



**INAOE**

# **Metodología de Selección de Caminos Críticos para Circuitos Nanométricos CMOS Tolerantes a Envejecimiento Debido a NBTI**

por:

**Ing. Andrés Felipe Gómez Chacón**

Tesis sometida como requisito parcial para obtener el grado de:

**MAESTRO EN CIENCIAS EN LA  
ESPECIALIDAD DE ELECTRÓNICA**

en el

Instituto Nacional de Astrofísica, Óptica y Electrónica

Julio 2013

Tonantzintla, Puebla

Director de tesis:

**Dr. Víctor Hugo Champac Vilela, INAOE**

©INAOE 2013

Derechos reservados

El autor otorga al INAOE el permiso de reproducir y distribuir  
copias de esta tesis en su totalidad o en partes.



# Resumen

El continuo escalado tecnológico de los dispositivos MOSFET ha permitido que los circuitos integrados VLSI alcancen una mayor densidad de integración y velocidad de operación, lo cual ha llevado a un incremento en la demanda de circuitos electrónicos, especialmente los circuitos digitales. Sin embargo, sobre el régimen de dimensiones nanométricas el impacto de algunos efectos que degradan el rendimiento del dispositivo es más significativo, disminuyendo su tiempo de vida útil. Los principales mecanismos de envejecimiento que afectan el rendimiento son el NBTI, HCI y TDDBD, siendo el NBTI identificado como el factor predominante de envejecimiento en transistores PMOS. El efecto que tiene el NBTI es el de aumentar la magnitud del voltaje umbral del transistor a través del tiempo, dependiendo de las condiciones de operación del dispositivo como temperatura y tensión de alimentación. El incremento del voltaje umbral ocasiona que el retardo de propagación de los caminos lógicos de señal en un circuito digital se incremente. Si este retardo sobrepasa las restricciones de temporización se presentará una falla en el sistema que en algunas aplicaciones como en el área automotriz o aeroespacial son inaceptables. Para tolerar el incremento del retardo de propagación debido al envejecimiento de los dispositivos, convencionalmente se introduce un tiempo adicional en el período de reloj que funciona como margen de seguridad (también llamado *banda de guarda*). Sin embargo, introducir márgenes de seguridad en el período de reloj no permite obtener un óptimo rendimiento del circuito durante su tiempo de vida inicial. Otra alternativa consiste en utilizar circuitos

---

adaptivos, que introducen sensores de envejecimiento en los caminos lógicos del circuito que podrían generar fallas en el sistema (caminos críticos del circuito) y de esta forma sintonizan o controlan las condiciones de operación, como el voltaje de alimentación o la frecuencia de operación, según el grado de envejecimiento del circuito. La identificación de estos caminos debe realizarse de forma estadística debido a los efectos de variaciones de proceso y es de gran importancia seleccionar adecuadamente sólo aquellos paths que realmente requieran ser monitoreados. Además, identificar a alto nivel cuales paths potencialmente deberán ser monitoreados permitiría reducir el esfuerzo para introducir los sensores requeridos, así como tomar decisiones en etapas tempranas del flujo de diseño.

En esta tesis de maestría, se presenta una metodología para la selección del conjunto de paths a monitorear en circuitos adaptivos considerando efectos de envejecimiento por NBTI y variaciones de proceso en los parámetros de fabricación. Además, la metodología parte de la descripción a alto nivel del circuito para estimar o predecir un primer conjunto de paths y de esta forma disminuir las modificaciones requeridas para insertar los sensores en etapas posteriores. Algunos paths que deben monitorearse, no son detectados mediante el análisis estimado. Entonces, se plantea tolerar el envejecimiento de éstos insertando una banda de guarda en el período de reloj, la cual será menor a la utilizada en un circuito convencional.

# Agradecimientos

**Al Concejo Nacional de Ciencia y Tecnología (CONACYT)**

Por el apoyo económico otorgado a través de la beca para estudios de maestría.

(No. de Registro: 420129)

**Al Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE)**

por abrirme las puertas para realizar los estudios de posgrado.

**A mi asesor: Dr. Victor Hugo Champac Vilela**

por el conocimiento y formación académica que he recibido de su parte.

**A los Doctores:**

Dr. Arturo Sarmiento Reyes

Dr. Luis Hernández Martínez

Dr. Alejandro Díaz Mendez

Por haberme apoyado con la revisión de esta tesis y por retroalimentarme con sus puntos de vista y opiniones.

**A los trabajadores del INAOE:**

Por el soporte y ayuda que prestan a los estudiantes.

*A Leidy Paola,  
Por tu apoyo incondicional  
en la distancia.*

# Contenido

Resumen	I
Agradecimientos	III
Lista de Acrónimos	IX
Prefacio	x
<b>1. Introducción</b>	<b>1</b>
1.1. Confiabilidad en circuitos integrados digitales . . . . .	2
1.1.1. Confiabilidad vs Yield . . . . .	2
1.1.2. Curva de tina de baño e impacto del escalado tecnológico .	3
1.2. Mecanismos que afectan la confiabilidad . . . . .	5
1.2.1. <i>Negative Bias Temperature Instability</i> (NBTI) . . . . .	6
1.2.2. <i>Hot Carriers Injection</i> (HCI) . . . . .	6
1.2.3. <i>Time Dependent Dielectric Break-down</i> (TDDBD) . . . . .	7
1.3. Banda de guarda . . . . .	8
1.4. Justificación de la tesis . . . . .	9
1.5. Organización de la tesis . . . . .	11
<b>2. Envejecimiento debido a NBTI y su impacto en tecnologías nanométricas</b>	<b>13</b>
2.1. Fundamentos . . . . .	14

---

2.2. Modelo predictivo de degradación del $V_{th}$ debido a NBTI . . . . .	20
2.3. Impacto en el rendimiento de un circuito digital . . . . .	23
<b>3. Técnicas para reducir el impacto del envejecimiento en circuitos integrados digitales</b>	<b>28</b>
3.1. Incremento de la movilidad efectiva de los huecos . . . . .	29
3.2. Disminución del voltaje de alimentación . . . . .	31
3.3. Control del vector de entrada . . . . .	31
3.4. Control de Nodos Internos . . . . .	32
3.5. Redimensionamiento de las compuertas . . . . .	33
3.6. Ordenar los nodos de entrada en las compuertas . . . . .	35
3.7. Circuitos Adaptivos . . . . .	37
3.7.1. Path réplica y sensores embebidos . . . . .	38
3.7.2. Predicción de error . . . . .	40
3.7.3. Detección y corrección de error . . . . .	41
<b>4. Estimación del retardo del circuito considerando efectos de NBTI y variaciones de proceso</b>	<b>46</b>
4.1. Consideraciones . . . . .	46
4.1.1. Variaciones en los parámetros . . . . .	46
4.1.2. Modelo de retardo de compuerta . . . . .	48
4.1.3. Correlación espacial . . . . .	51
4.2. Modelado del NBTI considerando variaciones de proceso . . . . .	53
4.3. Valor medio y varianza de retardo de una compuerta y covarianza de dos compuertas digitales . . . . .	57
4.4. Valor medio, varianza del retardo de un path y covarianza entre paths . . . . .	61
4.4.1. Correlación entre paths . . . . .	62

---

---

<b>5. Metodología de selección de caminos críticos a monitorear en un circuito adaptivo tolerante a envejecimiento por NBTI</b>	<b>64</b>
5.1. Detección de paths topológicos y cálculo de su retardo medio de propagación . . . . .	66
5.2. Consideración de máxima degradación en los paths debido a NBTI	67
5.3. Prefiltrado de paths . . . . .	70
5.4. Criterio de selección de paths a monitorear . . . . .	71
5.4.1. Variable de diferencia de retardos . . . . .	72
5.4.2. Efecto de la correlación estructural en la diferencia de retardos	73
5.4.3. Umbral de probabilidad . . . . .	75
5.4.4. Determinar redundancia de los paths a monitorear . . . . .	76
5.5. Estimación de varianza de DR sin información de layout para la predicción de paths a monitorear . . . . .	78
5.5.1. Correlación entre paths y Correlación global . . . . .	83
5.5.2. Efecto de la correlación espacial global en la desviación estándar de DR . . . . .	85
5.5.3. Efecto de la Profundidad del path bajo análisis . . . . .	86
5.5.4. Comportamiento de los términos de varianza de la diferencia de retardos . . . . .	89
5.6. Cobertura de los paths no detectados mediante uso de la banda de guarda . . . . .	90
5.7. Resultados en circuitos ISCAS combinacionales . . . . .	92
5.7.1. Predicción de paths . . . . .	92
5.7.2. Banda de guarda para tolerar paths no identificados . . . . .	95
5.7.3. Compromiso entre paths no identificados e identificados incorrectamente . . . . .	95
<b>6. Conclusión y trabajo futuro</b>	<b>99</b>
<b>Lista de figuras</b>	<b>104</b>

---



Lista de tablas	105
Bibliografía	106

# Lista de Acrónimos

**NBTI** *Negative Bias Temperature Instability*

**HCI** *Hot Carriers Injection*

**TDDBD** *Time Dependent Dielectric Break-down*

**PDF** *Probability Density Function*

**LTDM** *Long Term Degradation Model*

**RDF** *Random Dopant Fluctuation*

**AF** *Aging Factor*

**LCP** *Longet Critical Path*

**NCP** *Near Critical Paths*

**SSTA** *Statistical Static Timing Analysis*

**STA** *Static Timing Analysis*

# Prefacio

Debido a las condiciones de operación cada vez más extremas a las que están sometidos los dispositivos electrónicos, éstos tienden a desgastarse o envejecerse con el paso del tiempo. Como consecuencia, el tiempo de vida útil de un circuito electrónico es limitado.

Sin embargo, en algunas aplicaciones como lo son la industria automovilística y la aeroespacial, sus circuitos deben poder operar correctamente por largos períodos de tiempo, por ejemplo 10 años. Este tipo de aplicaciones requieren de una alta confiabilidad por parte de los dispositivos.

El fenómeno de *Negative Bias Temperature Instability* es uno de los mecanismos que más afectan la confiabilidad, especialmente en los transistores PMOS. Este fenómeno tiene como consecuencia un incremento del voltaje umbral del transistor y a su vez un incremento en el retardo de propagación de los caminos lógicos de un circuito digital.

El uso de alternativas, como lo son los circuitos adaptivos, para tolerar el envejecimiento por NBTI, resulta más llamativo que utilizar márgenes de seguridad para garantizar la correcta operación del circuito durante su tiempo de vida útil, ya que permiten controlar las condiciones de operación (frecuencia o voltaje) según el incremento del retardo debido al envejecimiento, manteniendo un rendimiento óptimo durante todo el tiempo de vida del dispositivo.

Las estrategias adaptivas requieren el uso de sensores de retardo que ocupan área y consumen potencia. Por tanto, el conjunto de caminos que deben ser

monitoreados debe ser cuidadosamente seleccionado de manera que el impacto de estos sensores sea el menor. Mas aún, insertar los sensores luego de realizar un primer prototipo de layout implicaría hacer modificaciones en el mismo y sería un retroceso en el flujo de diseño convencional.

En este trabajo se analiza una metodología para la selección de los caminos críticos que deben ser monitoreados debido a envejecimiento por NBTI en un circuito adaptivo. Mediante el análisis del criterio de selección establecido se determina una estrategia para identificar a alto nivel un conjunto potencial de paths a monitorear y facilitar las modificaciones en etapas posteriores del flujo de diseño.

# Capítulo 1

## Introducción

Durante los últimos años, el desarrollo de productos más rápidos y complejos en el campo de la electrónica se ha basado en el escalado de las dimensiones de los dispositivos, alcanzando actualmente el régimen nanométrico y sub-nanométrico tal como lo indica la tendencia observada por “la ley de Moore”.

Sin embargo, a medida que se alcanzan los límites tecnológicos y físicos de los dispositivos, nuevos efectos deben ser considerados durante la etapa de diseño. En los circuitos digitales nanométricos, incluso las más pequeñas variaciones en el proceso de producción pueden dar lugar a variaciones de los parámetros eléctricos de los dispositivos que podrían ocasionar que un circuito fabricado sea inutilizable [1]. Igualmente, la degradación de los dispositivos con el tiempo, debido a las condiciones extremas de operación (altas temperaturas y campo eléctrico, conmutación constante, etc.), es más propensa y tiene un gran impacto en el rendimiento esperado, limitando así el tiempo de vida útil de un circuito. Considerar e investigar estos efectos permite desarrollar nuevas técnicas de diseño para asegurar el correcto funcionamiento de los circuitos digitales de alta complejidad manteniendo un rendimiento y tiempo de vida prudente, bajo ciertas condiciones físicas, tecnológicas y económicas.

## 1.1. Confiabilidad en circuitos integrados digitales

Es un hecho que los circuitos y sistemas digitales tienen un tiempo de vida finito antes de que una falla crítica en el sistema ocurra. Sin embargo, si este tipo de fallas se produce una vez superado el tiempo de vida esperado por el usuario, éste estará satisfecho y se habrán cumplido las expectativas planteadas al producto [2].

La confiabilidad (en inglés *Reliability*) es la habilidad de los dispositivos electrónicos de realizar las funciones para las que fueron diseñados cumpliendo con las especificaciones requeridas, bajo las condiciones de operación presentes, como voltaje de polarización y frecuencia de operación, durante un período de tiempo específico [3]. El propósito del estudio de la confiabilidad de los dispositivos es el de buscar alternativas que permitan asegurar que el tiempo de vida del sistema será mayor que el tiempo de uso requerido.

En el mundo de los dispositivos CMOS, el criterio de operación para algunos productos puede ser muy diferente. Algunos son utilizados raramente y por tanto requieren muy poco tiempo de vida útil. Por otra parte existen aplicaciones donde garantizar la confiabilidad es de vital importancia y a la vez es requerido un alto rendimiento por parte de los dispositivos electrónicos. Por ejemplo, en el campo de la industria automotriz o espacial, donde la electrónica digital es cada vez más común y los circuitos se encuentran operando constantemente bajo condiciones extremas, se espera un correcto funcionamiento libre de fallas durante todo el tiempo de vida útil, por ejemplo 10 años.

### 1.1.1. Confiabilidad vs Yield

Confiabilidad y *Yield* son quizás los aspectos más importantes para el desarrollo de una nueva tecnología [4]. El yield está definido como la probabilidad de

falla de un dispositivo recién fabricado ( $t = 0$ ), mientras que la confiabilidad puede ser vista como una falla funcional del dispositivo durante su tiempo de operación ( $t > 0$ ). Un proceso de fabricación con bajo yield es inaceptable para comenzar a trabajar, pero aún un proceso con un muy alto yield (baja probabilidad de obtener defectos iniciales) y baja confiabilidad no es sostenible económicamente a largo plazo.

En la industria, la confiabilidad de un producto debe ser conocida antes de que éste sea vendido, de esta forma los costos de garantía pueden ser medidos y la satisfacción del consumidor asegurada. Cumplir con estos objetivos, implica calcular el impacto que tendrá cada mecanismo de falla y predecir el tiempo en el cual su impacto comienza a ser significativo.

### 1.1.2. Curva de tina de baño e impacto del escalado tecnológico

Los errores o fallas que pueden ocurrir en un circuito digital, pueden clasificarse en cuatro tipos diferentes, dependiendo del tiempo determinado en el que ocurren. La figura 1.1 es conocida como *curva de tina de baño*. Esta curva es una medición del número de fallas por unidad de tiempo (tasa de fallas) versus el tiempo de operación y es una estimación del comportamiento de la confiabilidad de un sistema electrónico fabricado.

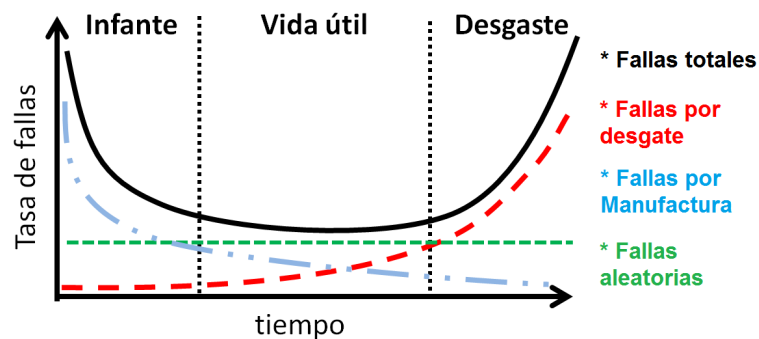


Figura 1.1: Tasa de Fallas representada en la Curva de tina de baño. [2]

- Fallas en el *yield*: Son defectos que se deben principalmente a las variaciones o imperfecciones en el proceso de manufactura que incluso en el tiempo cero de operación pueden ocasionar errores por mal funcionamiento del circuito. La gran mayoría de este tipo de defectos son detectados en la etapa de pruebas o *testing* realizado a los circuitos.
- Fallas infantiles: Son aquellos defectos que no fueron detectados en la etapa de test pero que el mal funcionamiento se presenta rápidamente. Comunmente un análisis de los defectos en edad infante revela estructuras extremadamente pobres, que de alguna forma sobrevivieron a los test iniciales de yield. A medida que el proceso de fabricación se madura para una nueva tecnología, la cantidad de estos defectos debe disminuir.

Una técnica para detectar este tipo de fallas es conocida como *burn-in* y consiste en aplicar condiciones de operación (como temperatura o voltaje) superiores a las nominales por un período de tiempo corto de manera que estos circuitos defectuosos se hagan evidentes y sean descartados. De esta forma, el consumidor recibirá sólo dispositivos que se encuentren en el tiempo útil de operación [2].

- Fallas durante el tiempo útil de operación: En esta etapa de operación, la tasa de fallas presentes es baja y debe estar contenidas dentro de las especificaciones establecidas para asegurar la viabilidad de la garantía establecida. Los mecanismos de falla presentes en esta etapa son de carácter aleatorio.
- Fallas por envejecimiento: A medida que pasa el tiempo, mayor cantidad de circuitos comienzan a ser defectuosos, estas fallas se deben al desgaste, degradación o envejecimiento de los dispositivos, es decir, a su confiabilidad. El momento en el que este tipo de fallas comienza a hacerse presente depende de la tecnología propia y las condiciones de operación como temperatura, el voltaje de alimentación o la frecuencia de uso del dispositivo,



entre otros factores. El objetivo de estudiar los mecanismos que afectan la confiabilidad es el de asegurar que las fallas por envejecimiento no estarán presentes hasta después de cumplir el tiempo de vida útil esperado. En la figura 1.2 se muestra un ejemplo donde luego de cierto tiempo de uso, debido a la degradación se alcanza el umbral máximo permitido de alguna especificación, pero este tiempo de uso es menor al requerido como garantía, lo cual representa una pérdida económica.

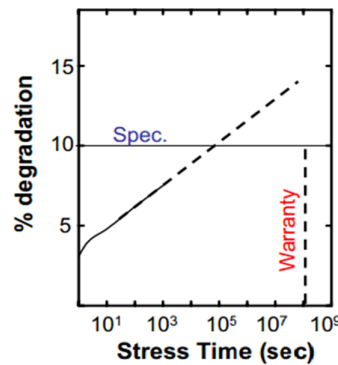


Figura 1.2: Degradación del rendimiento e incumplimiento del tiempo de vida requerido. [4]

A medida que se alcanzan los límites físicos y tecnológicos de las dimensiones en los dispositivos, tanto los defectos de manufactura como los fenómenos de degradación tienden a aumentar el número de fallas observadas y reducir la confiabilidad. Esto se observa en la tendencia marcada por las flechas en la figura 1.3.

## 1.2. Mecanismos que afectan la confiabilidad

Existen diferentes fenómenos físicos que afectan la confiabilidad de los dispositivos, ya sea debido a los materiales del sustrato y el óxido, así como las interfaces entre ellos. Los principales fenómenos de degradación que afectan la

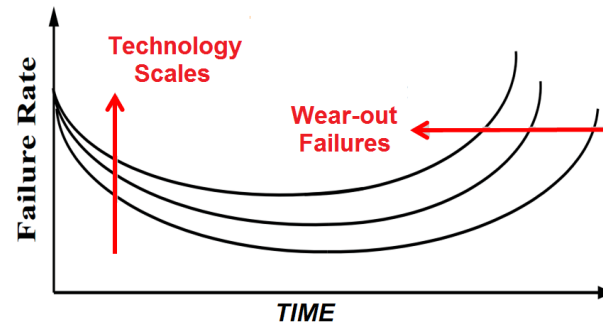


Figura 1.3: Tendencias en la curva de tina de baño

confiabilidad son NBTI, HCI y TDDDBD [4]. A continuación se mencionarán las principales características de cada uno de ellos.

### 1.2.1. NBTI

Negative Bias Temperature Instability, ha sido identificado como un fenómeno dominante de envejecimiento en tecnologías CMOS nanométricas, el cual afecta principalmente a los transistores PMOS. Los enlaces de hidrógeno y silicio en la interfaz, pueden ser fácilmente rotos debido al campo eléctrico perpendicular y el aumento de la temperatura, generando trampas o estados en la interface. Esto trae como consecuencia un incremento del voltaje umbral del dispositivo a lo largo del tiempo [9]. La condición de polarización propicia para este efecto es cuando la compuerta tiene un voltaje negativo respecto al drenaje y fuente.

En este trabajo, se hará énfasis en el fenómeno de NBTI y será explicado con mayor detalle más adelante.

### 1.2.2. HCI

Hot Carriers Injection es un fenómeno donde algunos portadores, que fluyen cerca de la región de estrangulamiento del canal, ganan energía significativa debido al alto campo eléctrico horizontal en esta zona.

En la región lineal,  $V_{ds}$  está distribuido a través del canal entero y el campo eléctrico está limitado a valores menores al campo crítico de saturación de la velocidad,  $V_{dsat}$ . A medida que  $V_{ds}$  se incrementa entrando en la región de saturación, el exceso de voltaje  $V_{ds} - V_{dsat}$  se distribuye en la región de estrangulamiento del canal cerca del drenador. En este punto, el campo eléctrico tiene una dependencia exponencial con la posición [2].

Como se muestra en la figura 1.4, estos portadores energéticos pueden perder su energía mediante impactos de ionización generando un par electrón–hueco y contribuyendo a la corriente de substrato. También pueden tener suficiente energía como para romper los enlaces entre silicio e hidrógeno en la interface, generando trampas en la interface. Igualmente pueden superar la barrera de potencial en la interface e inyectarse en el óxido de compuerta donde pueden ser atrapados o pueden generar defectos en el óxido.

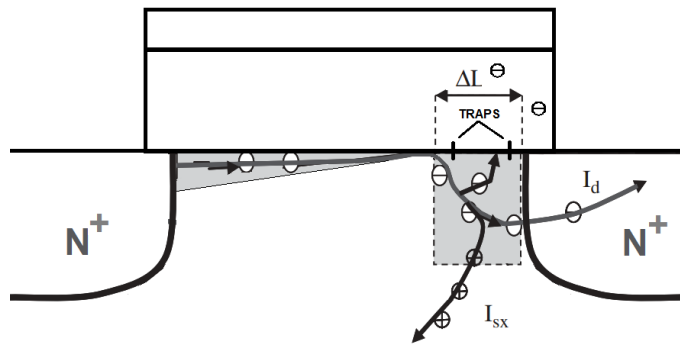


Figura 1.4: Mecanismos de Hot Carriers

### 1.2.3. TDDBD

El rompimiento del óxido de compuerta con el paso del tiempo se debe a que a éste es aplicado constantemente un campo eléctrico. Debido a efectos de tuneleo algunos huecos pueden quedar atrapados en el dieléctrico, mientras más carga es atrapada es posible formar un camino por el cual puede fluir corriente como se

muestra en la figura 1.5. Entonces, la compuerta y el canal estarían eléctricamente unidas, dañando completamente el concepto de funcionamiento.

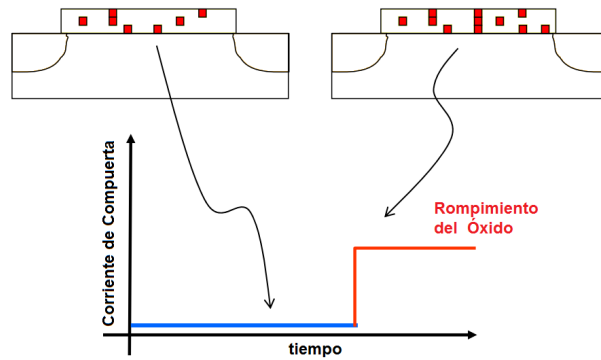


Figura 1.5: TDDBD

### 1.3. Banda de guarda

La mayoría de los circuitos digitales, se basan en la estructura mostrada en la figura 1.6. Un bloque con lógica combinacional recibe una o varias señales de entrada, realiza su operación en base a estas entradas y entrega el resultado a la siguiente etapa de lógica combinacional. El resultado de cada bloque es almacenado por Flip-Flops, los cuales mantienen el estado almacenado en un ciclo anterior hasta que el reloj indique el momento de almacenar la operación resultante. El tiempo en que el bloque combinacional tarda en recibir la señal de entrada, procesarla y entregar la función resultante al FF correspondiente determina el período de reloj necesario.

