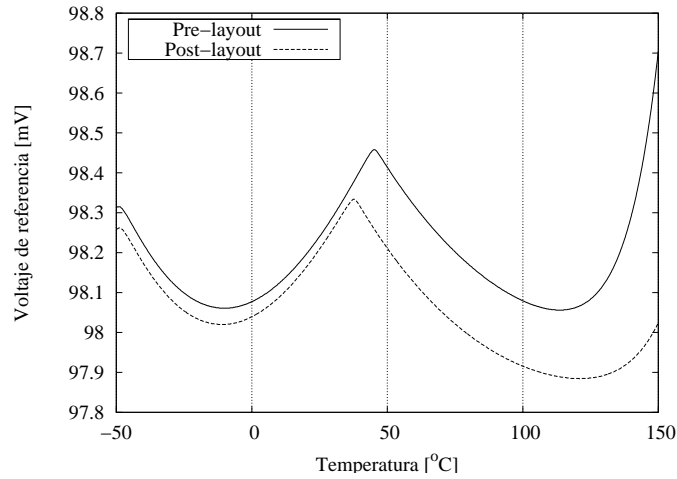


Figura 4.6: Voltaje de referencia *vs.* temperatura.

medio de la optimización vía programación geométrica. A pesar de que varios de los valores de desempeño reportados en [1] superen a los establecidos por este trabajo, el análisis estadístico estableció que a nivel de muestras el diseño propuesto aquí fue mejor que el de Mateus. La validación estadística de un circuito no solo permite evidenciar su variabilidad, sino el punto al cual corresponde la muestra característica del lote de fabricación y que representa la media de la distribución.

La gráfica del voltaje de referencia de la fuente tanto *pre-layout* como *post-layout* es mostrada en la figura 4.6. A pesar que el comportamiento altamente simétrico de la gráfica *pre-layout* con respecto a T_{med} , los resultados *post-layout* muestran una mejora en el coeficiente térmico. De la misma manera se destaca el desplazamiento de la temperatura media, en conjunto con una ligera pérdida de simetría. A nivel de variables de desempeño, la tabla 4.15 reporta una disminución de cerca de $10[dB]$ en la regulación de línea de la fuente para el caso de la validación en *layout*, el cual es atribuido ampliamente a las capacitancias parásitas inherentes a este, y que dado el aumento del área para control de variabilidad degradan enormemente la respuesta en frecuencia de la fuente de referencia.

Los histogramas de la tensión de referencia y del coeficiente térmico se muestran en la figura 4.7(a) y 4.7(b) respectivamente. El porcentaje de dispositivos que no cumplen con la varianza estimada por medio del análisis de Monte Carlo es de 0.25 %, representando un rendimiento para este caso del 99.75 %. A pesar de que el voltaje de referencia fluctuó cerca de $4mV$, el coeficiente térmico muestra un adecuado comportamiento, garantizando que algo más de la mitad de circuitos fabricados tengan un coeficiente térmico por debajo de $3,5\mu V/^{\circ}C$. Esto demuestra que las variaciones del voltaje de referencia no alteran significativamente la compensación térmica del circuito. En contraste, la significativa varianza exhibida por el voltaje de referencia no le permite garantizar al fabricante el valor de la tensión de salida con intervalos de confianza estrechos, originando diferencias sistemáticas en los puntos de