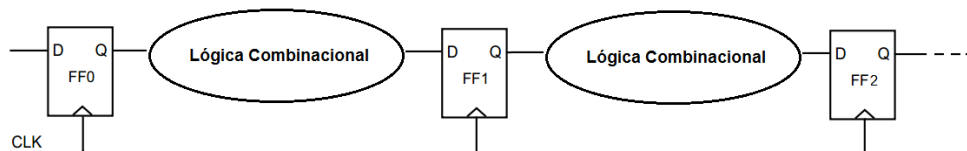


Figura 1.6: Estructura lógica de un circuito digital

Sin embargo, el período de reloj necesario depende de la variabilidad del retardo de propagación de la lógica combinacional, que pueden ser de diferente tipo: Estática como las variaciones de proceso, de degradación en el tiempo debido a los mecanismos de envejecimiento mencionados anteriormente o transitoria que son efectos temporales como caídas en alimentación, ruido por acoples capacitivos, radiación etc. Esta variabilidad impide estimar de forma determinista el retardo de propagación del bloque de lógica combinacional y el período de reloj necesario. Para asegurar que un período de reloj seleccionado, proporcione el suficiente tiempo de propagación a los caminos lógicos del circuito es necesario dejar un lapso de tiempo extra al período nominal, para que aún bajo el peor de los casos de estos efectos de variación de retardo, el circuito pueda operar correctamente. A este tiempo extra se le conoce como margen de seguridad o banda de guarda. La figura 1.7 ilustra esta idea, donde se ha dejado una banda de guarda (zona limitada por líneas verdes) exclusiva para tolerar el incremento del retardo del path crítico de un bloque de lógica combinacional debido al envejecimiento.

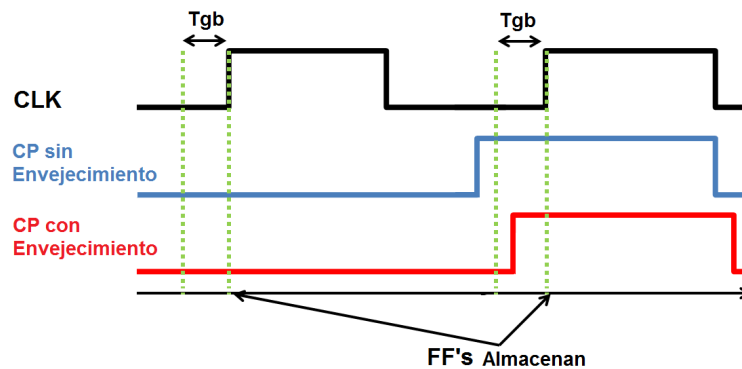


Figura 1.7: Margen de Seguridad para garantizar la correcta operación del circuito

## 1.4. Justificación de la tesis

Una de las especificaciones de diseño que recibe un mayor impacto por parte de los mecanismos de confiabilidad revisados anteriormente, es el período de op-

eración del reloj, ya que estos mecanismos generalmente se reflejan como un aumento en el tiempo de propagación de una señal digital a través de un path lógico en los boques combinatoriales. Las márgenes requeridas que se deben agregar al período de reloj para tolerar el efecto que tiene este tipo de mecanismos en el retardo de propagación y asegurar diseños robustos, pueden ser excesivamente grandes. Esto ha dado lugar a diseños conservativos con altas penalizaciones en consumo de potencia y rendimiento, limitando las ventajas obtenidas al avanzar a un nuevo nodo tecnológico.

Por otra parte, han surgido técnicas de diseños adaptivos [39] [34] [31] [30][29], los cuales sintonizan los parámetros del sistema (Frecuencia o Voltaje) según las condiciones de degradación de los dispositivos y permiten mitigar el impacto que éste tipo de variaciones en el tiempo tienen sobre el margen de seguridad requerido, obteniendo un rendimiento mayor y aprovechando al máximo los dispositivos cuando estos aún no se han desgastado. En este tipo de técnicas se utilizan sensores de retardo que monitorean periódicamente las transiciones de los paths lógicos que podrían ocasionar una falla funcional, para así detectar cuando un circuito a comenzado a desgastarse. Insertar los sensores implica consumo de área y potencia, por lo tanto, es de gran importancia identificar y sensar únicamente los caminos estrictamente necesarios y activar los sensores únicamente cuando éstos caminos se activen para evitar los mismos efectos de degradación en ellos.

Debido a los efectos de variaciones de proceso y a la interacción que tienen con el envejecimiento por NBTI, determinar cuales paths deben ser monitoreados en un circuito adaptivo se convierte en un problema estadístico donde conocer los valores de correlación espacial es posible sólo despues de realizar el primer prototipo de layout en el circuito y para insertar los sensores de retardo necesarios se debe realizar una modificación de éste, lo que implica un retroceso en el flujo de diseño convencional y un esfuerzo extra por parte de los ingenieros de layout.

Entonces, resultaría útil desarrollar metodologías de selección de caminos críticos que tengan en cuenta los efectos de confiabilidad a un alto nivel (antes de realizar

el layout) y permitan estimar o predecir el conjunto de paths que deberían ser monitoreados, ya que de esta forma es posible tomar decisiones sobre posibles soluciones en este nivel de diseño, reduciendo el esfuerzo y el costo requerido en etapas posteriores. En base a esto, una metodología que parte de alto nivel para estimar el conjunto de paths a monitorear en un circuito adaptivo considerando efectos de variaciones de proceso y envejecimiento, es propuesta.

## 1.5. Organización de la tesis

En esta introducción se ha dado una primera perspectiva de la importancia de la confiabilidad en los circuitos integrados digitales, los principales fenómenos que afectan la confiabilidad y el impacto de introducir márgenes de seguridad para tolerar este tipo de variaciones en el retardo de propagación de los caminos críticos de un circuito.

En el capítulo 2 se hará un mayor enfoque en el mecanismo de envejecimiento analizado en esta tesis, el cual es el NBTI. Se revisará la física de este fenómeno, el modelo matemático que lo describe y su impacto sobre el rendimiento de un circuito digital.

En el capítulo 3 se hace una revisión de las principales técnicas para reducir el impacto del envejecimiento por NBTI, bajo consideraciones de diseño y a nivel adaptivo.

En el capítulo 4, se plantean las bases para realizar un análisis estadístico de tiempos considerando el efecto del NBTI. En este capítulo se presenta el modelo lineal de compuerta, a partir del cual se desarrolla la metodología para el cálculo del valor medio y varianza del retardo de un path, así como la covarianza y correlación entre paths.

En el capítulo 5, se explicará la propuesta de una metodología de selección de caminos críticos a monitorear en circuitos adaptivos CMOS tolerantes a envejecimiento por NBTI. En este capítulo también se incluyen los resultados obtenidos.

Finalmente, en el capítulo 6 se exponen las conclusiones del trabajo de tesis y el trabajo futuro.



## Capítulo 2

# Envejecimiento debido a NBTI y su impacto en tecnologías nanométricas

En los transistores PMOS ha sido indentificado que su voltaje umbral cambia gradualmente a través del tiempo. Este cambio es ocasionado por el estrés que se genera en el óxido de compuerta debido al voltaje aplicado y su efecto se intensifica con la temperatura. La figura 2.1 muestra la condición de polarización donde el impacto de éste fenómeno es significativo, a esta condición se le llama *condición de stress*. La compuerta tiene una tensión negativa respecto al dreno y fuente de un transistor PMOS simultáneamente. Este efecto es conocido como *Negative Bias Temperature Instability* o NBTI.

En el diseño VLSI, una precisa predicción de la degradación del rendimiento debido a NBTI es muy importante, ya que permite a los diseñadores determinar posibles soluciones para garantizar el tiempo de vida útil especificado.

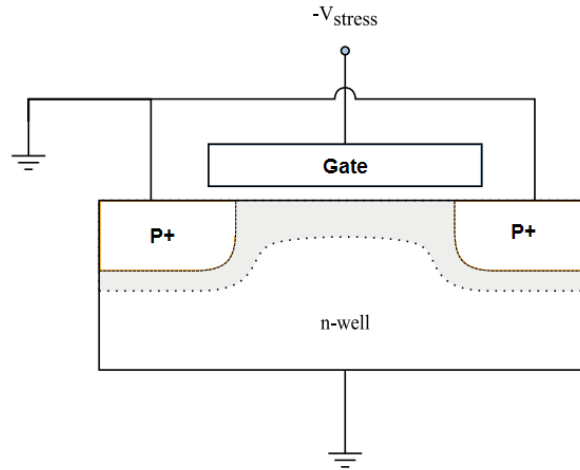


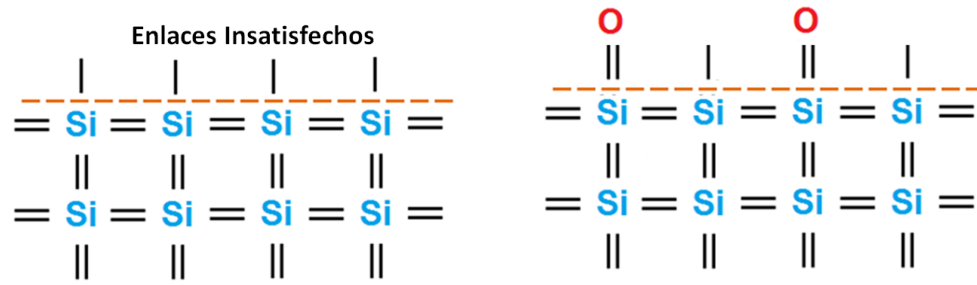
Figura 2.1: Condición de polarización donde el efecto de NBTI se presenta

## 2.1. Fundamentos

El fenómeno de NBTI fue observado por primera vez sobre los años 60, donde los transistores PMOS eran los dispositivos dominantes en los procesos de fabricación [4]. El efecto se genera debido a los enlaces insatisfechos (*dangling bonds*) en la interface  $Si/SiO_2$  que actúan como estados o trampas de huecos en la zona prohibida.

El átomo de silicio posee cuatro electrones de valencia y por tanto requiere también cuatro enlaces para completar su banda de valencia. Dentro de la estructura cristalina, cada átomo de silicio completa su banda de valencia estableciendo estos cuatro enlaces con sus átomos vecinos. Sin embargo, en la superficie del cristal de silicio no todos los enlaces pueden completarse como se muestra en la figura 2.2(a).

Luego del proceso de oxidación, gran parte de estos enlaces son completados con los átomos de oxígeno (figura 2.2(b)). Sin embargo, aún varios enlaces han quedado insatisfechos, para lo cual se realiza un proceso llamado *pasivación* con átomos de hidrógeno, formando enlaces  $Si-H$  como se muestra en la figura 2.2(c), mejorando de esta forma la calidad en la interface del dispositivo. Sin embargo,



(a) Enlaces insatisfechos en la superficie de silicio

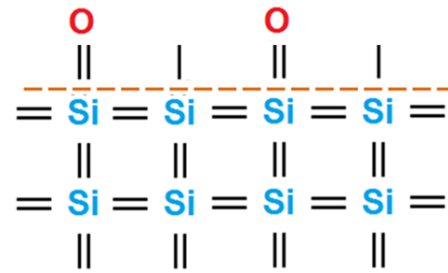
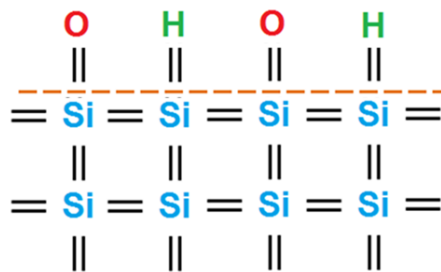
(b) Interface  $S_i - S_iO_2$  luego del proceso de oxidación(c) Interface  $S_i - S_iO_2$  luego del proceso de oxidación y pasivación

Figura 2.2: Interface del silicio y óxido de compuerta

son precisamente estos enlaces de hidrógeno y silicio los que causan el efecto del NBTI, ya que debido a su baja energía de acople pueden ser rotos fácilmente por los elevados campos eléctricos y la alta temperatura, reactivando las trampas en la interface [10].

Aunque el mecanismo exacto de generación de trampas en la interface es aún discutido, en [6] sugieren que algunos huecos en la capa de inversión llegan al óxido por mecanismos de tuneo, son atrapados e interactúan con los enlaces  $S_i - H$  que tienen baja energía de acople (enlaces débiles) de manera que por el alto campo eléctrico y temperatura pueden ser rotos fácilmente (ver figura 2.3).

Al romperse los enlaces  $S_i - H$ , los átomos de hidrógeno se difunden de la interface hacia el óxido de compuerta, dejando un enlace insatisfecho que actúa como trampa de huecos (figura 2.4). Una característica importante de la degradación

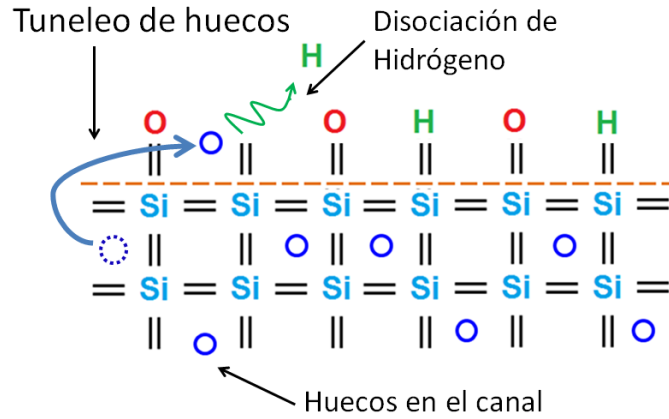


Figura 2.3: Tunelaje de huecos y rompimiento de enlaces  $S_i - H$  en la interfase (Reacción).

causada por NBTI, es la capacidad de recuperarse parcialmente cuando las condiciones de estrés (Campo eléctrico, temperatura, etc.) son disminuidas, ya que algunos enlaces  $S_i - H$  son reestablecidos como se ilustra en la figura 2.5.

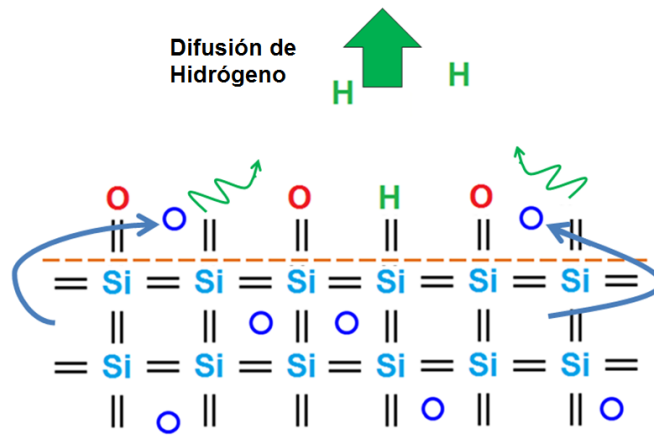


Figura 2.4: Difusión de los átomos de hidrógeno en el óxido (Difusión).

La figura 2.6 muestra la dependencia de la generación de trampas en la interfase durante el tiempo de estrés para varios valores de temperatura y campo eléctrico. Como se puede observar, el cambio en las trampas en la interfase incrementa con ambos parámetros y muestran una dependencia en forma de ley de potencia con el tiempo ( $t^n$ ). En la relación con la temperatura, es evidente que la

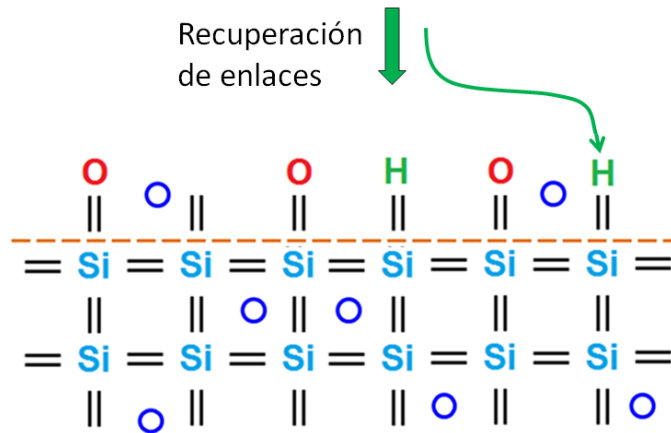


Figura 2.5: Pasivación de enlaces con átomos de hidrógeno que se recuperan (recuperación) .

tendencia cambia luego de cierto punto en el tiempo ( $t_{break}$ ) [6] a partir del cual la pendiente es mayor y que esta pendiente a su vez depende de la temperatura. Sin embargo, para el caso del campo eléctrico, tanto la pendiente como el punto de cambio de la tendencia son independientes del campo aplicado. El tiempo  $t_{break}$  está determinado por el momento en el cual el frente de difusión de Hidrógeno alcanza la interface entre el óxido de compuerta y el poly, ya que en éste la constante de difusión es mayor.

La generación de trampas en la interface en un transistor PMOS contribuyen con carga positiva a la banda de valencia, teniendo como consecuencia una desviación negativa en el voltaje umbral del transistor (se hace más negativo). La relación entre la densidad de trampas en la interface del silicio y óxido ( $N_{it}$ ) y el cambio en voltaje umbral puede representarse como:

$$\Delta V_{th} = -\frac{q}{C_{ox}} \Delta N_{it} \quad (2.1)$$

Entonces, la generación de trampas en la interface es proporcional con el voltaje umbral. La figura, 2.7 muestra la magnitud del cambio del voltaje umbral del transistor en función del tiempo, durante la etapa de estrés y recovery. Debe no-

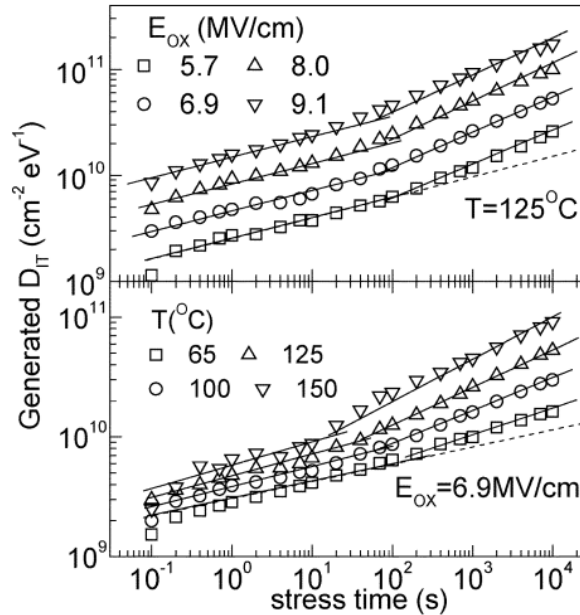


Figura 2.6: Dependencia del NBTI con el campo eléctrico y temperatura. [6]

tarse que aunque el tiempo dado a cada etapa es el mismo, la totalidad de los enlaces de  $Si-H$  no son reestalecidos. Por lo que el efecto general será aumento (en magnitud) del voltaje umbral. Si la proporción entre el tiempo en el que el dispositivo se encuentre en etapa de estrés aumenta respecto al tiempo en la etapa de recuperación, es de esperar que la desviación total del voltaje umbral sea aún mayor. Debe resaltarse de esta figura que el mayor incremento del voltaje umbral al comienzo de las etapas de estrés (recuperación) se debe a que el proceso de difusión del hidrógeno domina la generación (pasivación) de las trampas en la interface. Sin embargo, se llega a un momento en el cual la concentración de átomos de hidrógeno que pueden difundirse de la interface al óxido (del óxido a la interface) disminuye, haciendo que el cambio del voltaje umbral sea más lento [5].

Una característica importante del impacto del NBTI en el  $V_{th}$ , es que el aumento no depende de la frecuencia de conmutación (ver figura 2.8). Esto debido a que los mecanismos de difusión en la etapa de estrés son simétricos con respecto

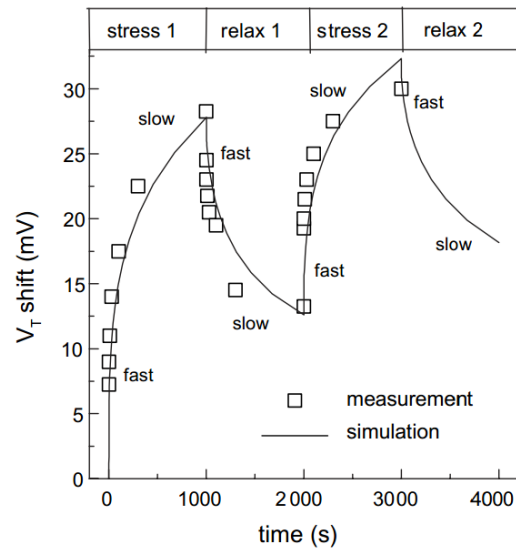


Figura 2.7: Desviación del voltaje umbral como consecuencia del NBTI. [5]

a los de la etapa de recuperación. La simetría de estos efectos se debe a que la cantidad de enlaces que son recuperados por la pasivación con el hidrógeno sólo dependen del mecanismo de difusión y no del campo eléctrico que sí es diferente entre una etapa y otra [7], esto debido a que el hidrógeno que se difunde tiene carga neutral. Para analizar la confiabilidad de un circuito digital, resultaría más realista considerar los efectos dinámicos del NBTI y no el peor caso de degradación estático.

Como consecuencia del incremento del voltaje umbral, la corriente de encendido de un transistor bajo una misma tensión de compuerta se disminuye. Esto causa una reducción del tiempo de propagación de las compuertas digitales, ya que la capacidad de carga de los nodos por parte de los transistores PMOS disminuye. Si el incremento en retardo es significativo, podría ocasionar un incumplimiento en las especificaciones de frecuencia de operación y un circuito dejaría de funcionar correctamente. Con el escalado tecnológico, el campo eléctrico que causa este fenómeno se ha incrementado, al igual que la temperatura debido al alto consumo de potencia, convirtiendo al NBTI en uno de los fenómenos de degradación que más afecta a los dispositivos.

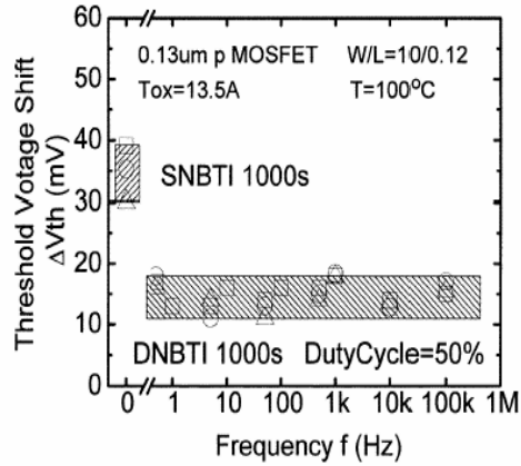


Figura 2.8: Dependencia del NBTI con la frecuencia de operación. [8]

## 2.2. Modelo predictivo de degradación del V<sub>th</sub> debido a NBTI

En [11] y [12] desarrollan un modelo para determinar un borde superior para la desviación de  $V_{th}$  debido a NBTI (como se muestra en la figura 2.9), basándose en el modelo físico de reacción y difusión[5] que describe matemáticamente los conceptos expuestos anteriormente. En este modelo, el cambio del voltaje umbral depende principalmente de las condiciones dinámicas de operación como el voltaje de alimentación ( $V_{dd}$ ), la temperatura ( $T$ ) y la probabilidad de estrés ( $\alpha$ ) que puede ser vista como un factor que determina la proporción del tiempo en el que el transistor será afectado por NBTI y el tiempo en el que se encontrará en estado de relajación ( $\alpha = t_{stress}/(t_{stress} + t_{rec})$ ). Esta aproximación resulta más realista que considerar que los dispositivos se encuentran todo el tiempo bajo degradación estática y permite realizar una mejor estimación del impacto del NBTI en el rendimiento de un circuito digital. Este modelo es llamado *Long Term Degradation Model* (LTDM). El LTDM describe el aumento del voltaje umbral debido al efecto de NBTI en un transistor PMOS por la ecuación 2.2.



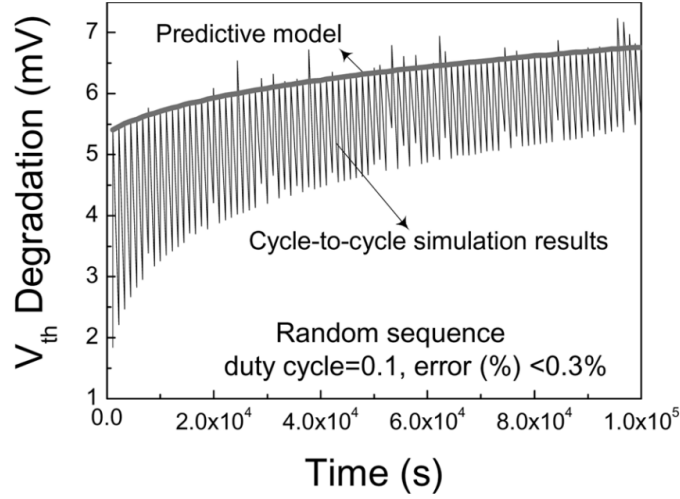


Figura 2.9: LTDM como borde superior de la degradación del  $V_{th}$  [13]

$$|\Delta V_{th_{NBTI}}| = \left( \frac{\sqrt{K_v^2 \alpha T_{clk}}}{1 - \beta_t^{1/2n}} \right)^{2n} \quad (2.2)$$

En esta ecuación,  $T_{clk}$ ,  $\alpha$  y  $n$  son el período de reloj, la probabilidad de estrés y una constante exponencial de tiempo con el valor de  $1/6$ , respectivamente.  $K_v$  describe la dependencia con el voltaje de alimentación  $V_{dd}$ , la temperatura  $T$ , el espesor del óxido  $t_{ox}$  y otros parámetros asociados con el efecto de NBTI [12].  $\beta_t$  refleja el impacto del tiempo de degradación.