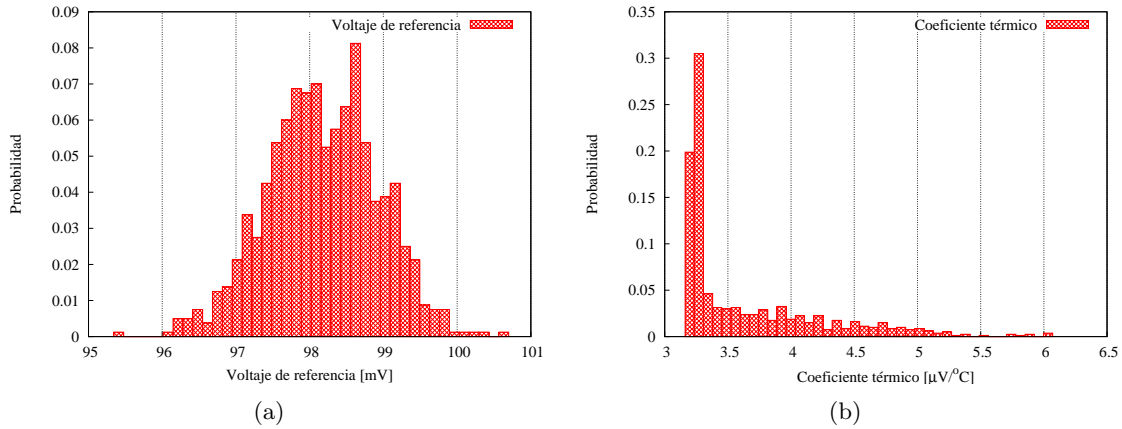


Figura 4.7: (a) Histograma para la tensión V_{ref} . (b) Histograma del coeficiente térmico.

polarización del circuito al cual esta fuente va a ser aplicada.

Finalmente, en la figura 4.8 se observa el *layout* implementado del circuito diseñado con la estrategia *DFM* propuesta en este trabajo. Las dimensiones del *layout* son de $411\mu\text{m} \times 393\mu\text{m}$, representando un aumento neto de área a nivel de *layout* de 4,16 con respecto a Mateus. Es importante notar el área que ocupan los transistores M_{17} y M_N , los cuales pueden ser vistos en la parte inferior derecha del *layout*. Los requerimientos de área fueron indispensables para lograr la obtención de las desviaciones de la temperatura media consignadas en la tabla 4.15.

4.9. Conclusiones

- Se plantea una metodología de diseño para manufactura aplicable a múltiples contextos de diseño y basada en una fundamentación clara, sencilla, y práctica. Esta se concentra en establecer los compromisos de diseño a nivel de desempeño y de rendimiento, excluyéndose de hacer rigurosas formulaciones matemáticas que tienden a confundir los propósitos de diseño. De esta manera el planteamiento expuesto en este trabajo es validado con el re-dimensionamiento de una fuente de voltaje de referencia en sub-umbral, en donde se evidencia una mejora en el rendimiento con respecto a la fuente de referencia de Mateus.
- Se analizaron los modelos de *mismatch* más relevantes para el diseño de circuitos analógicos basados en el establecimiento de cinco premisas esenciales. De esta manera, el uso del modelo Pelgrom en el diseño es tan sólo justificado para fases previas al dimensionamiento del circuito dada la sencillez en su formulación. De otro lado, se evidencia una carencia de modelos de *mismatch*, los cuales además de estimar acertadamente la variabilidad en procesos sub-micrón deben generar pautas de diseño correctas y concretas en el marco de la metodología *DFM*.