En las figuras 2.10(a) y 2.10(b) se muestra el comportamiento de la degradación del  $V_{th}$  durante 10 años, obtenida mediante el LTDM para diferentes valores de probabilidad de estrés y de voltaje umbral inicial representado como una desviación debida a variaciones de proceso. Del comportamiento con respecto a la probabilidad de estrés, se puede observar que cuando la probabilidad es de cero (el transistor nunca está en condición de estrés), la degradación por NBTI lógicamente es nula, a medida que se aumenta la probabilidad, mayor es el cambio en el voltaje umbral. De esta figura también es posible intuir que el cambio del voltaje umbral es más sensible a cambios en la probabilidad cuando esta tiene un valor pequeño que cuando tiene uno alto. Esto se debe a que si el valor de  $\alpha$

## 2.2. MODELO PREDICTIVO DE DEGRADACIÓN DEL VTH DEBIDO A NBTI

es alto, el transistor estará mayor tiempo en etapa de estrés y debido a que la densidad de enlaces  $Si - H$  que pueden romperse disminuyen con el tiempo, el impacto de un incremento en la probabilidad no será significativo. Mientras que en el caso de una probabilidad baja, siempre hay gran cantidad de enlaces que pueden romperse cuando el transistor se encuentre en estrés.

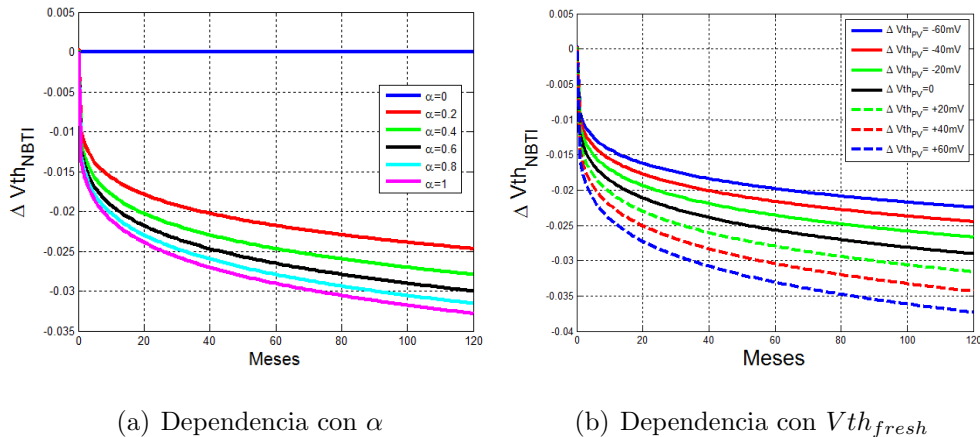


Figura 2.10: Curvas obtenidas al aplicar el LTDM

De la figura 2.10(b), teniendo en cuenta que el  $V_{th}$  de un transistor PMOS es negativo, es importante resaltar que transistores con valores de  $V_{th}$  iniciales mayores en magnitud ( $\Delta V_{th\_PV} < 0$ ) tendrán una menor degradación debido a NBTI. Como consecuencia, la diferencia inicial entre los voltajes de umbral de dos dispositivos disminuye con el tiempo gracias al envejecimiento. Debido a esta dependencia con el voltaje umbral inicial, el envejecimiento del circuito interactúa con las variaciones de proceso estáticas y se convierte en un fenómeno de carácter estadístico. En el capítulo 4 se expone un modelo que tiene en cuenta estos efectos y permite analizar el rendimiento de un circuito considerándolos simultáneamente.

## 2.3. Impacto en el rendimiento de un circuito digital

El incremento del  $V_{th}$  causado por las trampas en la interface generadas por el efecto de NBTI, tiene una relación directa con el retardo de propagación de las compuertas digitales en un path lógico. Este incremento puede ser visto a través del concepto de *sensibilidad*, que expresa el impacto que tiene la desviación de un parámetro de cada transistor de la compuerta, en el rendimiento de la misma [16]. El efecto del NBTI puede calcularse sumando la contribución que tiene la desviación de cada uno de los parámetros que se degradan (con el tiempo) de cada transistor por separado en el rendimiento nominal. Suponiendo que únicamente el voltaje umbral presenta degradación debido a NBTI, el retardo de una compuerta digital puede expresarse como sigue:

$$D_{age} = D_o + \sum_{k=1}^K S_{V_{th},k}^D \Delta V_{th_{k,nbti}} \quad (2.3)$$

En la anterior ecuación,  $K$  es el número de transistores en la compuerta y  $D_o$  es el retardo nominal de compuerta. El impacto de la degradación del  $V_{th}$  de cada transistor  $k$  en el tiempo de propagación, es el producto del coeficiente de sensibilidad del retardo debido a variaciones en el  $V_{th}$  del transistor  $S_{V_{th},k}^D$  con la cantidad de desviación del parámetro  $\Delta V_{th_{k,nbti}}$ .

Los coeficientes de sensibilidad pueden determinarse como la derivada del retardo respecto al voltaje umbral en el punto nominal:

$$S_{V_{th},k}^D = \left. \frac{\partial D}{\partial \Delta V_{th_k}} \right|_{\Delta V_{th}=0} \quad (2.4)$$

Es importante señalar en este punto, que la variación del  $V_{th}$  de la ecuación anterior, es puramente determinista y que más adelante en el capítulo 4, se explicará cómo considerar el impacto de los efectos de variaciones de proceso en esta variación.

La figura 2.11 muestra el cambio en el tiempo del retardo de propagación de una compuerta inversora obtenido con el anterior modelo. En la figura se observa que el incremento en el retardo es proporcional al cambio en el  $V_{th}$  visto en la figura 2.10(a).

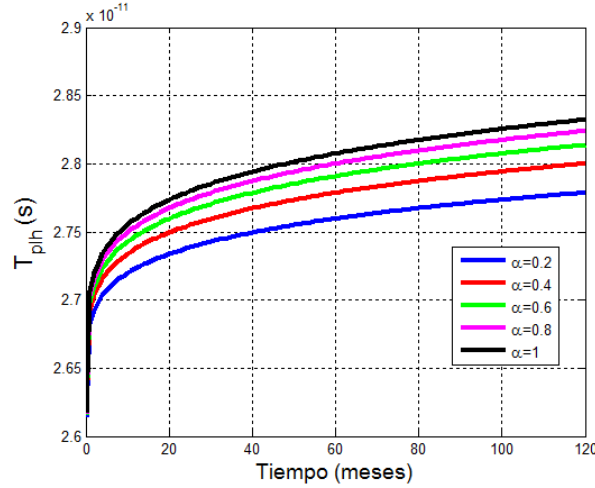


Figura 2.11: Variación del retardo de subida de una compuerta inversora

El impacto del NBTI en el retardo de un path lógico, puede calcularse sumando cada una de las contribuciones de las compuertas por separado:

$$D_{path,age} = \sum_{i=1}^N D_{age,i} \quad (2.5)$$

Donde  $N$  es el número de compuerta en el path y  $D_{age,i}$  es el retardo considerando efectos de envejecimiento de la  $i$ -ésima compuerta en el path.

En un circuito digital, la probabilidad de las señales de entrada principales puede tomar cualquier valor entre 0 y 1, esto implica que para cualquier combinación de valores, la degradación del retardo del circuito puede tomar un valor diferente. Como se muestra en la figura 2.12 para dos circuitos diferentes llamados “parity” y “9symml”, aplicar un patrón de entrada diferente causa una degradación de retardo diferente. Además la diferencia de retardo obtenida entre la degradación con un patrón de entrada y otro puede aumentar considerable-

### 2.3. IMPACTO EN EL RENDIMIENTO DE UN CIRCUITO DIGITAL

mente a través del tiempo [13]. Debido a la dificultad de conocer con exactitud el valor de probabilidad de las señales de entrada, es necesario analizar el impacto en el rendimiento bajo las peores condiciones de probabilidad que generarían una máxima degradación.

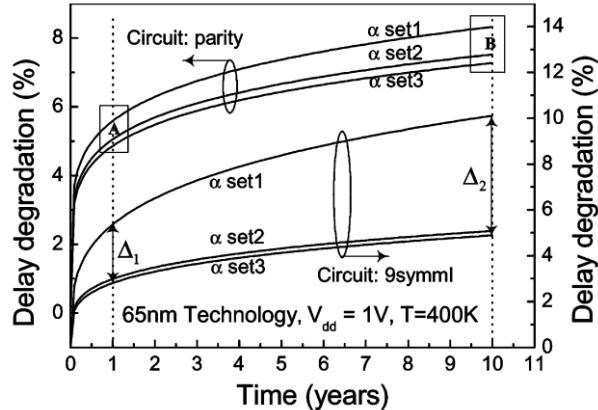


Figura 2.12: Degradación del retardo para diferentes conjuntos de probabilidad en las entradas principales de un circuito digital. [13]

Además, debe tenerse en cuenta que un patrón de entrada podría afectar en gran parte el retardo de un path determinado (supongamos un path no crítico) mientras que el impacto que tiene en otro (supongamos el path crítico) podría ser muy poco. En este caso, un path que bajo un análisis sin envejecimiento no se consideraba como path crítico, al pasar el tiempo puede convertirse en él o viceversa [13]. La figura 2.13 muestra este efecto en el circuito ISCAS *C17* donde un camino no crítico (*I10*) es más sensible al patrón de entrada aplicado que el camino crítico inicial *I9*. Entonces cuando se realiza un análisis de retardo no sólo debe considerarse el path crítico, sino también el conjunto de paths que potencialmente podrían convertirse en caminos críticos.

En [16] y [22] mencionan que debido a NBTI, el retardo del path crítico (CP) de un circuito digital puede incrementarse alrededor de un 10%. Si éste incremento es tal que el retardo del CP es mayor que el período completo reloj, podría ocurrir una falla en el sistema como se muestra en la figura 2.14. Inicialmente, cuando el

### 2.3. IMPACTO EN EL RENDIMIENTO DE UN CIRCUITO DIGITAL

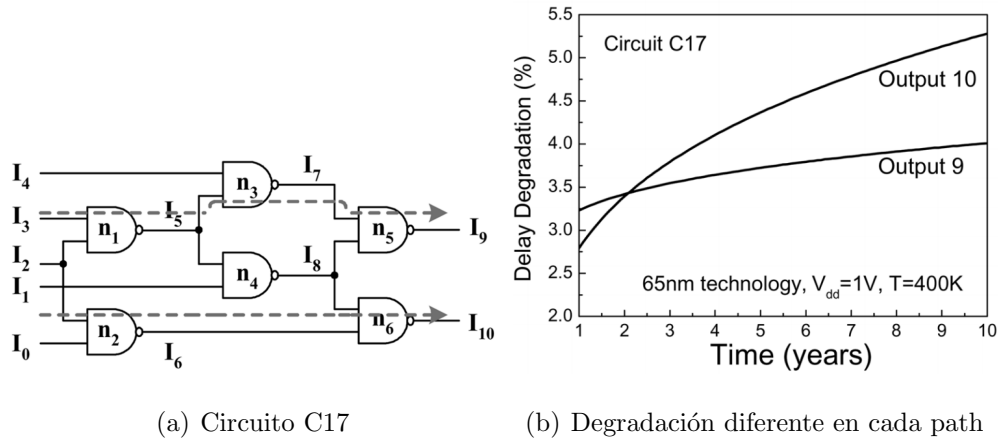


Figura 2.13: Degradación del retardo diferente para dos caminos del circuito C17. [13]

circuito no se ha envejecido y el CP realiza una operación, la transición que debe ser almacenada en los registros o FFs ocurre bastante antes del flanco positivo de reloj (donde los datos son almacenados). Sin embargo, a medida que pasa el tiempo, el circuito se envejece y la señal digital tarda más tiempo en propagarse y podría realizar la transición luego del flanco de reloj. Entonces, el dato almacenado no sería el correcto y ocurriría un error de operación.

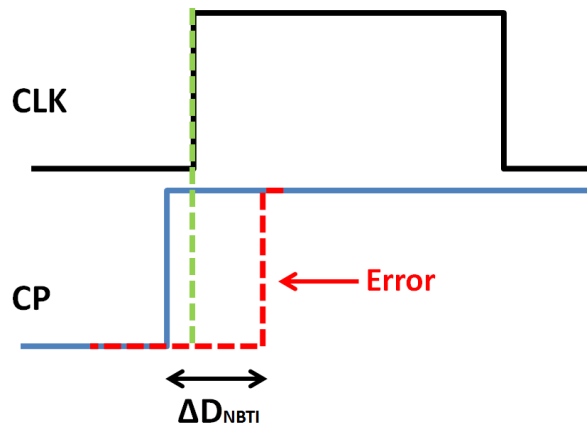


Figura 2.14: Error de operación debido al envejecimiento

Como se mencionó en el capítulo 1, para asegurar que el circuito opere correctamente durante todo su tiempo de vida útil esperado, usualmente se introduce

### 2.3. IMPACTO EN EL RENDIMIENTO DE UN CIRCUITO DIGITAL

una banda de guarda adicional en el período de reloj inicial, de forma que en los primeros años de operación el circuito opera más rápido de lo necesario, pero al pasar el tiempo, opera justo sobre la especificación de retardo. La banda de guarda requerida puede ser entre un 10% o 20% del retardo inicial del circuito [9] [20] [13].

## Capítulo 3

# Técnicas para reducir el impacto del envejecimiento en circuitos integrados digitales

En los capítulos anteriores, se han explicado los fundamentos del efecto de NBTI y cómo se ve reflejado este fenómeno ocasionando un incremento en el retardo de propagación de las señales digitales a través de una compuerta y un path lógico. De éstos análisis se concluyó que el incremento de retardo debido a NBTI depende principalmente de la probabilidad de que los transistores de una compuerta se encuentren bajo la condición de estrés, de la temperatura de operación y el voltaje de alimentación.

Para evitar que ocurran errores de propagación, éste incremento del retardo es considerado al ajustar el período de reloj inicial de un circuito digital, introduciendo la banda de guarda. Sin embargo, esto ocasiona que no se aproveche al máximo la capacidad de los dispositivos en su primera etapa del tiempo de vida.

Las principales técnicas o estrategias para disminuir el impacto del NBTI en un circuito digital y así reducir la banda de guarda requerida pueden clasificarse según el nivel en el que se realizan:



- Técnicas a nivel transistor: Consisten en modificar las propiedades físicas de los materiales utilizados en el proceso de fabricación, con el objetivo de que el impacto del NBTI sea menos significativo en las características eléctricas del dispositivo.
- Técnicas a nivel circuital: Se basan en estrategias ya sea para que el incremento del retardo debido a la degradación sea menor y así minimizar el margen de seguridad requerido o para optimizar el rendimiento a través del tiempo según el envejecimiento de los dispositivos. Las técnicas a nivel circuito pueden enfocarse en optimizar el desgaste durante la etapa de operación activa o durante los tiempos en el que el circuito no es utilizado y se encuentra en estado de standby.

A continuación se revisarán algunas de estas técnicas haciendo un mayor enfoque en las estrategias adaptivas que hacen parte de las estrategias a nivel circuito en estado activo de operación.

### 3.1. Incremento de la movilidad efectiva de los huecos

En [18] sugieren que la disminución de la corriente del transistor debido al aumento de  $V_{th}$  por NBTI, puede ser compensado si la movilidad de los huecos se incrementa a medida que el dispositivo se envejece. El fundamento de esta idea parte de que la corriente en la región lineal de un transistor PMOS puede ser expresada como  $|I_D| \sim \mu_{eff}(|VG| - |V_{th}|)$ . Por tanto, la variación en la corriente, el voltaje y la movilidad puede ser relacionada por:

$$\frac{\Delta I_D}{I_{D0}} = \frac{\Delta \mu_{eff}}{\mu_{eff0}} - \frac{|\Delta V_{th}|}{VG - V_{th0}} \quad (3.1)$$

Donde los parámetros con sub-índice (0) corresponden al valor inicial sin degradación.

### 3.1. INCREMENTO DE LA MOBILIDAD EFECTIVA DE LOS HUECOS

A medida que el transistor se degrada, las trampas en la interface son generadas y la magnitud del voltaje umbral del transistor aumenta, contribuyendo negativamente al cambio en la corriente a través del segundo término de la ecuación 3.1. Sin embargo, debido a la generación de carga positiva en la interface, el campo eléctrico efectivo tampoco puede permanecer constante ( $|E_{eff}| = \frac{Q_{dep} + Q_{inv}/3}{\epsilon_{Si}}$ ), de hecho debe disminuir, ya que las trampas en la interface disminuyen la carga de la capa de inversión y de la región de agotamiento. Entonces, el cambio en la corriente puede ser compensado aprovechando la relación entre el campo eléctrico y la movilidad.

Los autores proponen el uso de transistores de silicio forzado (*Strained Silicon*) ya que en ellos la pendiente entre la movilidad y el campo eléctrico es negativa como muestran en los resultados experimentales de la figura 3.1. En este caso, una reducción en el campo eléctrico debido a NBTI, ocasiona un aumento en la movilidad efectiva que podría compensar el incremento de la magnitud del  $V_{th}$  y reducir el cambio total de la corriente. En la figura se muestran curvas para dos materiales hipotéticos cuya relación de movilidad y campo eléctrico efectivo tiene la pendiente suficiente para compensar el envejecimiento.

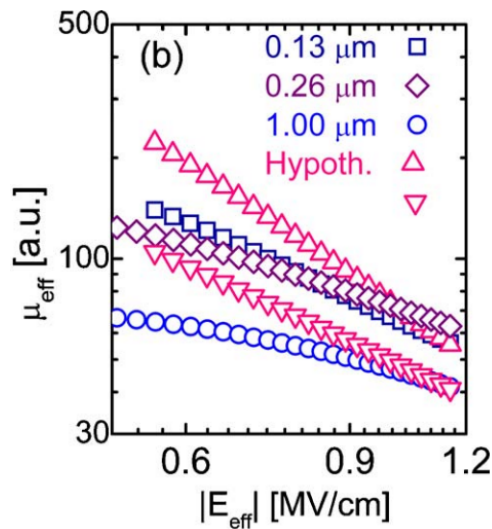


Figura 3.1: Curvas de movilidad vs campo eléctrico en silicio forzado [18]

La desventaja de esta estrategia sería que utilizar este tipo de materiales implicaría un alto costo en el proceso de fabricación.

## 3.2. Disminución del voltaje de alimentación

Como su nombre lo indica, consiste en aplicar una tensión de alimentación por debajo de la nominal mientras el circuito se encuentre en estado de reposo o standby, de esta forma se reduce tanto el efecto de NBTI como el consumo de potencia. Sin embargo, requiere insertar rieles extra de alimentación y convertidores de nivel. [19]

## 3.3. Control del vector de entrada

Esta técnica consiste en aplicar el vector de entrada al circuito digital que genere una menor degradación debido a NBTI cuando el circuito se encuentre inactivo o en standby. Para determinar un vector óptimo evitando realizar barridos exhaustivos, es posible utilizar un generador de vectores aleatorio y seleccionar el vector más adecuado. Debido a que el vector óptimo para disminuir el envejecimiento por NBTI puede no ser el mejor en cuanto a consumo de potencia, en [19] proponen analizar los vectores en un diagrama de pareto como el mostrado en la figura 3.2 y de esta forma seleccionar el vector que mejor optimice ambos valores. En esta técnica, la penalización requerida en área y tiempo por las memoria o flip-flops que almacenen el vector óptimo en las entradas principales puede ser despreciada en circuitos grandes.

Aunque la técnica de control de los vectores de entrada permite disminuir el impacto del NBTI, su eficiencia puede no ser suficiente en circuitos grandes [22] [23] ya que sólo permite controlar ciertos nodos internos.

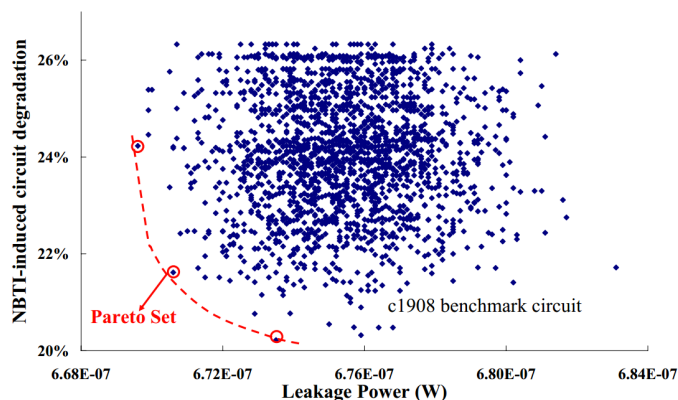


Figura 3.2: Conjunto de pareto para disminuir efecto de NBTI y consumo de potencia simultáneamente durante el período de standby [19]

### 3.4. Control de Nodos Internos

Otra estrategia para la disminución del efecto de NBTI durante el período de tiempo en standby de un circuito digital, consiste en reemplazar la estructura de algunas compuertas digitales e insertar una señal extra que controle el estado de algunos nodos internos, especialmente los de entrada de aquellas compuertas más afectadas por el NBTI durante el período de operación dinámico. La figura 3.3 ilustra el concepto de esta técnica. Los autores de [22] y [23] utilizan la estructura de la izquierda para controlar el nodo a la salida de las compuertas tipo NAND–AND, mientras que la derecha se utiliza para controlar las compuertas tipo NOR–OR.

Al activar la señal de *sleep* (únicamente en el estado de standby), se cambia el estado de algunos nodos internos, llevando ciertas compuertas (que tienen estos nodos como fan–in) de la condición de estrés a la de relajación o recuperación. Bajo esta metodología, una compuerta de dos entradas, se convierte en una de tres, por tanto, no se requiere estrictamente una estructura de compuertas diferentes a las de las celdas estándar. Sin embargo, la técnica intrduce penalizaciones en retardo, consumo de potencia y área. La penalización en retardo, es en parte compensanda a largo plazo por la reducción del impacto del NBTI. El consumo

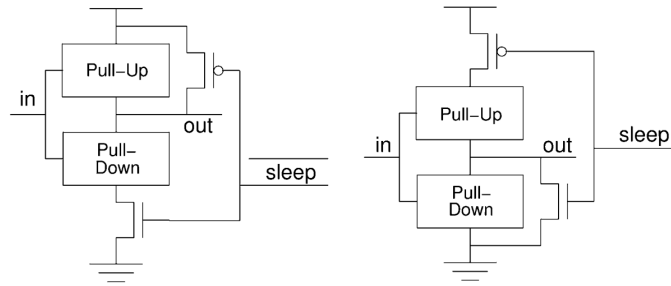


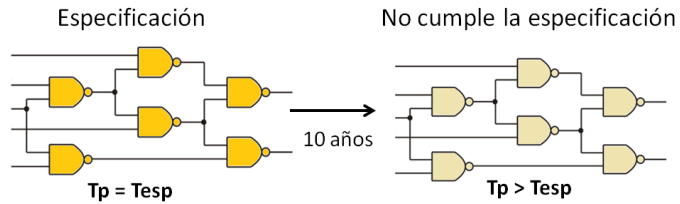
Figura 3.3: Técnica de reemplazo de compuertas para controlar el estado de algunos nodos internos

de potencia dinámico durante el lapso de tiempo de operación activo o en standby es despreciable ya que la señal de sleep permanece a un único estado lógico. Por otra parte, se incrementa la complejidad en la etapa de ruteado [23]. Tanto en [22] como en [23] presentan algoritmos para seleccionar adecuadamente las compuertas que deben ser reemplazadas y optimizar la reducción del efecto de NBTI junto con la técnica de control del vector de entrada.

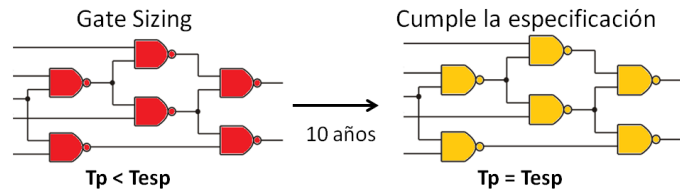
### 3.5. Redimensionamiento de las compuertas

El sobredimensionamiento de las compuertas es una forma de cumplir la especificación de retardo, ya que los dispositivos funcionarán más rápido durante su tiempo de operación [17], de forma que cuando se envejecen, aún cumplan la mínima restricción de tiempo. La metodología consiste, como se ilustra en la figura 3.4, en no realizar el diseño pensando en cumplir la especificación nominal, sino teniendo en cuenta el porcentaje de incremento del retardo que se tendrá durante el tiempo de vida útil esperado.

En este tipo de técnicas, resulta importante un cálculo adecuado de la probabilidad de las señales de entrada en cada compuerta digital y una selección precisa de las compuertas que serán redimensionadas ya que de lo contrario la penalización en área podría ser significativa.



(a) Diseño para la especificación



(b) Sobre dimensionamiento para cumplir la especificación  
luego de 10 años

Figura 3.4: Sobre dimensionamiento de las compuertas

En [20] asignan (suponen) los valores de probabilidad de las entradas principales y éstos son propagados a través de todo el circuito, de esta forma evitan realizar el sobre dimensionamiento para un peor caso de envejecimiento (donde se consideraría que la probabilidad de que todos los transistores se encuentren en condiciones de estrés es de la unidad,  $\alpha = 1$ ). Sin embargo, esta suposición no es viable, ya que los valores de probabilidad de entrada, en general no pueden ser estimados adecuadamente para cualquier circuito combinatorial.