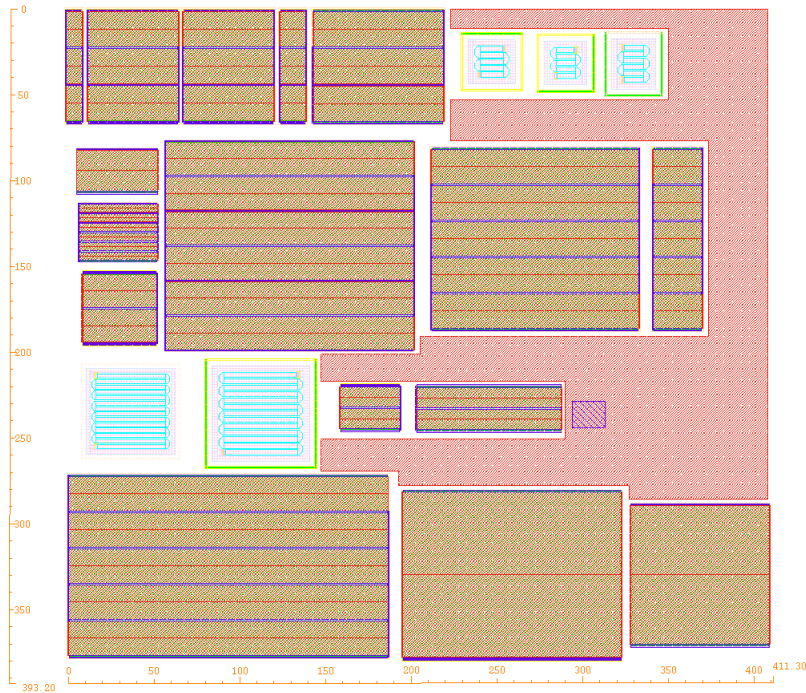


Figura 4.8: *Layout* de la fuente de referencia diseñada.

- Se reporta por primera vez un estudio de factibilidad del modelado en PG de las funciones comportamentales del transistor MOS polarizado en sub-umbral y bajo el modelo BSIM3v3, el cual establece una factibilidad para las variables: V_{GS} , V_{DS} , W y L . Paralelamente, se evidencia un comportamiento altamente no-convexo con la temperatura y la tensión fuente-cuerpo V_{SB} , que guía a este trabajo al modelado sobre intervalos estrechos de temperatura y de tensión V_{SB} . A pesar de esto, una fuente de referencia con un bajo coeficiente térmico es diseñada vía programación geométrica.
- La validación del desempeño en el ámbito de *DFM* debe ser siempre evaluada tomando las muestras características del lote de fabricación, las cuales a nivel de simulación corresponden a los valores de medias estimados bajo el análisis de Monte Carlo. De esta manera, la fuente de referencia de Mateus muestra un desempeño más bajo al expuesto en [1], revelando problemas de variabilidad inherentes a la región de polarización de sub-umbral.

4.10. Observaciones

- El análisis de sensibilidad de los parámetros de desempeño al no estar disponible en las herramientas de diseño, debe ser implementado de manera apropiada. Esta implementación parte de un correcto ajuste del simulador a usar, y debe estar en consonancia con el compromiso ligado al error de discretización de las ecuaciones diferenciales y al

mínimo error de convergencia determinado por el simulador.

- Se evidenció una tendencia de sobre-estimación del área guiada por la función de desempeño del *PSRR*. Además de estar influenciada por el error debido a estimación del simulador, el *PSRR* de manera más evidente que la regulación de línea, es dependiente de los polos de la salida, los cuales son altamente sensibles a las variables de diseño. Dado a que la aceptabilidad de rendimiento de este parámetro no es crucial en el funcionamiento global del circuito, sus condiciones pueden ser definidas con alto grado de distensión.
- La región de aproximación del espacio de diseño planteada por este trabajo a pesar de no tener una forma funcional conocida ni un centro establecido, permite efectuar un proceso de centralización de diseño enmarcado en la aplicación de la programación geométrica. Se prescinde de un estudio de convexidad dado a que los planos y superficies que la conforman son por si mismos conjuntos convexos bien definidos.
- Existe un compromiso entre el valor de los resistores tanto de *PTAT*, *CTAT* y de la rama final con el consumo de potencia global del circuito. Si el esquema de optimización busca reducir los valores de los resistores, se evidenciará un aumento de potencia en el circuito. De manera contraria, la disminución de potencia origina un aumento en el área de los resistores fomentado por el incremento en su valor. En la mayoría de las aplicaciones se propende por encontrar óptimos a nivel de potencia. Pero para este caso donde la potencia consumida es del orden de los μW , las exigencias de diseño tienden a establecer valores bajos de resistores con el fin de aumentar la regulación de línea y el *PSRR*.