En [17] sugieren optimizar las compuertas más cercanas a las salidas principales, ya que la señal digital que llega a estas compuertas presenta una mayor degradación en el slew rate (debido a el incremento del  $V_{th}$  de las compuertas anteriores) y por tanto, su impacto en el incremento del retardo de éstas últimas compuertas será mayor.

Por otra parte, en [21] proponen una estrategia para ranquear las compuertas en un path según su incremento del retardo y aquellas que tengan una degradación mayor a un porcentaje umbral, son optimizadas.

### 3.6. Ordenar los nodos de entrada en las compuertas

Como se explicó en el segundo capítulo, el modelo de degradación del  $V_{th}$  por NBTI, depende de la probabilidad de que el transistor se encuentre bajo condiciones de estrés ( $\alpha$ ). En una compuerta inversora o compuerta de tipo NAND, donde los transistores PMOS se encuentran en paralelo y su fuente está conectada directamente al riel de alimentación, la probabilidad de que el transistor se encuentre en condiciones de estrés es igual (depende únicamente) a la probabilidad de que la señal de entrada se encuentre en nivel lógico de cero (ver figura 3.5).

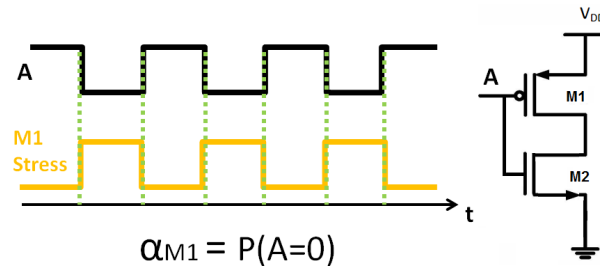


Figura 3.5: Condiciones de estrés en una compuerta inversora

Por otra parte, en compuertas donde hay dos o más transistores PMOS en serie (o *stack*), para que un transistor esté bajo condiciones de estrés, todos los demás transistores en una posición superior a él, también deben estarlo. Entonces, la probabilidad de que un transistor esté bajo estrés corresponde a la probabilidad de que su señal de entrada esté en cero lógico y que todos los demás transistores arriba de él estén bajo estrés. En la figura 3.6 se observa que aunque en algunos intervalos de tiempo, la señal B tiene el valor de cero lógico, el transistor  $M2$  no se encuentra polarizado en la condición de estrés, debido a que la tensión en su fuente no corresponde a  $V_{dd}$ , ya que el transistor  $M1$  se encuentra apagado.

Entonces, la idea de esta técnica es la de manipular las probabilidades de

### 3.6. ORDENAR LOS NODOS DE ENTRADA EN LAS COMPUERTAS

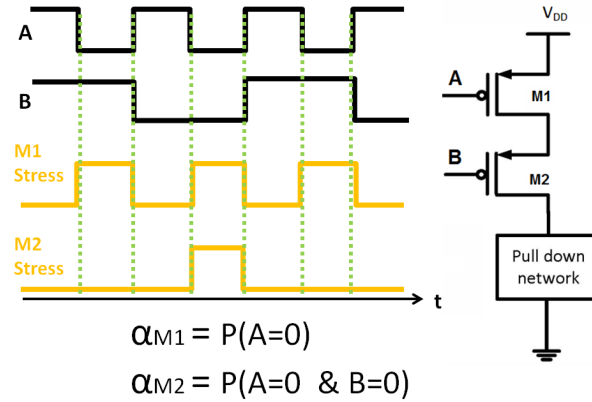


Figura 3.6: Condiciones de estrés en una compuerta NOR de dos entradas

estrés, ordenando los nodos de entrada de forma que el efecto del NBTI en las compuertas críticas pueda ser reducido [26].

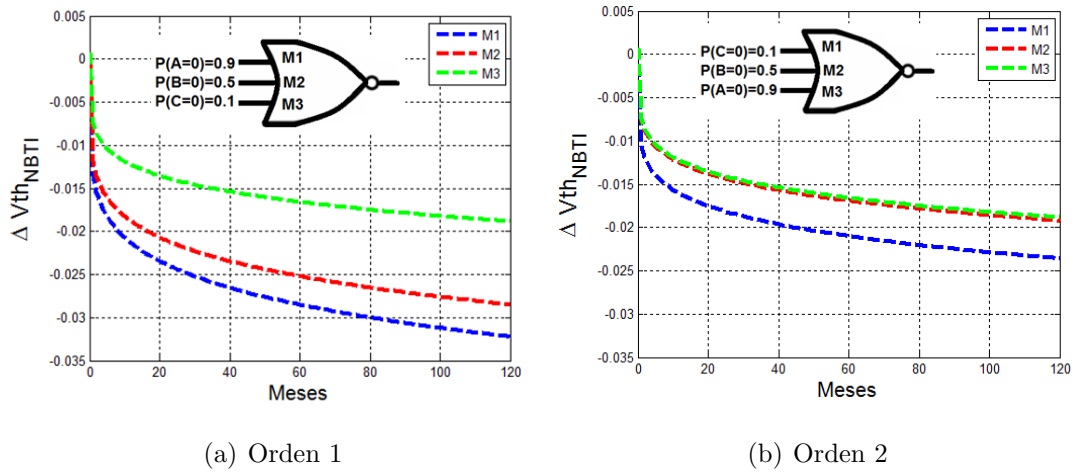


Figura 3.7: Ordenamiento de los nodos de entrada para disminuir el envejecimiento en una compuerta NOR3

La figura 3.7 muestra la disminución del efecto del NBTI en los transistores M2 y M3 obtenida en una compuerta NOR de tres entradas, tan sólo con realizar un adecuado ordenamiento de los nodos de entrada en función de su probabilidad. De esta figura también es obvio, que la degradación de los transistores arriba del stack siempre será mayor que los de abajo. Nótese además que el orden de los nodos adecuados es de menor a mayor probabilidad comenzando por el transistor



PMOS conectado directamente a  $V_{dd}$  ( $M1$ ).

### 3.7. Circuitos Adaptivos

La estrategia de utilizar circuitos adaptivos, busca mitigar la cantidad de banda de guarda requerida realizando una sintonización de los parámetros del sistema (frecuencia o voltaje) a medida que el circuito se desgasta con el tiempo.

El proceso de sintonización, puede llevarse a cabo de forma *offline* que consiste en realizar el proceso de sensado de envejecimiento únicamente cuando el circuito se encuentre en reposo (viable en sistemas con múltiples núcleos), o en los momentos de encendido (por ejemplo cada vez que se enciende un auto). Este tipo de sintonización puede ser imperceptible para el usuario y gran cantidad de patrones a la entrada pueden aplicarse para sensar el retardo de cada camino lógico deseado [28]. Sin embargo, existen aplicaciones que se encuentran operando constantemente y detener su funcionamiento para realizar un test de envejecimiento resulta inaceptable. Para este tipo de aplicaciones es útil realizar un monitoreo *on-line* de forma paralela al tiempo de operación, haciendo un seguimiento más exhaustivo.

Las variables del sistema que se pueden modificar según el envejecimiento de un circuito son:

- Voltaje de alimentación: La corriente que fluye por los dispositivos es proporcional al voltaje de alimentación utilizado. Entonces, para una especificación de frecuencia dada, se disminuye el voltaje de alimentación cuando el circuito se encuentra sin envejecimiento y opera rápidamente, reduciendo el consumo de potencia. Respectivamente, el voltaje de alimentación se aumenta a medida que éste se envejece.
- Frecuencia de Operación: De forma análoga al caso del voltaje de alimentación, se aumenta la frecuencia de operación cuando el circuito se encuentra joven,

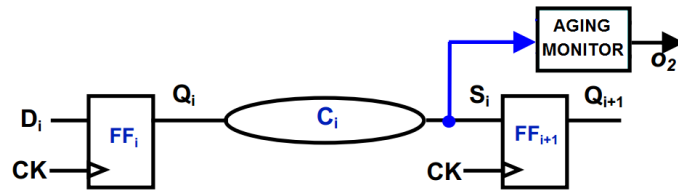
teniendo un máximo rendimiento, y luego se disminuye a medida que el circuito se desgasta. De esta forma, se obtiene un rendimiento óptimo durante todo el ciclo de vida útil.

- Efecto cuerpo: Esta idea consiste en aprovechar el efecto cuerpo para compensar el incremento del voltaje umbral.

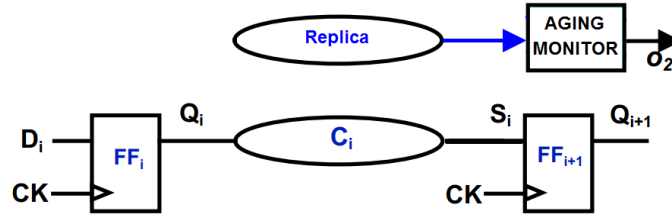
Las técnicas adaptivas requieren el uso de sensores o monitores para estimar que tanto se ha degradado el retardo de los caminos críticos del circuito debido al envejecimiento y así poder realizar la sintonización requerida. Al medir el envejecimiento en un path mediante estos sensores, es posible reducir significativamente la banda de guarda requerida, ya que a medida que el path de interés se envejece, el cambio en el retardo es detectado por los sensores y estos envían una señal de alerta a los controladores de los parámetros de operación mencionados anteriormente.

### 3.7.1. Path réplica y sensores embebidos

Los sensores pueden ser insertados ya sea directamente a la salida del path de interés (de forma embebida) o a la salida de un path réplica aislado que emula en lo posible las características del path de interés. La figura 3.8 ilustra estas posibilidades. El path réplica, es fácil de implementar y puede ser diseñado de forma que pueda calibrarse luego del proceso de fabricación (aumentando los costos de test), para así considerar posibles diferencias debido a variaciones de proceso [29]. Sin embargo, la degradación del path crítico de interés depende altamente de la probabilidad de las señales de entrada a la lógica combinacional y conectar éstas al path réplica trae complicaciones de ruteado e integridad de señal. Además, el path crítico de un circuito digital puede cambiar a través del tiempo y la réplica no puede emular correctamente a todos. Para solucionar los inconvenientes mencionados, al utilizar un path réplica, debe garantizarse que su retardo sea mayor y que éste se degrade más rápido que cualquier otro path en el circuito [29] [28].



(a) Sensor Embebido



(b) Path réplica + Sensor

Figura 3.8: Tipos de sensores

Por otra parte, insertar sensores directamente a la salida de los paths críticos del circuito, no requiere de calibraciones post layout [29], ya que el retardo es medido directamente del path de interés. Además permiten reducir al máximo la banda de guarda requerida, ya que cualquier incremento de retardo debido no sólo al envejecimiento será correctamente detectado por los sensores. Para reducir el consumo de potencia, resulta útil activar los sensores únicamente cuando el path crítico se active, ya que múltiples caminos pueden compartir una misma salida principal pero sólo algunos son potencialmente críticos. Igualmente, garantizar que el sensor se activa únicamente cuando el path crítico lo hace permite alternativas como el control de la velocidad, tanto aumentando como disminuyendo el período de operación cuando no se han detectado transiciones tardías por parte de los paths críticos durante largos períodos de tiempo [28]. Sin embargo, las desventajas de utilizar esta forma de sensado es que cargan directamente el path crítico, teniendo una penalización en el retardo de propagación. Además, el proceso de insertar los sensores y realizar una correcta selección de los paths que deben ser monitoreados es más complejo que en el caso del path réplica.

### 3.7.2. Predicción de error

Esta técnica consiste en predecir el punto operacional donde el path crítico del circuito falla en cumplir las especificaciones de retardo (cuando el retardo del path crítico está cerca del flanco de reloj) y garantizar la correcta operación añadiendo márgenes de seguridad (mucho menores a las requeridas en un circuito convencional) para predecir el punto de falla [30].

La figura 3.9 ilustra la estructura general de un sensor de predicción de error y su funcionamiento puede explicarse con las señales de la figura 3.10. Cuando el path crítico del circuito se encuentra sin envejecimiento (CP without aging), realiza una transición con suficiente tiempo de antelación al flanco positivo del reloj, donde los flip-flops muestrean y retienen el dato propagado, de esta forma, ambos FF almacenan el mismo valor y el circuito continúa operando de forma convencional.

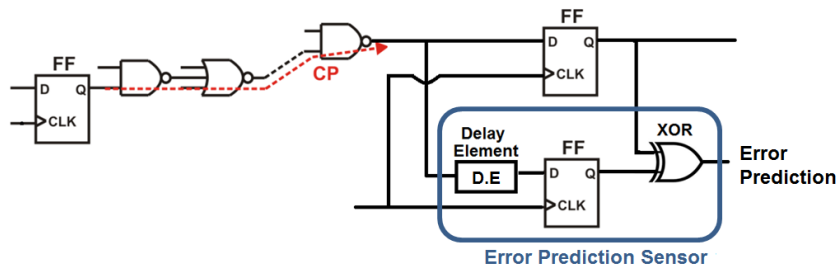


Figura 3.9: Sensor de predicción de error

Cuando el retardo del path crítico se incrementa debido al envejecimiento (CP with aging), la transición que realiza está más cerca del flanco de reloj (más cerca de generar una falla), si el tiempo entre el momento en el que ocurre esta transición y el flanco de reloj es menor que el retardo definido como margen por el elemento de retardo del sensor (Delay Element), la señal retardada a la salida de este elemento realizará su transición después del flanco de reloj y esta transición ya no será almacenada en el FF del sensor. Entonces, las señales retenidas en el FF principal y en FF del sensor serán diferentes y al compararse en la compuerta

XOR, activarán la señal de alerta (Error Prediction), indicando que el período de reloj debe aumentarse.

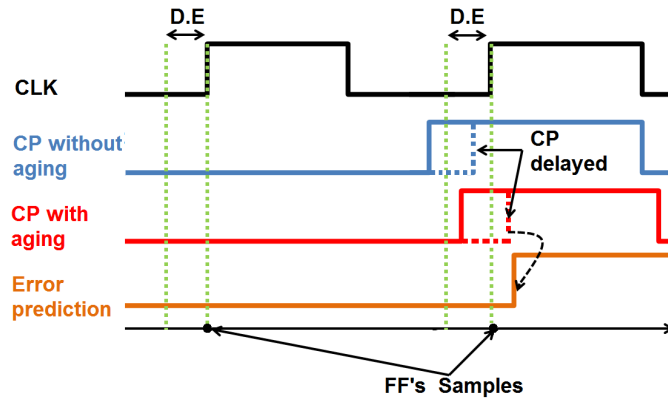


Figura 3.10: Funcionamiento del sensor de predicción de error

Las principales ventajas de esta técnica es que mantiene siempre un correcto funcionamiento del circuito, evitando propagar datos corruptos y puede ser aplicable a cualquier red combinatorial. Sin embargo, requiere aún de una margen de seguridad definida por el elemento de retardo y no puede predecir variaciones rápidas de retardo como caídas en  $V_{dd}$  o *soft errors* por radiación [31]. Igualmente, una falsa predicción de envejecimiento, debido a otro tipo de variaciones dinámicas de retardo, podría ocasionar una modificación de la frecuencia de operación, cuando no sea necesario.

### 3.7.3. Detección y corrección de error

El concepto base de estos esquemas es el de escalar los parámetros del sistema sólo hasta el punto en el que un error de tiempo ha ocurrido [32] [33]. Una vez el error es detectado, debe realizarse una lógica de recuperación del dato correcto, para que el circuito pueda continuar operando correctamente. Al permitir que el circuito falle, pueden reducirse al máximo las márgenes de seguridad requeridas para tolerar los peores casos de retardo de propagación, obteniendo un rendimiento significativamente mayor, incluso que las técnicas de predicción de error [29].

Aunque este tipo de técnicas pueden tolerar variaciones de retardo de todo tipo (alta frecuencia o baja frecuencia), existe un compromiso entre el beneficio de una menor frecuencia de operación y el tiempo extra requerido por los procesos de recuperación de error.

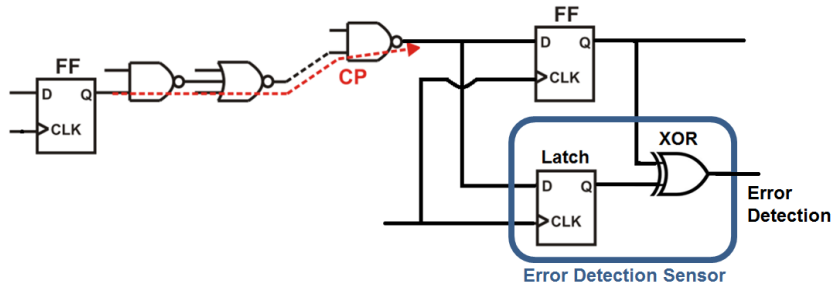


Figura 3.11: Sensor de detección de error

La figura 3.11 muestra la estructura general de este tipo de sensores y su funcionamiento puede ser entendido por el diagrama de tiempos de la figura 3.12. Cuando el path crítico se encuentra sin envejecimiento (CP without aging), el circuito opera de forma convencional (nótese que en este caso, la transición del CP puede ocurrir muy cerca del flanco de reloj, desde luego, manteniendo el tiempo de setup y hold requeridos por el FF). A medida que el retardo del path incrementa debido al envejecimiento de los dispositivos, eventualmente su retardo de propagación puede superar incluso el período de reloj, realizando una transición luego del flanco positivo del reloj (CP with aging), la cual no será correctamente almacenada en el FF principal. Sin embargo, durante el período de tiempo en el que la señal de reloj se encuentra en alto, el latch (también llamado “*shadow latch*”) se encuentra en estado de transparencia y éste sí almacena la señal tardía. De esta forma, el dato presente en el FF principal será diferente al dato en el latch, activando la señal de error al ser comparados en la compuerta XOR.

Una vez la señal de error es activada, la lógica de recuperación del dato debe llevarse a cabo. Para esto usualmente se repite nuevamente la operación realizada a una menor frecuencia de operación o mayor voltaje de alimentación [29].

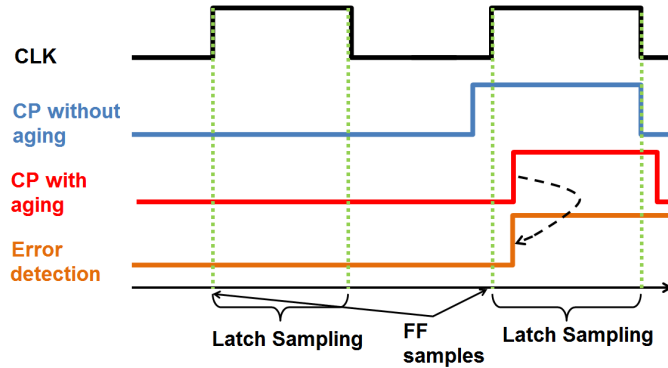


Figura 3.12: Funcionamiento del sensor de detección de error

La figura 3.13 muestra la arquitectura utilizada para la implementación de la estrategia de detección y corrección de errores en un microprocesador. Al iniciar un proceso los datos son propagados por cada una de las etapas de lógica combinacional conectadas en *pipeline*, si un path realiza una transición tardía que es detectada por el sensor de error visto anteriormente, la señal de error se activa y es propagada de etapa en etapa hasta generar la orden de que el dato propagado es inválido. La unidad de control de error entonces indica al controlador de los parámetros del sistema que aumente el período de reloj o la tensión de alimentación y así el procesador vuelve a ejecutar la tarea que esta vez se lleva a cabo correctamente.

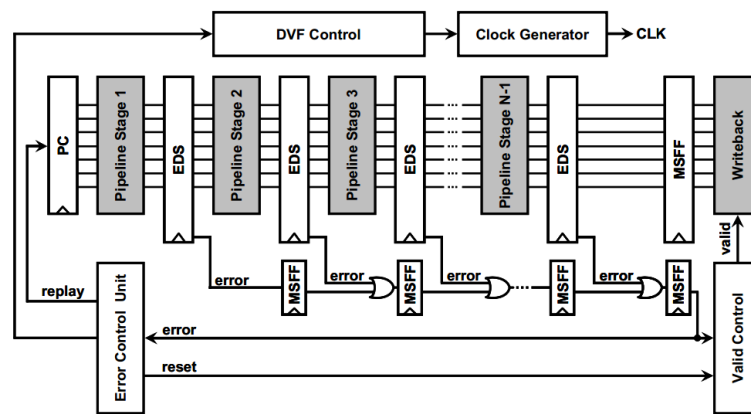


Figura 3.13: Corrección de error en un microprocesador [29].

Por otra parte, este tipo de estrategias tiene una mayor complejidad en el diseño, ya que todos los paths que terminan en un sensor de este tipo, deben satisfacer condiciones de mínimo y máximo retardo de propagación (ver figura 3.14). Si un path tiene un retardo de propagación muy bajo, podría llegar al sensor durante el intervalo de tiempo en el que el latch se encuentra en estado de transparencia, ocasionando una falsa señal de error. Por otra parte, si el retardo del path crítico de interés es muy grande de tal forma que la transición ocurre luego de que el latch salga del estado de transparencia, el error no podrá ser detectado [29]. Matemáticamente, estos requisitos son:

$$D_{CP,max} < T_{clk} + T_{Latch} \quad (3.2)$$

$$D_{CP,min} > T_{Latch}$$

Donde  $D_{CP}$  es el retardo de propagación de un path crítico de la lógica combinacional,  $T_{clk}$  es el período de reloj y  $T_{Latch}$  es el tiempo durante el cual el latch se encuentra en estado de transparencia.

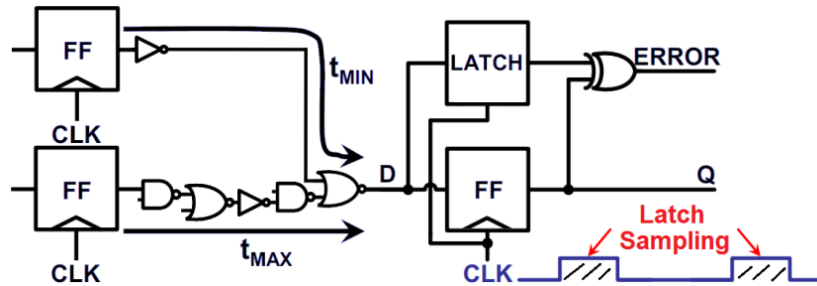


Figura 3.14: Funcionamiento del sensor de detección de error

Para cumplir con el requisito de retardo mínimo, una alternativa es la de insertar buffers en aquellos paths rápidos que podrían ocasionar una falsa señal de error hasta que su retardo supere al mínimo permitido [29]. Otra alternativa consiste en insertar decoders que permitan identificar cuando alguno de los paths que se requieren monitorear se activa y de esta forma activar también los sensores de retardo.



Además de reducir al máximo los márgenes de seguridad requeridas, una característica importante de esta técnica combinada con el uso de decores es que permite explorar nuevas alternativas, como explotar la probabilidad de activación de los paths. Si un path crítico tiene un retardo alto, pero también una baja probabilidad de activación, podría permitirse operar con un período de reloj bastante menor que el retardo del path y detectar y recuperar el dato erróneo cada vez que el path se active, ya que esto sucederá infrecuentemente y el tiempo de recuperación empleado en este caso, se verá ampliamente compensado por una frecuencia de operación mayor [29] [35].

# Capítulo 4

## Estimación del retardo del circuito considerando efectos de NBTI y variaciones de proceso

En este capítulo se explicará la metodología seguida para calcular el valor medio y la varianza del retardo de un path, así como la covarianza y correlación entre paths, considerando efectos de variaciones de proceso y envejecimiento por NBTI de forma conjunta. Esta metodología fue programada en *C++* y se utilizó para realizar los análisis que se expondrán en el capítulo 5 del presente trabajo.

### 4.1. Consideraciones

En esta sección se expondrán las consideraciones fundamentales que deben tenerse en cuenta en un análisis estadístico.

#### 4.1.1. Variaciones en los parámetros

Debido a que el impacto de las variaciones de proceso es cada vez mayor, los análisis deterministas clásicos de esquinas ya no son adecuados para estimar el

rendimiento de un circuito digital, ya que sobreestiman la varianza de los dispositivos. De hecho, un análisis de esquina es similar a asumir que la correlación entre todos los dispositivos de un mismo tipo (NMOS o PMOS) es cercana a la unidad, de manera que todas sus variaciones se orientan simultáneamente hacia la esquina lenta o rápida, lo cual no es realista.

Por otra parte, en un análisis estadístico se representa cada variación de un parámetro de proceso como una variable aleatoria con su correspondiente función de distribución de probabilidad (ej:  $\Delta V_{th} = N(\mu, \sigma)$ ). De esta forma, el retardo de propagación de una compuerta o de un path en el circuito serán una función de estas variables aleatorias y por tanto, también son una variable aleatoria determinadas por sus respectivas funciones de distribución de probabilidad.

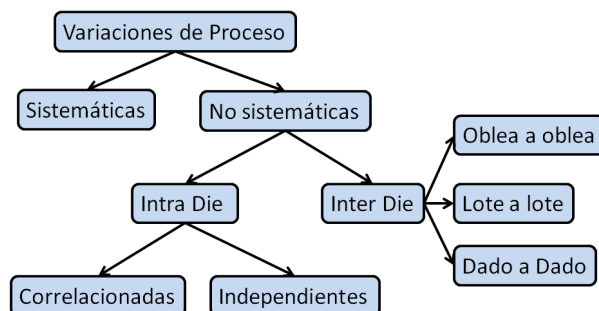


Figura 4.1: Clasificación de las variaciones de proceso

La figura 4.1 muestra una clasificación general de las variaciones de proceso [38]. Las variaciones sistemáticas surgen de procesos como el CMP (*Chemical Mechanical Planarization*) o el OPC (*Optical Proximity Correction*) y siguen un comportamiento que puede ser bien determinado analizando el layout diseñado. Las variaciones no sistemáticas se deben a inexactitudes del proceso de control durante la fabricación y pueden ser clasificadas en variaciones dentro de un mismo dado (intra die) o variaciones de dado a dado (inter die). Con el escalado tecnológico, las variaciones a nivel intra die, tienen un mayor aporte sobre la variación total de un parámetro. Estas variaciones pueden tener cierto grado de correlación debido a efectos de proximidad (Variaciones intra die globales) como

se explicará más adelante o pueden ser puramente aleatorias debido a fenómenos como el *Random Dopant Fluctuation* (RDF) (variaciones intra die locales).

#### 4.1.2. Modelo de retardo de compuerta

Generar los modelos de retardo de compuerta es considerado el primer paso en el análisis ya que de éstos representan las *celdas de compuerta estadísticas* que serán utilizadas en el *Statistical Static Timing Analysis* (SSTA). En un *Static Timing Analysis* (STA) convencional, cada celda es generada mediante la caracterización del retardo de cada compuerta como función del nodo de entrada y de su capacitancia de fan-out. En un análisis SSTA, la variabilidad del retardo respecto a cada uno de los parámetros de proceso considerados debe también tenerse en cuenta.

El retardo de compuerta es función de las variables aleatorias de los parámetros de proceso, por lo tanto, también puede ser considerado como una variable aleatoria que tiene una determinada función de distribución de probabilidad. La variable aleatoria de retardo de una compuerta puede ser representada como una combinación lineal de aquellos parámetros que depende (ver ecuación 4.1). Este modelo es llamado *modelo lineal de retardo de compuerta*.

$$D = D_0 + \sum_{k=1} S_k \cdot X_k \quad (4.1)$$

Donde  $D_0$  representa el retardo de la compuerta bajo condiciones nominales,  $S_i$  es el coeficiente de sensibilidad del retardo a variaciones en el parámetro  $k$  y  $X_k$  es la magnitud de dicha variación del parámetro.

Típicamente los parámetros cuya variación ( $X_k$ ) presenta un mayor impacto (mayor sensibilidad) en el retardo de compuerta son el ancho de canal ( $W$ ), la longitud de canal ( $L$ ), el espesor del óxido ( $T_{ox}$ ) y el voltaje umbral ( $V_{th}$ ). Ésta variación puede tener cierta correlación espacial con los demás dispositivos o puede ser puramente aleatorio para el caso del efecto de RDF en el voltaje umbral. En-

tonces, es necesario pre-caracterizar la sensibilidad del retardo de compuerta correspondiente a cada uno de estos parámetros junto con el retardo bajo condiciones nominales.

En este trabajo, la caracterización se llevó a cabo mediante simulaciones en HSPICE. La figura 4.2 ilustra un esquemático del procedimiento de simulación realizado para caracterizar el retardo nominal de cada compuerta. Para cada nodo de entrada de la compuerta bajo caracterización, se aplica una pulso de entrada cuyo slew rate ( $SR_{in}$ ) se varía dentro de un rango determinado. Igualmente, en el nodo de salida de la compuerta se introduce una capacitancia variable que emulará cada uno de los posibles valores de capacitancia de fan-out vistos por la compuerta. Para cada combinación de  $SR_{in}$  y  $CL$ , se obtiene el valor correspondiente de retardo de propagación de alto a bajo ( $t_{phl}$ ) y de bajo a alto ( $t_{plh}$ ). De esta forma el retardo de propagación de la compuerta estará en función de la capacitancia de carga vista y del SR de entrada. Igualmente se determinan los respectivos slew rates a la salida ( $SR_{ohl}$  y  $SR_{oh}$ ) ya que éstos corresponderán a los SR a la entrada con los que se deberá evaluar el retardo de las compuertas subsiguientes en el circuito, cuando se realice el SSTA [43]. Debe mencionarse que mientras se caracteriza el retardo de un nodo de entrada determinado, los demás se polarizan al valor de tensión que sensibilice a la compuerta por el nodo de entrada considerado ( $V_{dd}$  para las compuertas tipo NAND y AND, GND para las compuertas tipo NOR y OR, etc.).

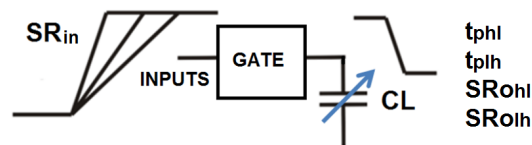


Figura 4.2: Caracterización del retardo de propagación de una compuerta digital

La figura 4.3 muestra un ejemplo de la matriz de datos obtenida para el tiempo de propagación de alto a bajo de una compuerta inversora en función de  $CL$  y

SRin. Para reducir el espacio de memoria que requiere almacenar cada uno de los puntos caracterizados en la herramienta de cómputo y también poder utilizar valores continuos de capacitancia de carga y SR a la entrada (la caracterización se realiza para valores discretos), la matriz obtenida es aproximada mediante un polinomio de dos variables.

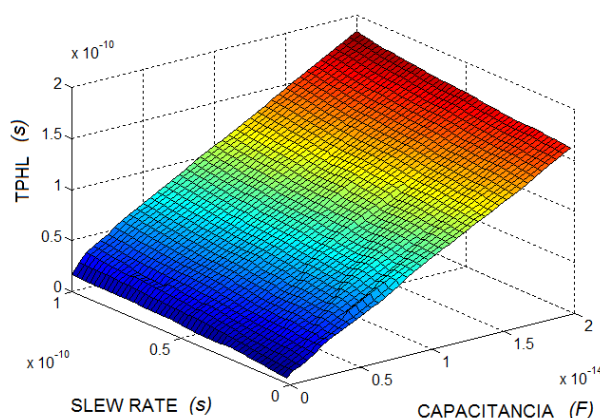


Figura 4.3: Matriz de TPHL para SR vs CL en un inversor

Para el caso de la caracterización de las sensibilidades del retardo a variaciones en los parámetros, el efecto del SR de entrada en la sensibilidad no fué considerado. En este caso, bajo un valor de SR constante a la entrada y para cada capacitancia de carga, se realizó un análisis de Monecarlo variando uno a uno los parámetros de interés por separado. Por simplicidad, ya que todos los transistores pertenecientes a una misma compuerta estarán muy cerca entre sí en el layout, sus variaciones intra die globales se consideraron completamente correlacionadas en la simulación (se utilizó la misma variable aleatoria para describir las variaciones en cada transistor simultáneamente). Para el caso del voltaje umbral, ya que cada transistor puede presentar un valor diferente tanto de envejecimiento como de variación local por RDF, la sensibilidad en el retardo de compuerta fue caracterizada para cada transistor por separado obteniendo un valor correspondiente a cada uno de los transistores presentes en la compuerta. Al igual que para el retardo medio, cada una de las matrices de caracterización obtenidas (ver figura 4.4) fue

aproximada por un polinomio de dos variables y la sensibilidad es calculada como la derivada de este polinomio respecto al parámetro correspondiente evaluada en el valor del parámetro nominal.

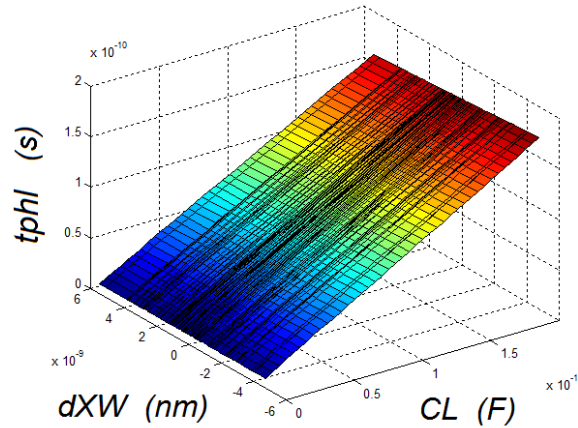


Figura 4.4: Matriz de  $t_{phl}$  vs  $CL$  vs  $DXW$  para una compuerta inversora

Una vez realizada esta caracterización, el circuito puede ser analizado a nivel de compuerta, donde cada una está descrita por su retardo nominal y sus respectivas sensibilidades a cada uno de los parámetros.

### 4.1.3. Correlación espacial

El grado de correlación es un indicador que representa que tan relacionada está la variación de una variable aleatoria con otra. Como se explicó anteriormente, cada compuerta puede ser representada como una combinación lineal de variables aleatorias que representan los parámetros eléctricos de los transistores en la compuerta. Entonces, la correlación espacial indica la relación entre las variaciones de un parámetro (o variable aleatoria, por ejemplo  $Tox$ ) en una compuerta y la variación que este mismo parámetro tiene en otra compuerta, separadas cierta distancia entre sí. En la figura 4.5 se muestra una cadena de tres inversores que hacen parte de un layout y las distancias entre ellos. Debido a que los inversores número 1 y 2 se encuentran más cerca, es de esperar que el grado de correlación

entre ellos dos sea más alto que el grado de correlación entre 1 y 3, lo que indicaría que se puede esperar que el valor que toma la variable aleatoria de los parámetros eléctricos en el inversor 1 es más cercano al valor en el inversor 2 que al valor en el inversor 3.

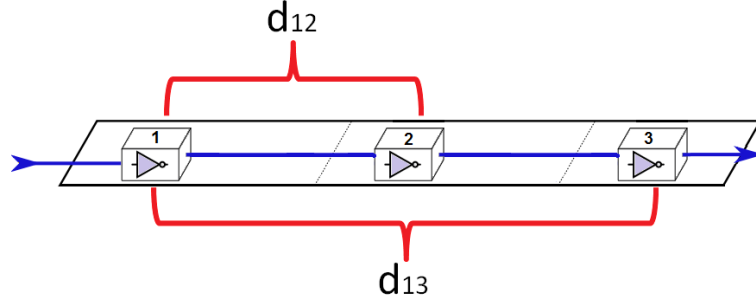


Figura 4.5: Correlación espacial entre compuertas

El grado de correlación espacial puede ser descrito mediante un modelo analítico que ajuste el comportamiento físico medido en una oblea como el desarrollado en [36]. En este trabajo, proponen que una función válida para obtener el grado de correlación entre dos parámetros de los transistores en dos compuertas A y B separadas cierta distancia ( $d$ ), es la mostrada en la ecuación 4.2.

$$\rho_{AB} = \exp(-d_{AB}/CD) \quad (4.2)$$

Donde  $CD$  es la distancia de correlación (*Correlation Distance*), un parámetro que regula la tasa de decaimiento de la correlación respecto de la distancia. Para valores de  $CD$  grandes, la tasa de decaimiento será lenta indicando que parámetros de transistores muy lejanos entre sí mantienen cierto grado de correlación, un comportamiento similar a las componentes  $D2D$  de variación. Para valores de  $CD$  pequeños, la tasa de decaimiento será rápida, haciendo que parámetros de transistores muy alejados no tengan un grado de correlación significativo, similar al comportamiento de las componentes  $WID$  de variación.

Entonces, para calcular el grado de correlación entre dos parámetros, considerando variaciones de proceso a nivel  $WID$  y  $D2D$  la anterior ecuación puede



ser expresada como:

$$\rho_{AB} = K_{WID} \cdot \exp\left(\frac{-d_{AB}}{CD_{WID}}\right) + K_{D2D} \cdot \exp\left(\frac{-d_{AB}}{CD_{D2D}}\right) \quad (4.3)$$

Donde  $K_{WID}$  y  $K_{D2D}$  indican el porcentaje de la variación total asignado a las componentes  $WID$  y  $D2D$  respectivamente y  $CD_{WID}$  y  $CD_{D2D}$  son las distancias de correlación para cada uno de los tipos de variación.  $CD_{D2D}$  tiene un valor alto, ya que parámetros de dispositivos muy alejados, incluso en otro dado, presentarán algún grado de correlación. Por lo tanto, al evaluar el exponencial en el segundo término de la ecuación anterior, sobre dos compuertas que pertenecen a un mismo dado (por ejemplo en un path) el valor obtenido será cercano a la unidad, indicando que las variaciones a nivel de dado a dado de las compuertas son casi las mismas.

## 4.2. Modelado del NBTI considerando variaciones de proceso

La figura 4.6 muestra mediciones de la degradación de la frecuencia de operación de un circuito realizadas en la tecnología de  $65nm$ . Como puede observarse, la degradación obtenida depende de las condiciones de operación ( $V_{dd}$  y  $T$ ). Además, se observa que la variabilidad de la degradación disminuye con el tiempo. Entonces, el envejecimiento interactúa con las variaciones de proceso y debe considerarse su efecto en la variabilidad en un análisis estadístico.

Las variaciones de proceso, que a diferencia del NBTI son variaciones de tipo estáticas (no se modifican con el tiempo), influyen en la cantidad de degradación obtenida, siendo el parámetro de  $V_{th}$  inicial ( $t = 0$ ) uno de los más influyentes [5]. La figura 4.7 muestra que el cambio en el tiempo del voltaje umbral debido a NBTI es diferente según la variación en el  $V_{th}$  inicial debido a proceso ( $\Delta V_{th_{PV}}$ ). La importancia de considerar estadísticamente el envejecimiento por NBTI radica en

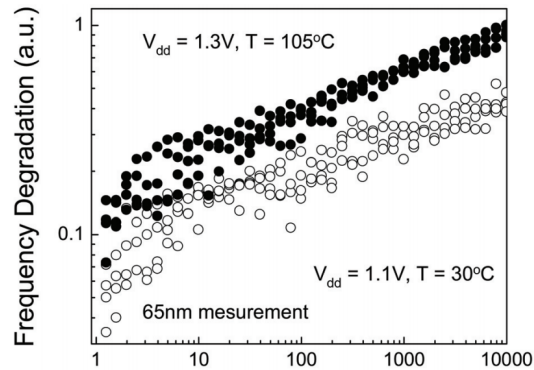


Figura 4.6: Degradación de la frecuencia en función del tiempo. [14]

que el tiempo de vida esperado de un circuito digital podría ser menor al esperado si se considera sólo el valor medio inicial del  $V_{th}$ .

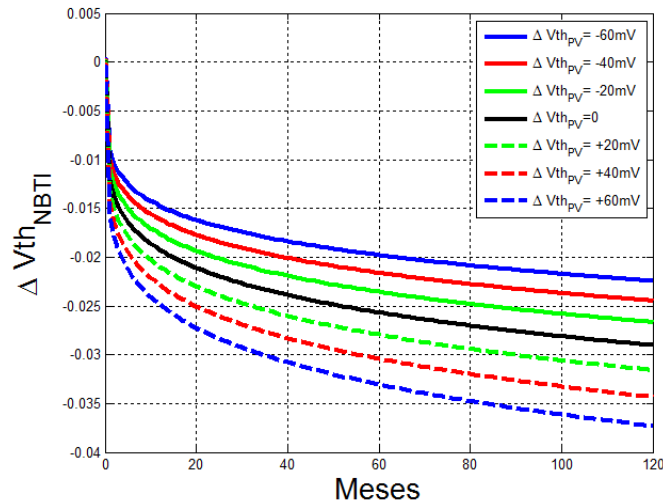


Figura 4.7: Desviación del  $V_{th}$  debido a NBTI para diferentes valores de desviación en el  $V_{th}$  inicial

Para considerar el efecto que tiene las variaciones de proceso del voltaje umbral en la degradación por NBTI, en [14] y [15] proponen linealizar el LTDM (ver ecuación 4.4) mediante el concepto de sensibilidad:

$$\begin{aligned}\Delta Vth_{NBTI} &= \Delta Vth_{NBTI,nom} + S_{Vth}^{\Delta Vth_{nbtI}} \cdot \Delta Vth_{PV} \\ \Delta Vth_{NBTI} &= A \cdot \alpha^n \cdot t^n - A \cdot \alpha^n \cdot t^n \cdot S_v \cdot \Delta Vth_{PV}\end{aligned}\tag{4.4}$$

En esta ecuación,  $S_{Vth}^{\Delta Vth_{nbtI}}$  es la sensibilidad de la degradación del Vth debido a NBTI respecto al voltaje umbral inicial, A y Sv pueden calcularse a partir del LTDM y son constantes de ajuste que dependen de los parámetros tecnológicos, además de la temperatura, el período de reloj y el Vth inicial nominal. El signo negativo en el segundo término indica que la pendiente de la degradación del Vth por NBTI es negativa. Como se puede ver en la figura 4.7 de éste capítulo o 2.10(b) del capítulo 2, un incremento en el Vth inicial de un transistor PMOS (menos negativo), genera una mayor degradación. Esta pendiente, que puede ser interpretada como la sensibilidad de la degradación por NBTI a variaciones de proceso, cambia en el tiempo como se observa en la figura 4.8 y es este cambio en la sensibilidad de la variación del Vth por NBTI el que implica que la varianza de la degradación sea menor con el paso del tiempo, mitigando el efecto de las variaciones de proceso en la degradación.

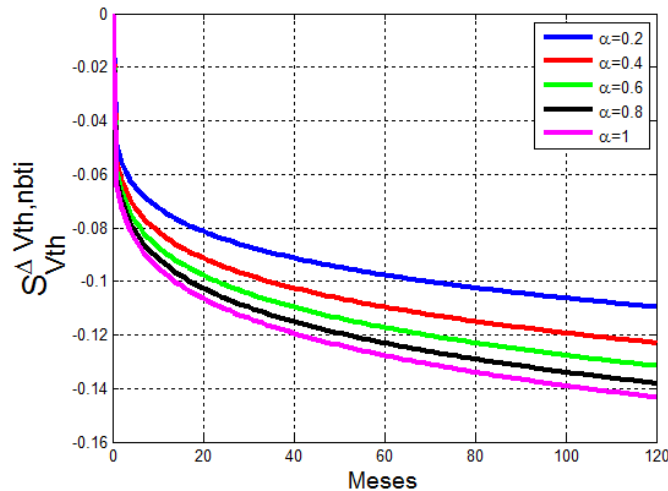


Figura 4.8: Sensibilidad de la degradación por NBTI a variaciones de proceso en el Vth.

La variación total del voltaje umbral de un transistor PMOS, considerando

efectos de variaciones de proceso y envejecimiento debido a NBTI de forma simultánea, puede calcularse como la suma de cada una de las componentes [14]:

$$\begin{aligned}
 \Delta Vth_{total} &= \Delta Vth_{PV} + \Delta Vth_{NBTI} \\
 \Delta Vth_{total} &= \Delta Vth_{PV} + A \cdot \alpha^n \cdot t^n - A \cdot \alpha^n \cdot t^n \cdot S_v \cdot \Delta Vth_{PV} \quad \text{reemplazando} \\
 \Delta Vth_{total} &= A \cdot \alpha^n \cdot t^n + (1 - A \cdot \alpha^n \cdot t^n \cdot S_v) \cdot \Delta Vth_{PV} \quad \text{agrupando términos}
 \end{aligned} \tag{4.5}$$

En la tercera ecuación de (4.5), el primer término corresponde a la desviación del Vth debido a NBTI bajo condiciones nominales (sin variaciones de proceso) y se le llamará  $\Delta Vth_{nbt,i,n} = A \cdot \alpha^n \cdot t^n$ , el segundo término muestra la interacción entre las variaciones de proceso y el efecto de NBTI. Al valor  $AF = 1 - A \cdot \alpha^n \cdot t^n \cdot S_v$  se le llamará en adelante factor de envejecimiento (*Aging Factor*), la figura 4.9 muestra cómo disminuye este factor a través del tiempo. Para  $t = 0$ , el factor de envejecimiento tiene el valor de la unidad y el envejecimiento nominal es de cero, entonces la variación en el voltaje umbral se debe únicamente al efecto de variaciones de proceso. A medida que el tiempo pasa ( $t > 0$ ), La degradación nominal aumenta mientras que factor de envejecimiento disminuye debido a que la sensibilidad de la degradación por NBTI disminuye, esto puede ser visto como una reducción de la variación total del voltaje umbral ocasionada por el efecto de NBTI. De esta forma, estadísticamente hablando, el impacto que tiene el NBTI es el de incrementar el valor medio del voltaje umbral a la vez que disminuye la varianza del mismo [14].

Teniendo en cuenta que las variaciones de proceso en el Vth tienen una componente sistemática y otra completamente aleatoria,  $\Delta Vth_{PV} = \Delta Vth_{sys} + \Delta Vth_{rand}$ . La variación total del voltaje umbral de un transistor PMOS puede escribirse como sigue [14]:

$$\Delta Vth_{total} = \Delta Vth_{nbt,i,nom} + AF(\Delta Vth_g + \Delta Vth_l) \tag{4.6}$$

Donde  $\Delta Vth_g$  y  $\Delta Vth_l$  corresponden a las variaciones globales y locales en el

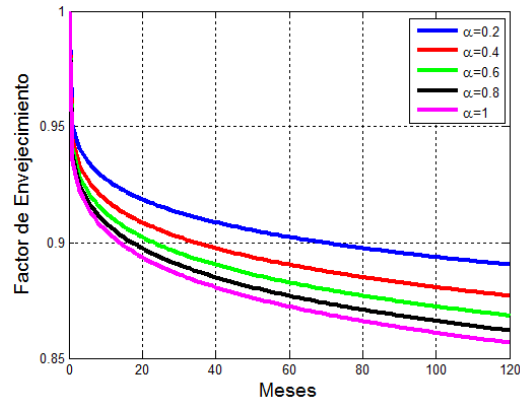


Figura 4.9: Cambio del factor de envejecimiento con el tiempo.

voltaje umbral, respectivamente.

### 4.3. Valor medio y varianza de retardo de una compuerta y covarianza de dos compuertas digitales

Habiendo definido anteriormente el impacto que tiene el envejecimiento por NBTI y su relación con las variaciones de proceso en la variación del  $V_{th}$  total, es posible aplicar el modelo lineal de retardo de compuerta para evaluar el valor medio y la varianza de la PDF de una compuerta digital así como la covarianza entre dos compuertas.

Al considerar el efecto del NBTI en la varianza del retardo, es importante tener en cuenta que a diferencia de variaciones intra die globales en los parámetros de  $W$ ,  $L$  y  $T_{ox}$  que se pueden asumir completamente correlacionadas espacialmente y con el mismo valor de varianza en cada transistor, para el caso del  $V_{th}$ , la varianza de cada transistor puede ser diferente y debe considerarse el efecto de cada uno por separado.

El modelo estadístico lineal de una compuerta que presenta variaciones en los

### 4.3. VALOR MEDIO Y VARIANZA DE RETARDO DE UNA COMPUERTA Y COVARIANZA DE DOS COMPUERTAS DIGITALES

---

parámetros  $W$ ,  $L$ ,  $Tox$  y  $V_{th}$  de los transistores (el cual fue explicado en la sección anterior), permite representar el retardo de una compuerta digital ( $D_i$ ) como:

$$D_i = D_{n,i} + S_{W_i} \cdot \Delta W_i + S_{L_i} \cdot \Delta L_i + S_{Tox_i} \cdot \Delta Tox_i + \sum_k S_{V_{th},ki} \cdot \Delta V_{th_{total},ki} \quad (4.7)$$

Donde  $D_{n,i}$  es el valor de retardo medio de la compuerta cuando no presenta envejecimiento ni variaciones de proceso,  $S_{W_i}$ ,  $S_{L_i}$ ,  $S_{Tox_i}$  son las sensibilidades del retardo a variaciones globales en el parámetro correspondiente. Debido a que los transistores en la compuerta están altamente correlacionados, es posible asumir que la variación global en cada uno está modelado por la misma variable con desviación  $\Delta W$ ,  $\Delta L$  o  $\Delta Tox$  respectivamente y por tanto sólo se tiene un único valor de sensibilidad con respecto a cada parámetro. Para el caso del  $V_{th}$ , la variación en cada transistor queda modelada por una variable diferente debido al NBTI y al RDF, por tanto su efecto se tiene en cuenta sumando por separado la contribución de cada uno de los  $k$  transistores que operan en la carga o descarga del nodo de salida de la compuerta.