4.11. Recomendaciones para trabajos futuros

- Tanto la aproximación del espacio de diseño como la estimación del área mínima para *mismatch* no tienen en cuenta las posibles correlaciones entre las variables de desempeño con cada una de las variables de dimensionamiento. Para evitar sobre-estimaciones de variabilidad y por ende de dimensionamiento, futuras implementaciones de esta metodología deberán incluir las dependencias entre las dimensiones de los elementos con cada uno de sus parámetros eléctricos y variables de desempeño globales del circuito.
- La fundamentación de esta metodología fue llevada a cabo completamente bajo un modelo comportamental de transistor MOS no continuo. Para extender las expectativas, modelos, métodos y medios aquí planteados, se recomienda la implementación en todo el proceso de diseño de un modelo totalmente continuo.

- Al descubrimiento del espacio de diseño puede ir ligada la implementación de funciones de automatización que permitan establecer rangos de polarización (V_{GS} , I_D , V_{DS} , k , etc), de manera tal que el modelado se vea definido en el marco de la funcionalidad del circuito. Además, debe ser implementado un programa geométrico que permita medir indistintamente valores de desempeño y a la vez analizar la viabilidad en la consecución de un diseño a través de un análisis de factibilidad de PG.
- Dado a que gran parte de la posible fabricación de un circuito integrado analógico depende del proceso que involucra la implementación a nivel de *layout*, se requiere que futuras metodologías tengan presentes desde sus fases tempranas los problemas ligados a la ubicación, el *mismatch* y los compromisos entre área y los componentes parásitos del *layout*.

Bibliografía

- [1] J. C. M. Ardila, “Diseño de una Fuente de Referencia Bandgap Integrada en tecnología CMOS,” Proyecto de pregrado de la Universidad Industrial de Santander, Octubre 2007.
- [2] A. J. Strojwas, “Design for Manufacturability and Yield,” in *26th ACM/IEEE Design Automation Conference*. Carnegie Mellon University, 1989, pp. 454–459.
- [3] N. Jakatdar, “DFM: What, Why, When & How,” Solid State Technology and Devices Seminar EECS 298-12, UC at Berkeley, February 2007.
- [4] B. P. Wong, A. Mittal, Y. Cao, and G. Starr, *Nano CMOS Circuit and Physical Design*, 1st ed. John Wiley & Sons, Inc., 2005.
- [5] M. del Mar Hershenson, “CMOS Analog Circuit Design Via Geometric Programming,” Ph.D. dissertation, Department of Electrical Engineering of Stanford University, November 1999.
- [6] J. Oliveros, D. Cabrera, E. Roa, and W. V. Noije, “An Improved and Automated Design Tool for the Optimization of CMOS OTAs Using Geometric Programming,” in *Symposium on Integrated Circuits and System Design - SBCCI*, September 2008.
- [7] W. Carvajal, E. Roa, and W. V. Noije, “A 23 MHz GBW 460 uW Folded Cascode OTA for a Sample and Hold Circuit Using Double Sampling Technique,” in *Conference on Integrated Circuits and Systems, DCIS*, November 2008.
- [8] C. Galup-Montoro, M. Schneider, H. Klimach, and A. Arnaud, “A compact model of MOSFET mismatch for circuit design,” *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 8, pp. 1649–1657, Aug. 2005.
- [9] K. P. White, W. J. Trybula, and R.Ñ. Athay, “Design for Semiconductor Manufacturing-Perspective,” in *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 20, no. 1, January 1997, pp. 58–72.
- [10] S. S. Sapatnekar, P. M. Vaidya, and S.-M. Kang, “Convexity-Based Algorithms for Design Centering,” vol. 13, no. 12, pp. 1536–1549, December 1994.