Reemplazando la ecuación 4.6 en 4.7 y separando términos, se tiene:

$$D_i = D_{nom,i} + S_{W_i} \cdot \Delta W_i + S_{L_i} \cdot \Delta L_i + S_{Tox_i} \cdot \Delta Tox_i + \sum_k S_{V_{th},ki} \cdot \Delta V_{th_{nbt},nom,ki} + \sum_k S_{V_{th},ki} \cdot AF_{ki} \cdot (\Delta V_{th_{gki}} + \Delta V_{th_{lki}}) \quad (4.8)$$

Los valores de la primera sumatoria  $S_{V_{th},ki} \cdot \Delta V_{th_{nbt},nom,ki}$  no depende de variaciones de proceso y pueden agregarse al retardo nominal, este término corresponde al aumento de retardo debido al efecto de NBTI bajo condiciones nominales. Entonces, el retardo medio de propagación de una compuerta digital que presenta envejecimiento debido a NBTI es el retardo nominal inicial de la compuerta, más la contribución al retardo de la degradación de cada uno de los transistores en la compuerta:

$$\mu_{age,i} = D_{nom,i} + \sum_k S_{V_{th},ki} \cdot \Delta V_{th_{nbt},nom,ki} \quad (4.9)$$

### 4.3. VALOR MEDIO Y VARIANZA DE RETARDO DE UNA COMPUERTA Y COVARIANZA DE DOS COMPUERTAS DIGITALES

---

De esta forma, el modelo lineal de compuerta queda dado por la siguiente expresión:

$$D_{gate} = \mu_{age,i} + S_{Wi} \cdot \Delta Wi + S_{Li} \cdot \Delta Li + S_{Toxi} \cdot \Delta Toxi + \sum_k S_{Vth,ki} \cdot AF_{ki} \cdot \Delta Vth, gki + \sum_k S_{Vth,ki} \cdot AF_{ki} \cdot \Delta Vth, lki \quad (4.10)$$

Para obtener la covarianza entre el retardo de dos compuertas digitales (i,j), se puede aplicar la siguiente propiedad de la covarianza entre dos funciones lineales de variables aleatorias:

$$COV(a_1 \cdot X_1 + a_2 \cdot X_2 + \dots + a_n \cdot X_n; b_1 \cdot Y_1 + b_2 \cdot Y_2 + \dots + b_n \cdot Y_n) = \sum_i \sum_j a_i \cdot b_j \cdot COV(X_i, Y_j) \quad (4.11)$$

Teniendo en cuenta que  $COV(X_i, Y_j) = \rho \cdot \sigma_{X_i} \cdot \sigma_{Y_j}$ , que la correlación entre componentes aleatorias sólo existe cuando se analiza exactamente el mismo transistor de la misma compuerta (covarianza entre una compuerta y ella misma) y asumiendo que sólo existe correlación entre parámetros del mismo tipo (W con W, Tox con Tox, etc.) y que entre diferentes parámetros la correlación es de cero, la varianza de una compuerta puede calcularse al obtener la covarianza de ésta, con ella misma:

$$\begin{aligned} COV(D_i, D_j) &= VAR(D_i) && \text{si } i=j \\ &= S_{Wi} \cdot S_{Wj} \cdot \rho_{ij} \cdot \sigma_{Wi} \cdot \sigma_{Wj} \\ &\quad + S_{Li} \cdot S_{Lj} \cdot \rho_{ij} \cdot \sigma_{Li} \cdot \sigma_{Lj} \\ &\quad + S_{Toxi} \cdot S_{Toxj} \cdot \rho_{ij} \cdot \sigma_{Toxi} \cdot \sigma_{Toxj} \\ &+ \sum_{ki} \sum_{kj} S_{Vth,ki} \cdot S_{Vth,kj} \cdot \rho_{ij} \cdot \sigma_{Vth,gki} \cdot AF_{ki} \cdot \sigma_{Vth,gkj} \cdot AF_{kj} \\ &\quad + \sum_k S_{Vth,ki}^2 \cdot \sigma_{Vth,lki}^2 \cdot AF_{ki}^2 \end{aligned} \quad (4.12)$$

En esta ecuación,  $S_W, S_L, S_{Tox}$  son las sensibilidades del retardo de la compuerta indicada al parámetro correspondiente,  $\sigma_W, \sigma_L, \sigma_{Tox}$  son las desviaciones

### 4.3. VALOR MEDIO Y VARIANZA DE RETARDO DE UNA COMPUERTA Y COVARIANZA DE DOS COMPUERTAS DIGITALES

---

estándar del parámetro correspondiente en la compuerta indicada,  $S_{Vth,ki}$  y  $S_{Vth,kj}$  son las sensibilidades de retardo debido a variaciones en el  $V_{th}$  del transistor  $ki$  en la compuerta  $i$  y al transistor  $kj$  en la compuerta  $j$ ,  $\sigma_{Vth,gki}$  y  $\sigma_{Vth,gkj}$  son las desviaciones estándar del  $V_{th}$  de los transistores correspondientes a nivel intradie global,  $\rho_{ij}$  es la correlación espacial entre las compuertas (en este caso, como  $i = j$  sería de la unidad),  $AF_{ki}$  y  $AF_{kj}$  es el factor de envejecimiento de los transistores  $ki$  en la compuerta  $i$  y  $kj$  en la compuerta  $j$  respectivamente y  $\sigma_{Vth,lki}$  es la desviación estándar de las variaciones puramente aleatorias debido al efecto de RDF y se obtiene de la siguiente expresión:

$$\sigma_{Vth,l} = \sigma_{Vth,g} \sqrt{\frac{W_{eff}L_{eff}}{W_{min}L_{min}}} \quad (4.13)$$

En la ecuación 4.12 la última sumatoria corresponde al efecto de las componentes puramente aleatorias, como se puede detallar, en este caso se recorre cada uno de los transistores en la compuerta ( $i$  o  $j$  es la misma compuerta) y se suma su componente aleatoria.

Cuando se evalúa la covarianza entre dos compuertas diferentes ( $i \neq j$ ), no existe ningún tipo de relación entre las componentes puramente aleatorias y la ecuación 4.11 produce:

$$\begin{aligned} COV(D_i, D_j) = & S_{Wi} \cdot S_{Wj} \cdot \rho_{ij} \cdot \sigma_{Wi} \cdot \sigma_{Wj} && \text{si } i \neq j \\ & + S_{Li} \cdot S_{Lj} \cdot \rho_{ij} \cdot \sigma_{Li} \cdot \sigma_{Lj} \\ & + S_{Toxi} \cdot S_{Toxj} \cdot \rho_{ij} \cdot \sigma_{Toxi} \cdot \sigma_{Toxj} \\ & + \sum_{ki} \sum_{kj} S_{Vth,ki} \cdot S_{Vth,kj} \cdot \rho_{ij} \cdot \sigma_{Vth,ki} \cdot AF_{ki} \cdot \sigma_{Vth,kj} \cdot AF_{kj} \end{aligned} \quad (4.14)$$

En este caso, el valor de correlación espacial no es de la unidad ya que las compuertas  $i$  y  $j$  deben estar ubicadas en diferentes posiciones en el layout. En este caso se debe aplicar la ecuación 4.3 para determinar el grado de correlación existente.



## 4.4. Valor medio, varianza del retardo de un path y covarianza entre paths

El valor medio de un path y la covarianza entre paths se calculan asumiendo, al igual que se hizo a nivel compuerta, que el retardo del path es una función lineal (la suma) de los retardos individuales de cada una de las N compuertas en el path:

$$D_{path} = SUM(D_1 + D_2 + \dots + D_N) \quad (4.15)$$

El valor medio de retardo del path será la suma del valor medio de retardo de cada compuerta (ecuación 4.9), considerando el envejecimiento bajo condiciones nominales de los parámetros:

$$\begin{aligned} \mu_{p,age} &= \sum_{i=1}^N \mu_{age,i} \\ &= \sum_{i=1}^N (D_{nom} + \sum_k S_{Vthg,k} \cdot \Delta Vth_{nbt_i,nom,k})_i \end{aligned} \quad (4.16)$$

La covarianza entre dos paths A y B, por ejemplo como los mostrados en la figura 4.10(a) se calcula aplicando igualmente la propiedad ya mencionada en la ecuación 4.11 de los modelos lineales. Aplicando esta propiedad se llega a que la expresión para la covarianza entre los dos paths es:

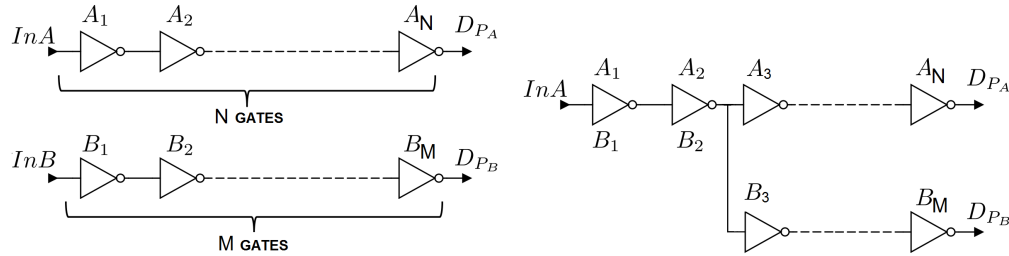
$$cov(D_{PA}, D_{PB}) = \sum_{i=1}^N \sum_{j=1}^M cov(D_{iA}, D_{jB}) \quad (4.17)$$

Donde  $cov(D_{iA}, D_{jB})$  es la covarianza entre la compuerta i en el path A y la compuerta j en el path B. El valor de covarianza entre las dos compuertas se calcula por la ecuación 4.14. Cuando los paths comparten compuertas como los mostrados en la figura 4.10(b), en alguno de los sumandos de la ecuación 4.17, la compuerta i evaluada en el path A es la misma compuerta j en el path B, sólo en estos sumandos debe aplicarse la ecuación 4.12 y considerar el efecto de las

#### 4.4. VALOR MEDIO, VARIANZA DEL RETARDO DE UN PATH Y COVARIANZA ENTRE PATHS

---

variaciones aleatorias. Cuando se calcula la covarianza entre un path y él mismo, todas las compuertas se comparten y se obtiene la varianza del retardo del path.



(a) Dos paths que no comparten compuertas (b) Dos paths que comparten compuertas

Figura 4.10: Dos paths a analizar

Al evaluar la ecuación 4.17 es necesario conocer el grado de correlación espacial entre todas las compuertas presentes en los paths analizados. A cada término de la sumatoria de esta ecuación, le corresponde un valor de correlación espacial, el cual es obtenido a partir de la siguiente matriz:

$$\rho_{ij} = \begin{pmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1N} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{M1} & \rho_{M2} & \dots & \rho_{MN} \end{pmatrix} \quad (4.18)$$

Vale la pena mencionar que al evaluar la varianza de un path por medio de la ecuación 4.17 los valores de correlación en la diagonal de la ecuación 4.18 es de la unidad.

##### 4.4.1. Correlación entre paths

El grado de correlación que existe entre el retardo de dos paths puede calcularse mediante la siguiente expresión:

$$\rho_{AB} = \frac{COV(D_{pA}, D_{pB})}{\sigma_{DpA} \cdot \sigma_{DpB}} \quad (4.19)$$

#### 4.4. VALOR MEDIO, VARIANZA DEL RETARDO DE UN PATH Y COVARIANZA ENTRE PATHS

---

La correlación entre paths indica que tan parecidas son las variaciones de retardo en un path con respecto a la de otro y como se verá en el siguiente capítulo resulta de gran utilidad para seleccionar adecuadamente los paths a monitorear en un circuito adaptivo.

## Capítulo 5

# Metodología de selección de caminos críticos a monitorear en un circuito adaptivo tolerante a envejecimiento por NBTI

En el diseño de circuitos digitales adaptivos, resulta de gran importancia la correcta identificación de aquellos paths que requieren ser monitoreados, con el objetivo de minimizar el consumo de área y potencia que implica insertar los sensores correspondientes. Además debe garantizarse que los sensores sólo detecten transiciones tardías generadas por los paths críticos del circuito para evitar falsas detecciones. Sin embargo, el criterio de selección de estos paths debe hacerse de forma estadística [34], por lo que requiere el conocimiento de la posición final de las compuertas en el layout para estimar la correlación espacial entre ellas. Esto hace que insertar los sensores implique modificar el layout realizado, lo que resulta en un retroceso en el flujo de diseño.

Se ha desarrollado una herramienta que parte de alto nivel para realizar la selección de caminos críticos a monitorear en un circuito adaptivo. Además, se

---

busca estimar a alto nivel, el conjunto de paths a monitorear, de forma que se reduzcan las modificaciones requeridas en etapas posteriores.

La estrategia implementada en esta herramienta se ilustra en la figura 5.1. Partiendo de la descripción a alto nivel del circuito a analizar, es decir, la información del layout en esta fase es desconocida, se detectan todos los caminos de señal o paths existentes desde una entrada principal a una salida principal. Sobre estos paths se realiza un prefiltrado escogiendo sólo a aquellos que bajo las peores condiciones de envejecimiento tendrían alguna posibilidad de tener un retardo mayor al del path más lento del circuito (*Longest Critical Path, LCP*). Luego, se realiza una predicción o selección de los paths que tentativamente deben ser monitoreados mediante un análisis estadístico de tiempos estimado sin información del layout. Una vez determinado este conjunto predicho de paths a monitorear, se debe realizar el primer prototipo de layout, incluyendo la circuitería adicional de los sensores en los paths seleccionados. Luego, incluyendo la información de las distancias entre compuertas, es posible realizar un análisis fino o real para detectar aquellos paths que fueron incorrectamente seleccionados y aquellos que no fueron seleccionados en la etapa previa. De esta forma se espera reducir la cantidad de modificaciones requeridas en el layout si no se hubiera realizado la etapa de predicción y así facilitar el proceso de insertar sensores de envejecimiento en el flujo de diseño. Además, es planteada una estrategia que consiste en penalizar el período de reloj, introduciendo una banda de guarda para tolerar aquellos paths que no fueron identificados en el análisis a alto nivel, reduciendo al máximo las modificaciones requeridas.

## 5.1. DETECCIÓN DE PATHS TOPOLÓGICOS Y CÁLCULO DE SU RETARDO MEDIO DE PROPAGACIÓN

---

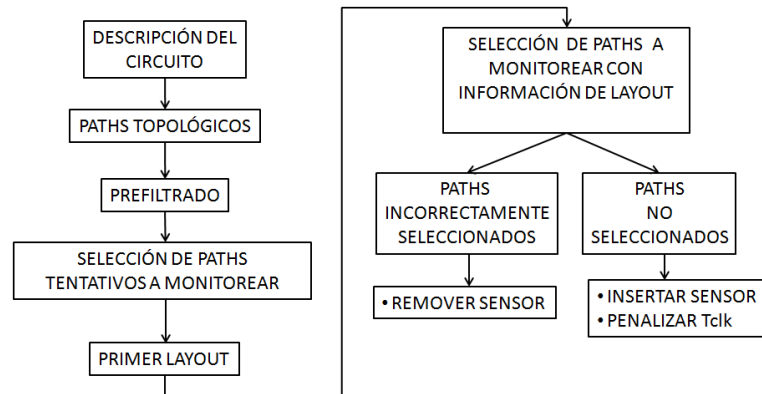


Figura 5.1: Diagrama de flujo de la metodología propuesta

### 5.1. Detección de paths topológicos y cálculo de su retardo medio de propagación

Para detectar los paths topológicos se parte de una compuerta que maneje una salida principal (fin del path) y se recorre hacia atrás el trayecto de una posible señal digital, manteniendo como prioridad que la señal entre por el primer nodo de fan-in. Una vez encontrada una entrada principal, se ha detectado un path y mediante una bandera que consiste en marcar ciertos nodos por donde no puede volver a pasar la señal, se indica que ese camino recorrido ya ha sido detectado para evitar repetirlo. Entonces, se retorna a la salida principal para repetir el procedimiento cumpliendo las restricciones de las marcas. Una vez todos los paths que terminan en una salida principal se han encontrado, se cambia hacia la siguiente salida principal y se repite el procedimiento.

La figura 5.2 ilustra la metodología de detección de los dos primeros paths para un circuito prueba. Al detectar el path azul, se ha marcado la entrada principal correspondiente al nodo 1 de la compuerta  $X1$ , por tanto para el segundo path, la señal se ha propagado por el nodo 2 de la misma y ha pasado por el nodo 1 de la compuerta  $X0$ . El tercer path detectado, marcaría el segundo nodo de la compuerta  $X0$ , haciendo que se marque a su vez el segundo nodo de la compuerta

$X1$  y el primero de la compuerta  $X2$  ya que los nodos de las compuertas precedentes estarían completados. En este momento se pasaría a repetir el análisis para el segundo nodo de la compuerta  $X2$ .

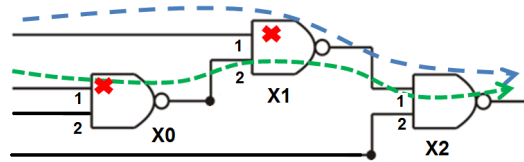


Figura 5.2: Detección de paths en un circuito prueba

Una vez detectados todos los caminos existentes en el circuito, se evalúa la capacitancia de carga en cada uno de los nodos del circuito sumando la capacitancia de entrada de cada una de las compuertas manejadas por un mismo nodo.

Con esta información y asignando un valor de Slew-Rate y tipo de transición (alto a bajo, bajo a alto) de la señal digital a las entradas principales del circuito, se evalúan los polinomios de retardo, slew rate a la salida y sensibilidad correspondientes a una señal que se propaga por los nodos de cada path encontrado. El slew rate a la salida de una compuerta es utilizado como slew rate a la entrada de la compuerta subsiguiente.

## 5.2. Consideración de máxima degradación en los paths debido a NBTI

Uno de los principales cuellos de botella en los análisis de envejecimiento, es que no es posible conocer con certeza la probabilidad de las señales en las entradas principales para cualquier circuito en general [21], haciendo difícil determinar con exactitud la degradación del retardo de un path en específico. Por esto, es necesario trabajar bajo condiciones conservativas como suponer condición de degradación estática.

Para abordar este inconveniente, en [21] y [34] proponen una estrategia para

## 5.2. CONSIDERACIÓN DE MÁXIMA DEGRADACIÓN EN LOS PATHS DEBIDO A NBTI

---

determinar el máximo incremento de retardo de un path debido a NBTI. Como se muestra en la figura 5.3, los autores suponen que sólo conmuta la señal de entrada del path bajo análisis, asignando a cada una de las entradas laterales de las compuertas (*Side Inputs*), una probabilidad de degradación de  $\alpha = 0$  para las compuertas AND o NAND (ya que para que se sensibilice a un nodo de entrada, los demás deben estar en 1 lógico) y de  $\alpha = 1$  para las compuertas tipo OR o NOR (0 lógico en las demás entradas para que se sensibilice). De esta forma, el valor de probabilidad de entrada es propagado hacia las compuertas subsiguientes y será alternado entre los valores de  $\alpha$ ,  $1 - \alpha$ ,  $\alpha$ , etc. Debido a que se asume que las demás entradas se encuentran en un valor estático, el valor de probabilidad propagado corresponde a la probabilidad de que el transistor de entrada se encuentre bajo envejecimiento por NBTI. De esta forma, se obtiene un retardo de propagación máximo sin necesidad de considerar el peor caso donde cada uno de los transistores se encuentra bajo degradación estática.

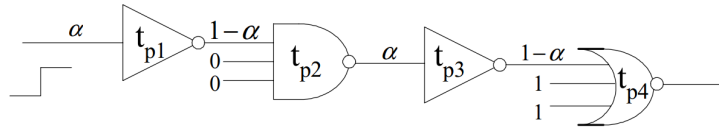


Figura 5.3: Probabilidades de degradación para máximo envejecimiento del path [21]

El valor de  $\alpha$  de entrada al path que genera una máxima degradación se asigna según el tipo de transición supuesto en la entrada principal. La figura 5.4 permite analizar cómo seleccionar este valor. En la cadena de 5 inversores mostrada, si la transición en la entrada es de bajo a alto (figura superior) los inversores en color son aquellos en los que los transistores PMOS están realizando la operación de carga del nodo de la salida, por tanto, serán éstos los que contribuyen a la degradación del retardo de propagación de este tipo de transición, ya que los transistores NMOS no presentan efecto de NBTI (al menos no significativo). Para que estos inversores tengan una mayor degradación, el valor de  $\alpha_{in}$  debe ser bajo



## 5.2. CONSIDERACIÓN DE MÁXIMA DEGRADACIÓN EN LOS PATHS DEBIDO A NBTI

---

(0.1), de forma que el valor de  $\alpha$  de éstos sea alto (0.9). De forma similar puede analizarse el caso en que la transición de entrada es de alto a bajo (figura inferior). La figura 5.5 muestra el retardo obtenido para diferentes valores de  $\alpha_{in}$  en la cadena de 5 inversores analizada.

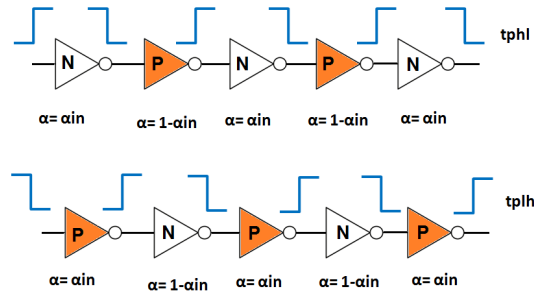


Figura 5.4: Valor de  $\alpha$  en las compuertas con transición de bajo a alto en la salida

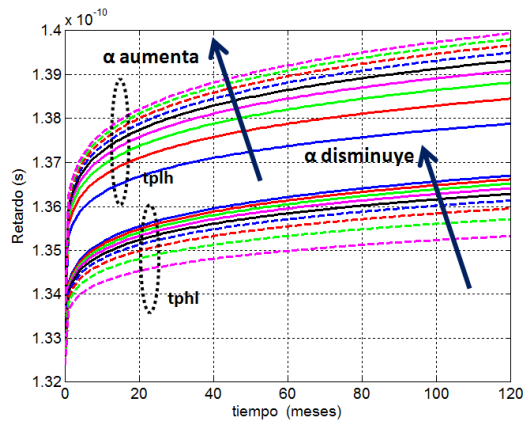


Figura 5.5: Probabilidades de degradación para máximo envejecimiento del path

Entonces, los valores de  $\alpha_{in}$  que deben ser asignados a la señal de entrada del path, para generar una máxima degradación del retardo de propagación son:

$$\begin{aligned}
 \alpha_{in} &= 0.1 & \text{si } \text{input-tran}=0\text{to}1 \\
 \alpha_{in} &= 0.9 & \text{si } \text{input-tran}=1\text{to}0
 \end{aligned}
 \tag{5.1}$$

### 5.3. Prefiltrado de paths

Para disminuir el conjunto de paths que inicialmente se analiza, se realiza un prefiltrado donde se escogen como paths a analizar sólo aquellos cuyo retardo de propagación podrían eventualmente ser mayor al retardo del path más lento (LCP) del circuito. El prefiltrado se realiza estimando un borde superior de la esquina  $3\sigma$  del retardo de cada uno de los  $i$  paths considerando máxima degradación en ellos y comparando este valor con el borde inferior de la esquina  $3\sigma$  del retardo del LCP sin envejecimiento. De esta forma, se seleccionan como paths que podrían ser críticos (se les llamará Near Critical paths NCP) a aquellos que cumplan con la siguiente condición:

$$\mu_i + 3\sigma_{max,i} \geq \mu_{LCP} - 3\sigma_{max,LCP} \quad (5.2)$$

La figura 5.6 muestra el criterio del prefiltrado mencionado.

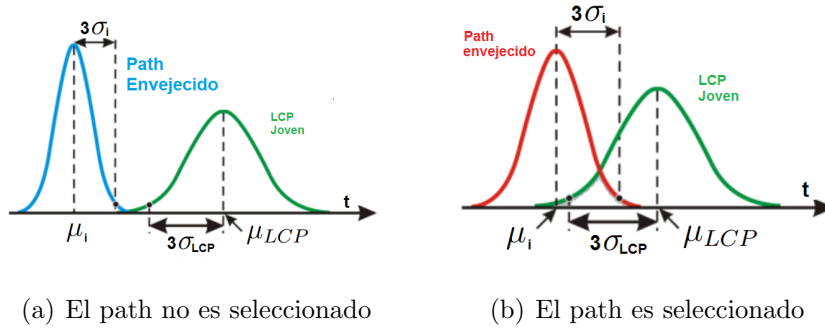


Figura 5.6: Criterio de prefiltrado de paths

Los valores de desviación estándar máximos  $\sigma_{max,i}$  y  $\sigma_{max,LCP}$  se calculan utilizando un valor de correlación de la unidad en todas las posiciones de la matriz de correlación (ecuación 4.18), al aplicar la ecuación 4.14 en la ecuación 4.17. De esta forma se considerarán todos aquellos paths que tengan alguna probabilidad, por mínima que sea, de tener un retardo mayor que el LCP.

## 5.4. Criterio de selección de paths a monitorear

El criterio de selección establecido, toma como referencia la probabilidad de que el retardo de un path NCP bajo máximo envejecimiento por NBTI sea mayor al retardo del LCP sin envejecimiento:

$$P(D_{NCP,age} \geq D_{LCP,fresh}) \quad (5.3)$$

Esta probabilidad puede calcularse por medio de la función de densidad de probabilidad conjunta (ecuación 5.4), la cuál está determinada por los valores de desviación estándar del retardo de cada path y la correlación entre ellos. La probabilidad de interés corresponde a aquella que se obtiene al integrar la función de probabilidad conjunta en la región sombreada, como se muestra en la figura 5.7, la cual muestra un diagrama de contornos de esta función para dos paths.

$$P(D_{PA} \cap D_{PB}) = \frac{1}{2\pi\sigma_{PA}\sigma_{PB}\sqrt{1-\rho_{AB}^2}} \cdot \exp\left[\frac{-Q}{2\cdot(1-\rho_{AB}^2)}\right] \quad (5.4)$$

$$Q = \frac{(Da-\mu_{PA})^2}{\sigma_{PA}^2} - \frac{2\rho_{AB}(Da-\mu_{PA})\cdot(Db-\mu_{PB})}{\sigma_{PA}\sigma_{PB}} + \frac{(Db-\mu_{PB})^2}{\sigma_{PB}^2}$$

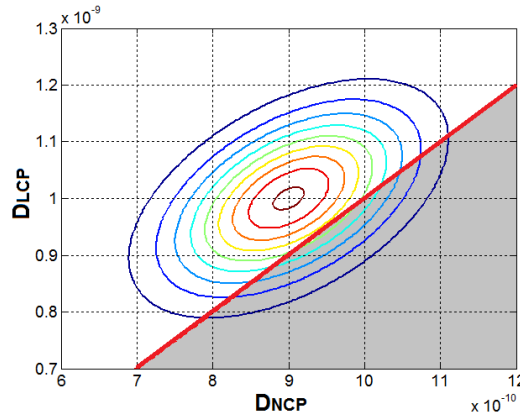


Figura 5.7: Región de integración

Sin embargo, a partir de la función de densidad de probabilidad conjunta, es difícil intuir el comportamiento que tiene la probabilidad de interés en función de

algunas características de los paths, ya que no representa directamente el fenómeno que se desea analizar. Por esto, resulta más conveniente definir una nueva variable aleatoria que dé una mejor perspectiva del problema, esta será la variable aleatoria de diferencia de retardos DR.

### 5.4.1. Variable de diferencia de retardos

La variable aleatoria de diferencia de retardos (DR) entre un NCP y LCP está definida por la ecuación 5.5 y su valor medio y varianza se pueden calcular por las ecuaciones 5.6a y 5.6b, respectivamente. Esta variable puede utilizarse para calcular la probabilidad de que el retardo del NCP bajo envejecimiento por NBTI sea mayor al del LCP sin envejecimiento o fresco, integrando su función de densidad de probabilidad en la región donde toma valores mayores a cero,  $P(DR \geq 0)$ .

$$DR = D_{NCP,age} - D_{LCP,fresh} \quad (5.5)$$

$$\mu_{DR} = \mu_{NCP} - \mu_{LCP} \quad (5.6a)$$

$$\sigma_{DR}^2 = \sigma_{NCP}^2 + \sigma_{LCP}^2 - 2 \cdot cov(D_{NCP}, D_{LCP}) \quad (5.6b)$$

Integrar la función de densidad de probabilidad conjunta o la función de densidad de probabilidad de DR en sus regiones correspondientes proporciona el mismo valor de probabilidad del evento a analizar: “El retardo del NCP con envejecimiento es mayor al retardo del LCP sin envejecimiento”. En otras palabras, DR contiene la misma información de interés que se puede extraer de la función de probabilidad conjunta y además permite visualizar más fácilmente cada una de las componentes que impactan en el valor de probabilidad obtenida. Esta probabilidad que depende de la varianza de DR se ve reducida con el aumento de la

## 5.4. CRITERIO DE SELECCIÓN DE PATHS A MONITOREAR

correlación o covarianza entre paths, como se puede observar de las figuras 5.8 a 5.10.

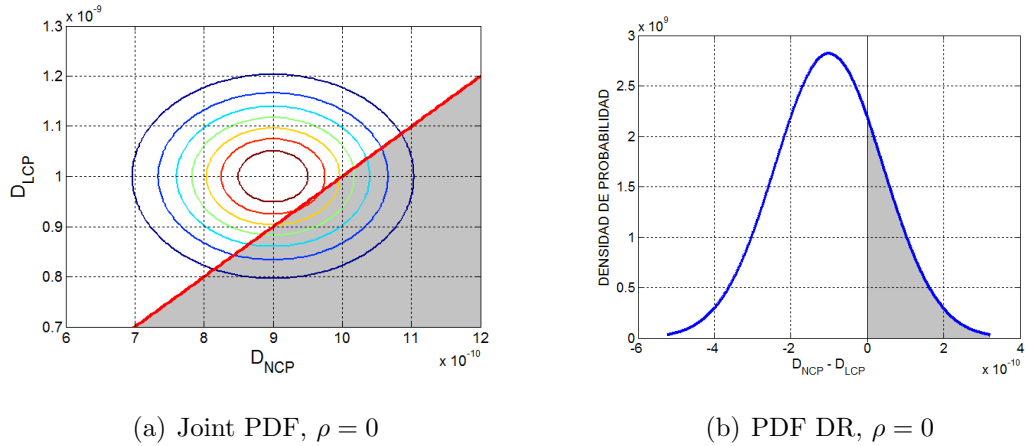


Figura 5.8: Probabilidad de  $D_{NCP} \geq D_{LCP}$

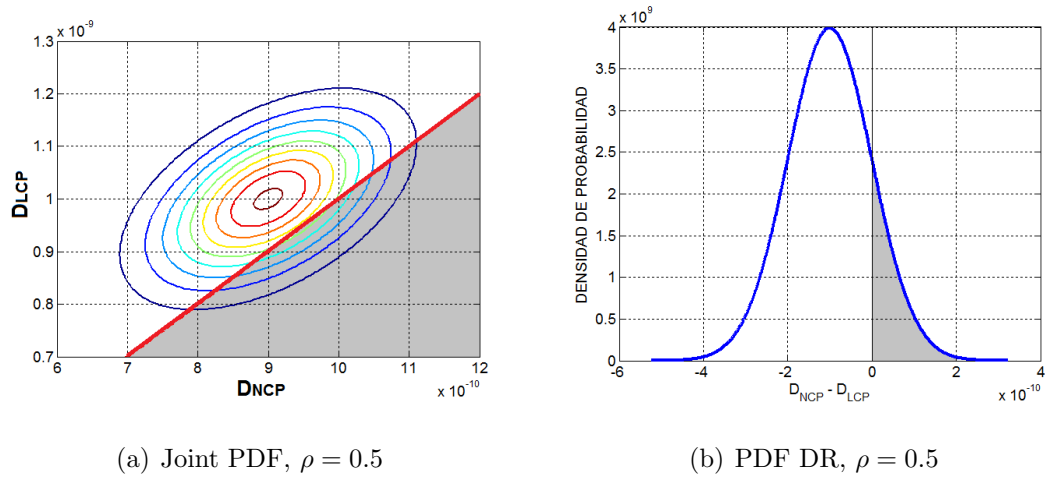


Figura 5.9: Probabilidad de  $D_{NCP} \geq D_{LCP}$

### 5.4.2. Efecto de la correlación estructural en la diferencia de retardos

Una de las primeras características que son evidentes al utilizar la variable de diferencia de retardo en lugar de la función de densidad de probabilidad conjunta

## 5.4. CRITERIO DE SELECCIÓN DE PATHS A MONITOREAR

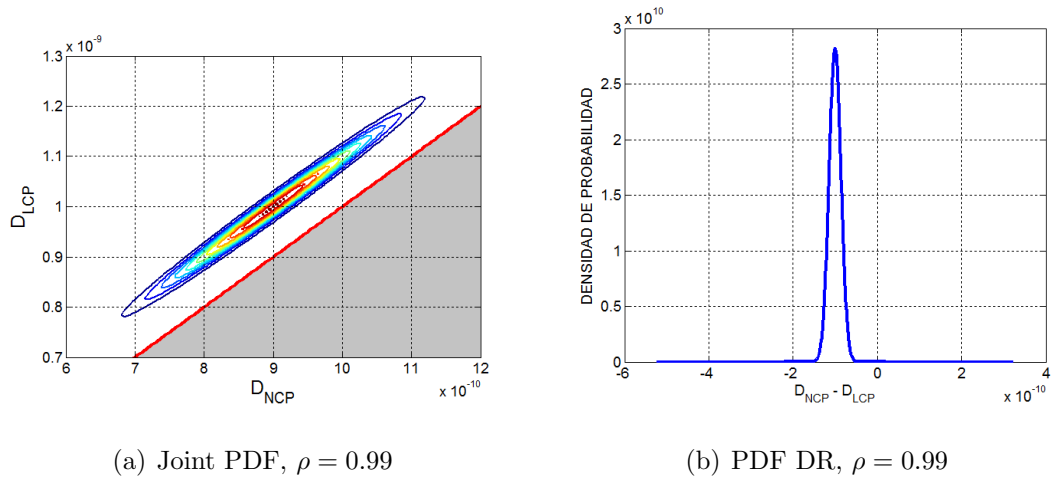


Figura 5.10: Probabilidad de  $D_{NCP} \geq D_{LCP}$

es el efecto que tiene la correlación estructural entre dos paths, por ejemplo los mostrados en la figura 5.11 donde cada uno comparte una compuerta y otra no.

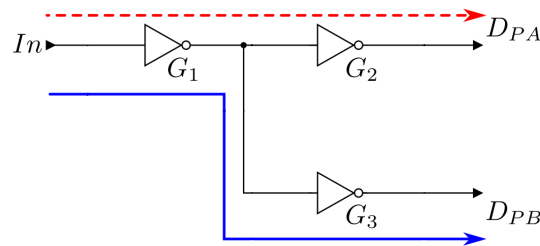


Figura 5.11: Dos paths con correlación estructural

Teniendo en cuenta que la covarianza entre una compuerta y ella misma corresponde a su varianza ( $COV(X, X) = VAR(X)$ ), la varianza de DR puede calcularse como sigue:

$$\begin{aligned}
 \sigma_{DR}^2 &= \sigma_{PA}^2 + \sigma_{PB}^2 - 2 \cdot COV(PA, PB) \\
 &= \sigma_{D1}^2 + \sigma_{D2}^2 + 2 \cdot COV(D1, D2) + \sigma_{D1}^2 + \sigma_{D3}^2 + 2 \cdot COV(D1, D3) \\
 &\quad - 2 \cdot COV(D1, D1) - 2 \cdot COV(D1, D3) - 2 \cdot COV(D2, D1) - 2 \cdot COV(D2, D3) \\
 &= \underline{2 \cdot \sigma_{D1}^2} + \sigma_{D2}^2 + \sigma_{D3}^2 + \underline{2 \cdot COV(D1, D2)} + \underline{2 \cdot COV(D1, D3)} \\
 &\quad - \underline{2 \cdot \sigma_{D1}^2} - \underline{2 \cdot COV(D1, D3)} - \underline{2 \cdot COV(D2, D1)} - 2 \cdot COV(D2, D3)
 \end{aligned} \tag{5.7}$$

En el desarrollo anterior, los términos subrayados se eliminan entre sí, estos términos son aquellos que se generan debido a la compuerta compartida por los dos paths, la compuerta  $G1$ . Al eliminar los términos respectivos se tiene:

$$\sigma_{DR}^2 = \sigma_{D2}^2 + \sigma_{D3}^2 - 2 \cdot COV(D2, D3) \tag{5.8}$$

Queda así demostrado que la correlación estructural, no contribuye ni a la desviación estándar ni al valor medio de la diferencia de retardos, es decir, el valor medio y la desviación estándar de DR estarán determinados únicamente por las contribuciones de aquellas compuertas que no son compartidas por los paths. Esto se debe a que como se mostró anteriormente, la varianza de DR disminuye con la correlación entre paths y en las zonas compartidas los paths están completamente correlacionados.

### 5.4.3. Umbral de probabilidad

Para determinar si un path debe ser monitoreado o no, la probabilidad de que su retardo envejecido sea mayor al retardo del LCP debe ser mayor a un umbral de decisión requerido  $\varepsilon_R$ .

$$Monitor = \begin{cases} 1 & \text{si } P(DR \geq 0) \geq \varepsilon_R \\ 0 & \text{si } P(DR \geq 0) < \varepsilon_R \end{cases}$$

Tomando como referencia la probabilidad que hay por fuera de la esquina  $\mu + 3\sigma$  de una función de probabilidad con distribución normal, el valor de  $\varepsilon_R$  se tomó como 0.135 %.

#### 5.4.4. Determinar redundancia de los paths a monitorear

Hasta ahora, en el criterio de selección de paths planteado se ha tomado como referencia el path más lento del circuito. Sin embargo, al introducir sensores en algunos de los NCP, éstos pueden cubrir a su vez otros de los NCP que no son cubiertos por el LCP. Lo anterior puede explicarse mediante el ejemplo de la figura 5.12.

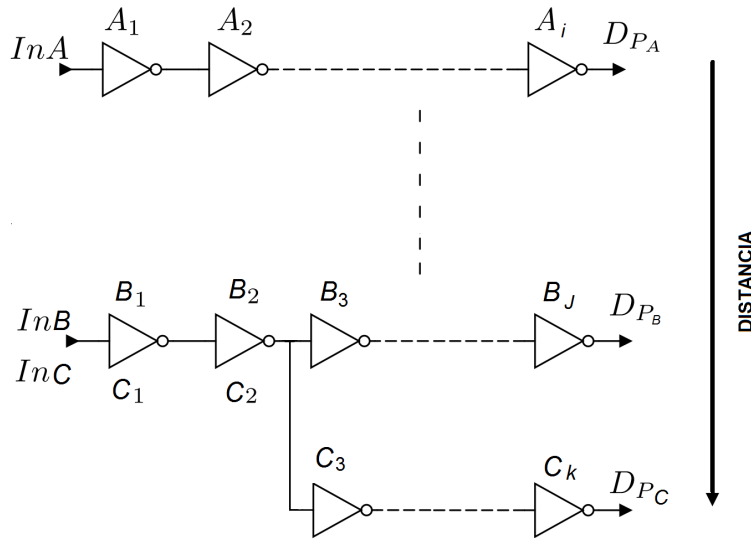


Figura 5.12: Tres paths en un circuito digital

Suponiendo que en este ejemplo el path A es el LCP y que los paths B y C son dos NCP a analizar, y que retardo de los paths tiene la siguiente relación:

$$D_{PA} > D_{PB} > D_{PC} \quad (5.9)$$

Al tomar únicamente como referencia el LCP (A), debido a la baja correlación existente entre éste y los NCP (B y C), puede que tanto B y C se seleccionen para



ser monitoreados. Sin embargo, si B es monitoreado, debido a su alta correlación con C, es poco probable que el retardo de este último sea mayor al retardo de B. Por tanto, realmente no sería necesario monitorear C.

Entonces, es necesario realizar iterativamente el criterio de selección sobre el conjunto de paths seleccionado inicialmente partiendo del LCP y así determinar si alguno de los paths que se ha seleccionado para monitorear, puede ser cubierto por otro que también se esté monitoreando.

La estrategia seguida es la siguiente:

1. Obtener el conjunto de paths a monitorear tomando como referencia el LCP.
2. Rankear el conjunto de paths obtenido según su retardo medio sin envejecimiento ya que éste es el tomado como referencia. La primera posición corresponde al LCP, la segunda al NCP con mayor retardo y así sucesivamente.
3. Considerar ahora el NCP con mayor retardo (posición 2) como si fuera el LCP, considerando que no presenta degradación debido a NBTI y determinar la probabilidad de que el retardo con envejecimiento de los NCP con menor retardo que este (posiciones 3 a N) sea mayor a su retardo. Si la probabilidad es menor que el umbral, puede dejarse de monitorear el NCP correspondiente, ya que será cubierto por el *NCP2*.
4. Realizar el punto anterior para cada uno de los NCP que se continúan monitoreando.

## 5.5. Estimación de varianza de DR sin información de layout para la predicción de paths a monitorear

En esta sección se explicarán los análisis y las consideraciones realizadas para poder estimar el conjunto de paths a monitorear sin información de la correlación espacial del layout.

La probabilidad de que DR de un path NCP tome valores superiores a 0, se calcula a partir de su valor medio y varianza, que pueden obtenerse por las ecuaciones 5.6a y 5.6b, respectivamente. El valor de retardo medio de DR es la resta de los retardos medios de propagación de los paths bajo análisis. Los valores de retardo medio de los paths, están determinados por la librería de compuertas utilizadas, su valor se conoce a alto nivel y no se verá afectado una vez realizado el layout (omitiendo los efectos de interconexiones). Entonces, antes de realizar el layout el parámetro que queda a la incertidumbre ya que depende de la correlación espacial y que se requiere para estimar la probabilidad deseada es la varianza de DR. Para garantizar la cobertura o detección del mayor conjunto de paths a monitorear, la desviación estándar de DR que se utilice para la estimación de la probabilidad debe corresponder a un límite superior de los posibles valores que pueda tomar una vez realizado el layout.

Expandiendo cada uno de los términos de la ecuación 5.6b mediante el uso de la ecuación 4.17 y ordenando se tiene la ecuación general:

$$\sigma_{DR}^2 = \sigma_{NCP}^2 + \sigma_{LCP}^2 - 2 \cdot \text{cov}(D_{NCP}, D_{LCP})$$

$$\sigma_{DR}^2 = \sum_{i=1}^N \sigma_{D_{i,NCP}}^2 + \sum_{i=1}^N \sum_{j=i+1}^N 2 \cdot \text{cov}(D_{i,NCP}, D_{j,NCP})$$

## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

---

$$\begin{aligned}
 & + \sum_{i=1}^M \sigma_{D_{i,LCP}}^2 + \sum_{i=1}^M \sum_{j=i+1}^M 2 \cdot \text{cov}(D_{i,LCP}, D_{j,LCP}) \\
 & - \sum_{i=1}^N \sum_{j=1}^M 2 \cdot \text{cov}(D_{i,NCP}, D_{j,LCP})
 \end{aligned}$$

En la ecuación anterior, los términos sumados corresponden a las varianzas de cada uno de los paths y se calcula por medio de la ecuación 4.17 ( $\text{COV}(X, X) = \text{VAR}(X)$ ), el término restado corresponde al efecto de la covarianza o correlación entre paths (que como se mencionó anteriormente tiende a disminuir la varianza de DR y con ella la probabilidad de selección) y también se calcula por medio de la ecuación 4.17. N y M son el número de compuertas en el NCP analizado y el LCP respectivamente,  $\sigma_{D_{i,NCP}}^2$  y  $\sigma_{D_{i,LCP}}^2$  son la varianza de la i-ésima compuerta en cada uno de los paths correspondientes y se calculan por medio de la ecuación 4.12,  $\text{cov}(D_{i,NCP}, D_{j,NCP})$  y  $\text{cov}(D_{i,LCP}, D_{j,LCP})$  es la covarianza entre las compuertas i y j pertenecientes a cada uno de los respectivos paths y se calcula por medio de la ecuación 4.14 y  $\text{cov}(D_{i,NCP}, D_{j,LCP})$  es la covarianza entre las compuertas i en el NCP y j en el LCP e igualmente se calcula por medio de la ecuación 4.14. Es importante mencionar que al evaluar las ecuaciones 4.12 o 4.14, el factor de envejecimiento AF de las compuertas presentes en el LCP es de la unidad, ya que se consideran sin envejecimiento.

Al analizar esta ecuación podemos ver que la contribución de los términos de varianza de las compuertas individuales ( $\sigma_{D_{i,LCP}}^2$  y  $\sigma_{D_{i,NCP}}^2$ ) será la misma sin importar cómo se ubiquen las compuertas en el layout y este valor depende únicamente de la topología de compuerta y su capacitancia de Fan Out. Por otra parte, sólo dependen de la posición de las compuertas, los términos correspondientes de covarianza entre dos compuertas diferentes.

En un SSTA convencional, partiendo de la información del layout, la covarianza entre el retardo de cualesquiera dos compuerta i y j diferentes en el circuito se calcula por la ecuación 4.14 que se repite a continuación. En esta ecuación  $\rho_{ij}$

## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

---

corresponde al grado de correlación espacial entre las compuertas.

$$\begin{aligned}
 COV(D_{gi}, D_{gj}) &= S_{Wi} \cdot S_{Wj} \cdot \rho_{ij} \cdot \sigma_{Wi} \cdot \sigma_{Wj} \\
 &\quad + S_{Li} \cdot S_{Lj} \cdot \rho_{ij} \cdot \sigma_{Li} \cdot \sigma_{Lj} \\
 &\quad + S_{Toxi} \cdot S_{Toxj} \cdot \rho_{ij} \cdot \sigma_{Toxi} \cdot \sigma_{Toxj} \\
 &\quad + \sum_{ki} \sum_{kj} S_{Vth,ki} \cdot S_{Vth,kj} \cdot \rho_{ij} \cdot \sigma_{Vthi} \cdot AF_{ki} \cdot \sigma_{Vthj} \cdot AF_{kj}
 \end{aligned}$$

Es posible establecer un límite superior e inferior para la varianza del retardo de un path, suponiendo un valor constante de correlación entre las compuertas diferentes en el path, denotado como *correlación global*  $\rho_g$ :

$$\rho_{ij} = \rho_g \tag{5.11}$$

Para el caso en que se tome  $\rho_g = 1$  ( $\rho_g = 0$ ), el valor de varianza del path obtenido será máximo (mínimo), ya que cada uno de los términos de covarianza sumados será también máximo (mínimo). En otras palabras, el valor real de varianza del path está contenido entre el rango de varianzas obtenidas al evaluar la ecuación 4.17 con un valor de correlación espacial global.

Debido a que la desviación estándar de DR depende tanto de la varianza del NCP y del LCP como de la covarianza entre ellos, la correlación global puede utilizarse también para estimar un límite superior de la desviación estándar de DR que permitirá predecir el conjunto de paths a monitorear. Analizando la ecuación 5.6b, puede observarse que una primera aproximación para el caso de mayor desviación estándar de DR es:

- Calcular la varianza del NCP y del LCP suponiendo una correlación espacial global de la unidad, de esta forma la varianza de cada uno será máxima.
- Calcular la covarianza entre el NCP y el LCP suponiendo un valor de correlación espacial global de cero. de esta forma el término de covarianza será mínimo ya que sólo se sumarán las componentes correspondientes a la correlación estructural.

## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

---

Esta forma de aproximación proporciona una sobre estimación de la desviación estándar de DR y del conjunto de paths a monitorear ya que existe una dependencia entre los términos  $\sigma_{D_{NCP}}^2$ ,  $\sigma_{D_{LCP}}^2$  y  $cov(D_{NCP}, D_{LCP})$  de la ecuación 5.6b. La dependencia puede ser vista intuitivamente si se utiliza la correlación entre paths  $\rho_{AB} = cov(D_A, D_B)/\sigma_A\sigma_B$  en la ecuación 5.6b:

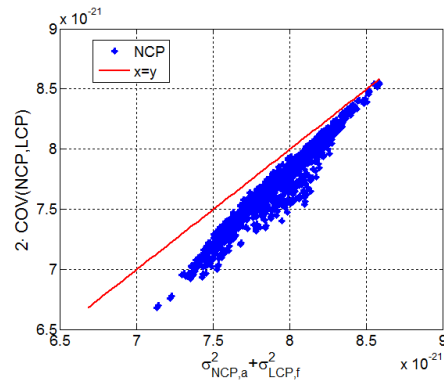
$$\sigma_{DR}^2 = \sigma_{NCP}^2 + \sigma_{LCP}^2 - 2 \cdot \rho_{ip} \cdot \sigma_{NCP} \cdot \sigma_{LCP} \quad (5.12)$$

En la anterior ecuación, independientemente del valor de correlación entre paths existente, un incremento de la desviación estándar de alguno de los paths analizados, también aumenta el término restado. La correlación entre paths únicamente define que tan proporcional es este efecto.

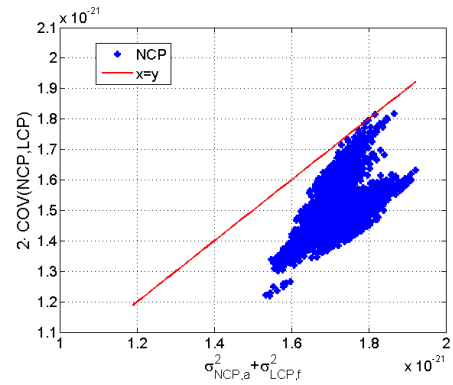
En las figuras 5.13 se muestran los puntos de  $\sigma_{NCP}^2 + \sigma_{LCP}^2$  vs  $2COV(NCP, LCP)$  obtenidos mediante una caracterización con información del layout de algunos circuitos combinatoriales analizados. En la figura sólo se muestran los puntos correspondientes a los paths que son seleccionados para monitorear a partir del LCP, se observa que la relación mencionada anteriormente es evidente. La recta roja define el punto donde ambos términos son iguales, es decir, donde la correlación entre paths es de la unidad. La pendiente de la recta que une uno de los puntos con el punto (0,0) corresponde a la correlación entre los paths y como se mencionó anteriormente indica que tan proporcional es el valor del término sumado con el término restado en la ecuación 5.12. De estas figuras, teniendo en cuenta que la varianza del LCP es constante, puede intuirse además que aquellos paths con mayor varianza (puntos más a la derecha) tienden en general a tener una mayor correlación con el LCP.