-
- [11] M. Pelgrom, A. Duinmaijer, and A. Welbers, "Matching properties of MOS transistors," *Solid-State Circuits, IEEE Journal of*, vol. 24, no. 5, pp. 1433–1439, Oct 1989.
- [12] C. Galup-Montoro, M. Schneider, A. Cunha, F. de Sousa, H. Klimach, and O. Siebel, "The Advanced Compact MOSFET (ACM) Model for Circuit Analysis and Design," *Custom Integrated Circuits Conference, 2007. CICC '07. IEEE*, pp. 519–526, Sept. 2007.
- [13] M. C. Bernardo, R. Buck, L. Liu, and W. A. Nazaret, "Integrated Circuit Design Optimization Using a Sequential Strategy," in *IEEE Transactions on Computer-Aided Design*, vol. II, no. 3, March 1992.
- [14] G. Taguchi, *Introduction To Quality Engineering*. Asian Productivity Organization, 1986.
- [15] D. Ñam, Y. D. Seo, L. Park, C. H. Park, and B. Kim, "Parameter Optimization of an On-Chip Voltage Reference Circuit Using Evolutionary Programming," in *IEEE Transactions on Evolutionary Computation*, vol. V, no. 4, August 2001.
- [16] A. Tajalli, M. Atarodi, A. Khodaverdi, and F. S. Esfanjani, "Design and Optimization of a High PSSR CMOS Bandgap Voltage Reference," in *ISCAS*, 2004.
- [17] B. R. Gregoire, "Optimum Area Allocation for Minimum Mismatch," in *IEEE Custom Integrated Circuits Conference*, 2004, pp. 643–646.
- [18] C. Roma, P. Daglio, G. D. Sandre, M. Pasotti, and M. Poles, "How Circuit Analysis and Yield Optimization Can Be Used To Detect Circuit Limitations Before Silicon Results," in *Proceedings of the 6th International Symposium on Quality Electronic Design*, 2005.
- [19] T. Serrano-Gotarredona, B. Linares-Barranco, and J. Velarde-Ramirez, "A precise CMOS mismatch model for analog design from weak to strong inversion," *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on*, vol. 1, pp. I-753–6 Vol.1, May 2004.
- [20] K. Lakshmikumar, R. Hadaway, and M. Copeland, "Characterisation and modeling of mismatch in MOS transistors for precision analog design," *Solid-State Circuits, IEEE Journal of*, vol. 21, no. 6, pp. 1057–1066, Dec 1986.
- [21] H. Yang, V. Macary, J. Huber, W.-G. Min, B. Baird, and J. Zuo, "Current mismatch due to local dopant fluctuations in MOSFET channel," *Electron Devices, IEEE Transactions on*, vol. 50, no. 11, pp. 2248–2254, Nov. 2003.
- [22] J. Croon, M. Rosmeulen, S. Decoutere, W. Sansen, and H. Maes, "An easy-to-use mismatch model for the MOS transistor," *Solid-State Circuits, IEEE Journal of*, vol. 37, no. 8, pp. 1056–1064, Aug 2002.

-
- [23] P. Drennan and C. McAndrew, “Understanding MOSFET mismatch for analog design,” *Solid-State Circuits, IEEE Journal of*, vol. 38, no. 3, pp. 450–456, Mar 2003.
- [24] O. G. Filho, “Um Modelo Compacto do Transistor MOS para Simulaco de Circuitos,” Ph.D. dissertation, UFSC, Florianópolis, Sep. 1999.
- [25] R. Coitinho, L. Spiller, M. Schneider, and C. Galup-Montoro, “A simplified methodology for the extraction of the ACM MOST model parameters,” 2001, pp. 136–141.
- [26] P. Kinget, “Device mismatch and tradeoffs in the design of analog circuits,” *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 6, pp. 1212–1224, June 2005.
- [27] S. Director and G. Hachtel, “The simplicial approximation approach to design centering,” *Circuits and Systems, IEEE Transactions on*, vol. 24, no. 7, pp. 363–372, Jul 1977.
- [28] K. Low and S. Director, “A new methodology for the design centering of IC fabrication processes,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 10, no. 7, pp. 895–903, Jul 1991.
- [29] A. Seifi, K. Ponnambalam, and J. Vlach, “A unified approach to statistical design centering of integrated circuits with correlated parameters,” *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 46, no. 1, pp. 190–196, Jan 1999.
- [30] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming,” Internet, February 2009, build 711. [Internet]. Visite: <http://stanford.edu/~boyd/cvx>