Para determinar qué relación deben tener los valores de correlación global asignados al cálculo de los valores de varianza de cada path y de covarianza entre paths, se realizó una prueba que consistía en determinar el valor de correlación global que al ser utilizado en la ecuación 4.17 permitía calcular el mismo valor de

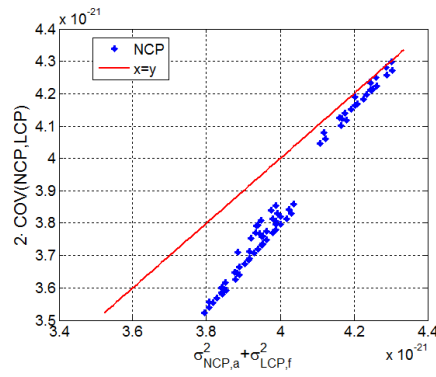
## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR



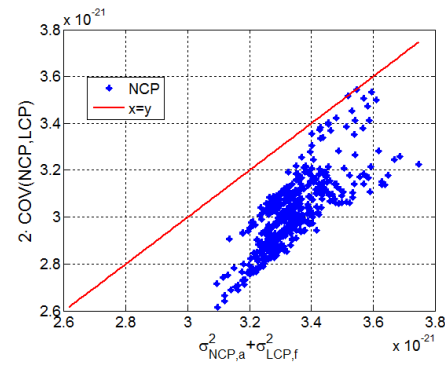
(a) *C432*



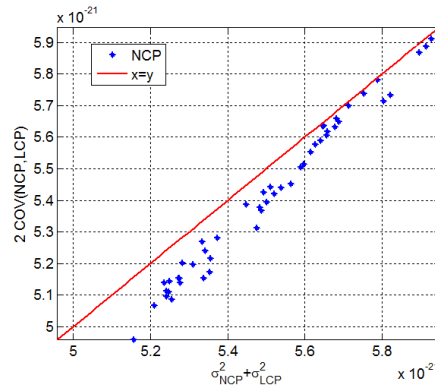
(b) *C499*



(c) *C880*



(d) *C1908*



(e) *C2670*

Figura 5.13: Valores de  $\sigma_{NCP}^2 + \sigma_{LCP}^2$  y  $2COV(NCP, LCP)$  de los paths críticos de circuitos ISCAS analizados

varianza del path o de covarianza entre paths que el obtenido a partir del layout, a éste valor se le llamó correlación equivalente (peq). Es decir, se buscó una matriz de correlación que tiene la siguiente forma:

$$\rho_{eq,ij} = \begin{pmatrix} 1 & \rho_{eq} & \cdots & \rho_{eq} \\ \rho_{eq} & 1 & \cdots & \rho_{eq} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{eq} & \rho_{eq} & \cdots & 1 \end{pmatrix} \quad (5.13)$$

Tal que al evaluar con ella la ecuación 4.17 se obtenga el mismo valor de varianza o de covarianza calculado mediante la matriz de correlación espacial. En la figura 5.14 se muestran los resultados obtenidos, donde se puede ver que igualmente existe una relación entre estos valores. Basándose en esta relación se decidió por utilizar un mismo valor de correlación global para estimar la varianza y covarianza de los paths y así predecir el valor de varianza de la diferencia de retardos de una forma menos pesimista.

Sin embargo, es necesario aún seleccionar el valor de correlación global que permite predecir adecuadamente el mayor conjunto de paths a monitorear. A continuación se muestran los resultados obtenidos de algunos experimentos realizados para determinar cuál es el valor de correlación global que proporciona un máximo valor de varianza de diferencia de retardos teniendo en cuenta que en base a las pruebas anteriores, se utilizará un único valor de correlación global entre todas las compuertas del circuito.

### 5.5.1. Correlación entre paths y Correlación global

La figura 5.16 muestra cómo cambia el valor de correlación entre el retardo de dos cadenas de inversores como la mostrada en la figura 5.15 en función de la correlación espacial global utilizada para realizar el análisis estadístico. En este experimento se consideró que el LCP corresponde a la cadena con  $M = 20$  inversores y el NCP corresponde a la otra cadena con N variable número de

## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

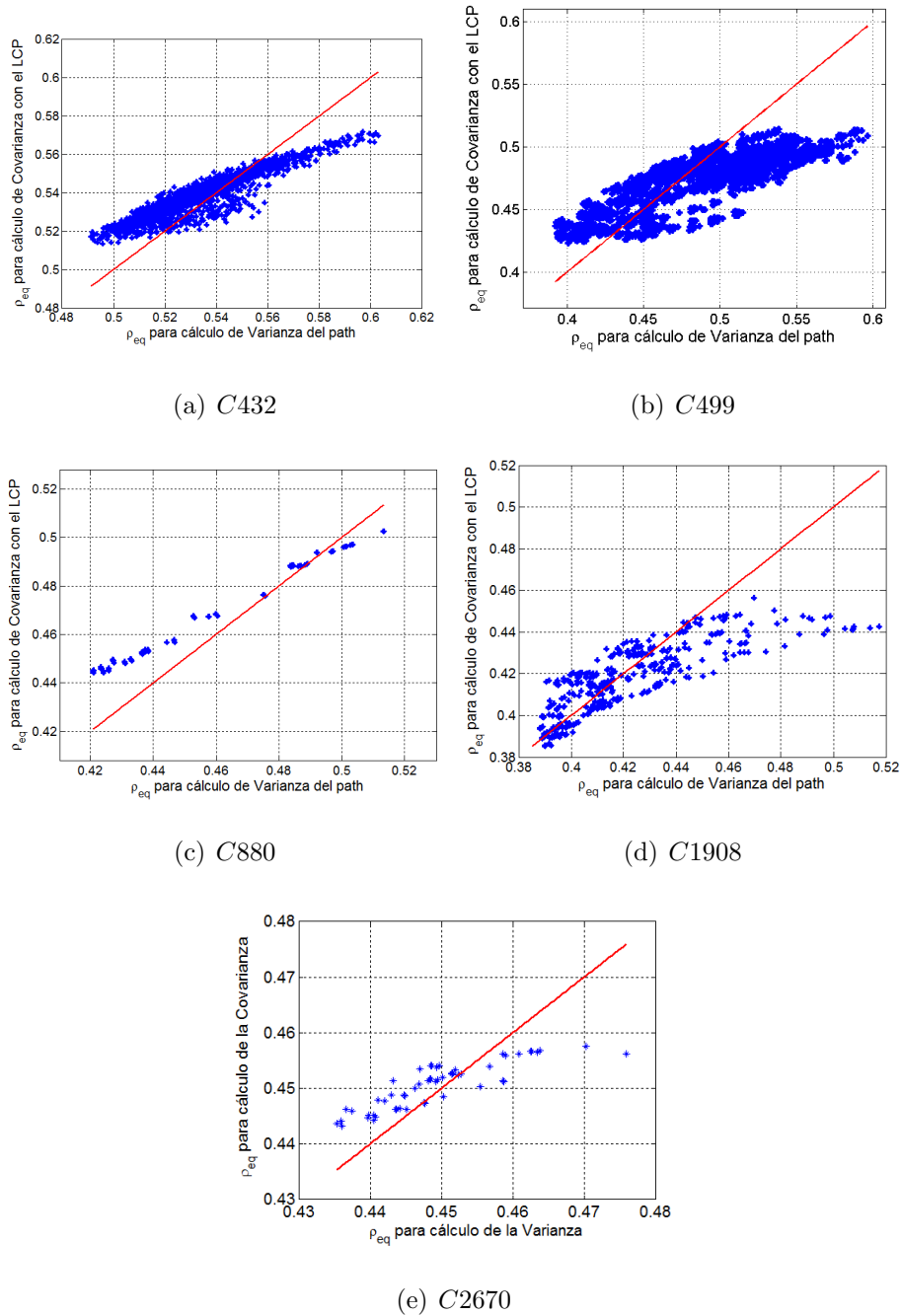


Figura 5.14: Valores de  $\rho_{eq}$  para el cálculo de la varianza de los NCP y su covarianza con el LCP

inversores.

Como se observa en la figura 5.16, el valor de correlación entre paths aumenta



## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

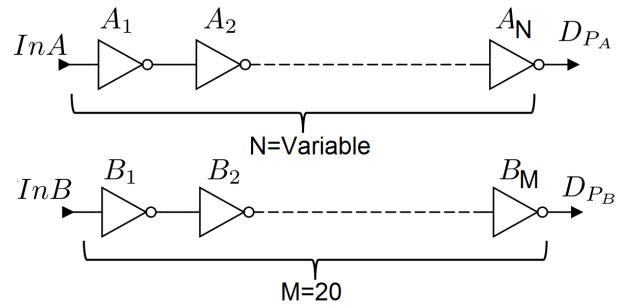


Figura 5.15: Dos paths con  $M = 20$  y  $N = Variable$  Inversores

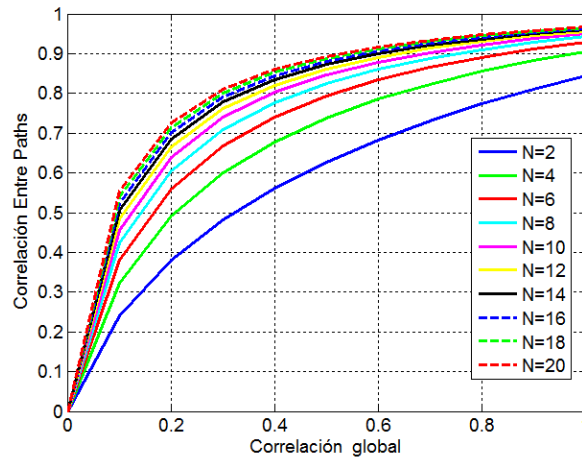


Figura 5.16: Correlación entre paths vs Correlación espacial global

a medida que se aumenta tanto la profundidad del NCP como la correlación global espacial utilizada. Por tanto, al analizar los paths críticos del circuito de esta forma, se supone que su grado de correlación es mayor debido a que son en general los paths más profundos.

### 5.5.2. Efecto de la correlación espacial global en la desviación estándar de DR

Se calculó la desviación estándar de DR en función de la correlación espacial global asumida entre todas las compuertas del circuito, para el mismo ejemplo anterior, se obtuvieron las curvas mostradas en la figura 5.17. Se observa que a

## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

---

partir de cierta profundidad del NCP hay un cambio en la tendencia (aumentar o disminuir) de la desviación estándar de DR en función de la correlación espacial global asumida. Este cambio puede explicarse como sigue: Para un valor fijo de correlación global, se tiene que los paths más profundos tendrán una mayor correlación entre paths como se concluyó de la figura 5.16, esto hace que el término de covarianza en la ecuación 5.6b aumente. A medida que N aumenta, la varianza del retardo del NCP aumenta, pero debido al aumento de la correlación entre paths, el efecto total que se tiene en la desviación estándar de DR es cada vez menos significativo. A medida que la profundidad del NCP es más cercana al LCP, el término de covarianza entre ambos paths ( $cov(D_{NCP}, D_{LCP}) = \rho_{ip} \cdot \sigma_{NCP} \cdot \sigma_{LCP}$ ) aumenta más rapido que lo que aumenta  $\sigma_{NCP}^2$ , el efecto total será un valor menor de  $\sigma_{DR}$ . Si se observa la zona donde  $\rho_g \approx 0$  de la figura 5.17, puede verse que para paths NCP con profundidad cercana al LCP, el valor de desviación estándar máximo se encuentra en este punto.

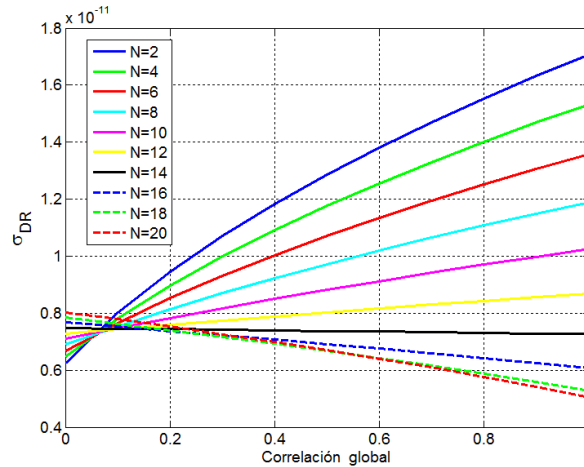


Figura 5.17:  $\sigma_{DR}$  vs  $\rho_g$

### 5.5.3. Efecto de la Profundidad del path bajo análisis

La figura 5.18 muestra cómo cambia la varianza de DR en función de la profundidad N del path NCP bajo análisis en el ejemplo de la figura 5.15. Para calcu-

## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

---

lar la varianza de cada path y la covarianza entre paths, se utilizó la matriz de correlación global, la figura muestra el comportamiento para diferentes casos de correlación global asumida.

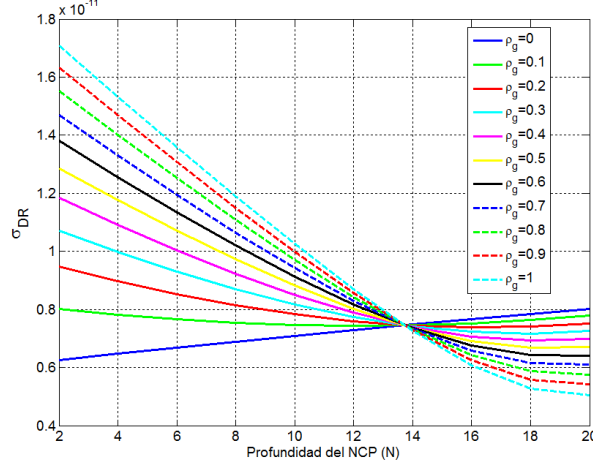


Figura 5.18:  $\sigma_{DR}$  vs Path depth

La tendencia a disminuir de la desviación estándar de DR se hace más débil a medida que se aumenta la profundidad del path, este comportamiento se explicó anteriormente. Claramente puede verse que todas las curvas se intersectan cerca de un valor de profundidad del path NCP, para este caso de  $N \approx 14$ , a partir de este punto, el valor de desviación estándar de DR obtenido mediante  $\rho_g = 0$  es el mayor.

Para el ejemplo analizado, el punto de intersección de las curvas puede estimarse, asumiendo que el retardo de todos los inversores tiene exactamente el mismo valor de desviación estándar e igualando la desviación estándar de DR para dos casos de correlación global diferentes. Aplicando las siguientes ecuaciones válidas para una cadena de inversores idénticos:

$$\begin{aligned}
 \sigma_{PA}^2 &= [N + \rho_g(N^2 - N)]\sigma_{inv}^2 \\
 \sigma_{PB}^2 &= [M + \rho_g(M^2 - M)]\sigma_{inv}^2 \\
 cov(PA, PB) &= N \cdot M \cdot \rho_g \cdot \sigma_{inv}^2
 \end{aligned} \tag{5.14}$$

## 5.5. ESTIMACIÓN DE VARIANZA DE DR SIN INFORMACIÓN DE LAYOUT PARA LA PREDICCIÓN DE PATHS A MONITOREAR

---

Entonces, reemplazando en la ecuación 5.6b para dos los valores de correlación espacial global de  $\rho_g = 0$  y  $\rho_g = 1$ :

$$\sigma_{DR_{\rho=1}}^2 = (N^2 + M^2) \cdot \sigma_{inv}^2 - 2N \cdot M \cdot \sigma_{inv}^2 \quad (5.15a)$$

$$\sigma_{DR_{\rho=0}}^2 = (N + M) \cdot \sigma_{inv}^2 - 0 \quad (5.15b)$$

De la ecuación 5.15a se tiene que a medida que  $N$  aumenta, el primer término aumenta en función de  $N^2$ , mientras que el término que se resta aumenta en función de  $N \cdot M$ , dado que  $M \geq N$ , éste término aumentará más rápido y por tanto la desviación estándar de DR debe disminuir. De la ecuación 5.15b se tiene que el primer término siempre aumenta en función de  $N$ , mientras que no se resta ningún valor a este término. A partir de cierta profundidad la desviación estándar de DR cuando se considera  $\rho = 0$  se hace más grande que cuando se considera  $\rho = 1$ , igualando las ecuaciones 5.15a y 5.15b y aplicando álgebra se obtiene que el valor de profundidad donde sucede esto es:

$$N \approx \frac{2M + 1 - \sqrt{8M + 1}}{2} \quad (5.16)$$

Para el ejemplo analizado con  $M = 20$ , se obtiene que  $N \approx 14$ , lo cual concuerda con los resultados obtenidos en la figura 5.18. De la ecuación 5.16 puede analizarse que el valor de  $N$  a partir del cual  $\rho_g = 0$  es un peor caso, es siempre menor que  $M$  y que la diferencia entre estos dos valores ( $M - N$ ) aumenta en función de  $M$ . Dado que los NCP son en general paths con una profundidad lógica cercana al LCP, es correcto utilizar el valor de cero correlación espacial global para estimar el valor máximo de la desviación estándar de DR bajo la suposición de que entre todas las compuertas existirá este mismo valor de correlación global.

#### 5.5.4. Comportamiento de los términos de varianza de la diferencia de retardos

Un análisis cualitativo sobre los términos de DR para el ejemplo mostrado en la figura 5.19, permite inferir por qué siempre que se analicen dos paths con igual profundidad, considerar  $\rho_g = 0$  proporciona un valor máximo de la desviación estándar de DR. En este caso se tiene que la varianza de DR es:

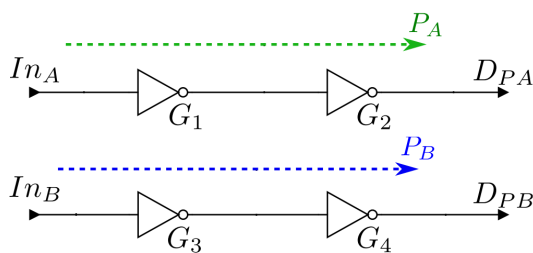


Figura 5.19: Dos paths de dos inversores

$$\begin{aligned}
 \sigma_{DR}^2 &= \sigma_{PA}^2 + \sigma_{PB}^2 - 2 \cdot COV(PA, PB) \\
 &= \sigma_{D1}^2 + \sigma_{D2}^2 + \sigma_{D3}^2 + \sigma_{D4}^2 + 2 \cdot COV(D1, D2) + 2 \cdot COV(D3, D4) \\
 &\quad - 2 \cdot COV(D1, D3) - 2 \cdot COV(D1, D4) - 2 \cdot COV(D2, D3) - 2 \cdot COV(D2, D4)
 \end{aligned}
 \tag{5.17}$$

A medida que se aumenta la correlación global, de los términos sumados sólo aumentan los correspondientes a  $COV(D1, D2)$  y  $COV(D3, D4)$  (ya que las componentes de varianza propias de cada compuerta son constantes sin importar la correlación utilizada) mientras que todos los términos restados aumentan en magnitud, por esto el efecto general será de una disminución en la varianza de DR.

## 5.6. Cobertura de los paths no detectados mediante uso de la banda de guarda

En la sección anterior se explicó el criterio de selección de paths a monitorear. Además se explicó que suponer un valor de correlación global igual para el cálculo de la varianza y covarianza entre paths, permite obtener una estimación más realista de la varianza de la diferencia de retardos y que utilizar correlación de cero proporciona un valor máximo de ésta y un conjunto de paths más cercano al real. Sin embargo, debido a la suposición del valor de correlación global, que puede no ser válido para todos los paths en el circuito, puede que algunos de ellos que requieran ser monitoreados (analizando el layout) no hayan sido identificados en el análisis a alto nivel. En esta sección se expone una estrategia para cubrir estos paths sin necesidad de insertar sensores y modificar el layout realizado.

En los circuitos convencionales, el período de reloj se selecciona en función del retardo del LCP. En el diseño estadístico es común tomar la esquina  $\mu + 3\sigma$  como punto de referencia para el período de reloj. Sin embargo, debido al envejecimiento, se introduce una banda de guarda tal que aún en el peor caso de degradación del LCP, incluso luego de un largo tiempo de funcionamiento, el retardo de éste path no sea mayor que el período de reloj establecido. En la figura 5.20 se ilustra lo dicho anteriormente. De esta forma la banda de guarda que se requiere para tolerar el envejecimiento del circuito puede calcularse por la siguiente ecuación:

$$T_{GB} = \mu_{LCP,age} + 3\sigma_{LCP,age} - \mu_{LCP,fresh} - 3\sigma_{LCP,fresh} \quad (5.18)$$

En los circuitos adaptivos, el período de reloj es seleccionado de forma que la banda de guarda para envejecimiento se elimina casi por completo. Sin embargo, al realizar la predicción y el análisis estadístico de tiempos fino basado en el layout, algunos de los paths que no fueron “predecidos”, sí requieren ser monitoreados. Puede evitarse tener que hacer modificaciones en el layout realizado si se penaliza

5.6. COBERTURA DE LOS PATHS NO DETECTADOS MEDIANTE USO DE LA BANDA DE GUARDA

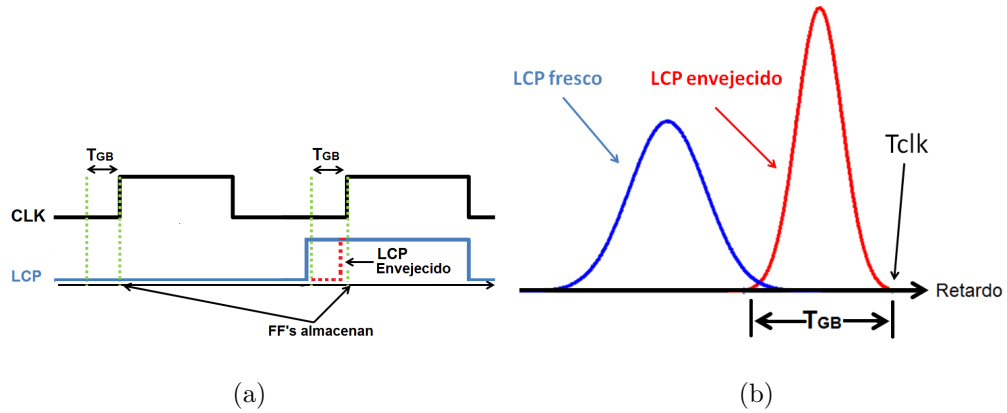


Figura 5.20: Definición de la banda de guarda para envejecimiento

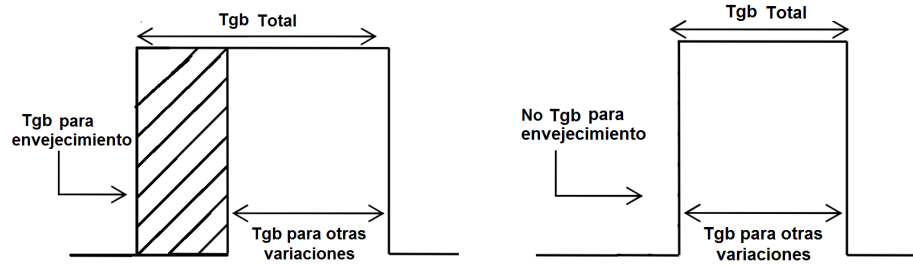
el período de reloj inicial, introduciendo una banda de guarda de la cantidad necesaria para tolerar el envejecimiento únicamente de estos paths que no fueron predecidos. La figura 5.21 ilustra la banda de guarda total de un circuito con diseño convencional, de un circuito adaptivo y de un circuito adaptivo donde no todos los paths requeridos son monitoreados.

La cantidad de tiempo a aumentar en la banda de guarda puede obtenerse mediante el uso de la variable aleatoria de diferencia de retardos DR, el path no detectado, puede dejarse de monitorear sin afectar la confiabilidad del circuito si la probabilidad de que su retardo envejecido exceda al retardo de algún otro path que sí es monitoreado (no necesariamente el LCP) en  $\Delta_{GB}$  es menor que el umbral requerido para que sea monitoreado  $\varepsilon_R$

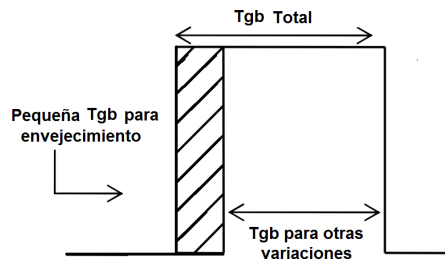
$$P(DR \leq \Delta_{GB}) \leq \varepsilon_R \quad (5.19)$$

Cada uno de los paths no detectados, requieren un valor de banda de guarda diferente respecto a cada path que sí es monitoreado (dependiendo de su grado de correlación) de todos estos valores, se escoge el mínimo, de manera que el path no monitoreado será cubierto por al menos un path monitoreado.

El valor de banda de guarda que debe ajustarse al período de reloj será el máximo del valor requerido por todos los paths no identificados.



(a) Banda de guarda convencional (b) Banda de guarda circuito adaptivo sin penalización



(c) Banda de guarda circuito adaptivo con penalización

Figura 5.21: LTDM y Modelo Ajustado

## 5.7. Resultados en circuitos ISCAS combinacionales

A continuación se resumen los resultados obtenidos del trabajo de tesis. En la tabla 5.1 se muestran la cantidad de paths topológicos y prefiltrados obtenidos mediante el criterio mencionado. El circuito C432 es el que tiene mayor cantidad de paths de los analizados.

### 5.7.1. Predicción de paths

La tabla 5.2 muestra la cantidad de paths a monitorear seleccionados mediante el uso de la variable aleatoria de diferencia de retardos, utilizando una correlación global espacial de  $\rho_g = 0$  para realizar un SSTA estimado a alto nivel, obteniendo el valor de varianza de cada path y su covarianza con el retardo del LCP y con los demás paths seleccionados.



Circuito	Paths topológicos	Paths prefiltrados
C432	82364	66781
C499	9440	8952
C880	4935	2643
C1908	15638	10079
C2670	3381	950

Tabla 5.1: Path topológicos y prefiltrados en cada circuito

En esta tabla y las siguientes, PMP es el número de paths obtenido mediante la predicción o estimación a alto nivel de la varianza de la diferencia de retardos, PML corresponde al número de paths a monitorear obtenido mediante el análisis del layout realizado. PNI son los caminos no identificados en la estimación y PII son los que fueron indentificados incorrectamente como paths a monitorear.

Para todos los casos el conjunto de paths estimado es cercano al *verdadero* obtenido luego de realizar el análisis del layout. Además, el conjunto de paths no detectado es pequeño respecto al total requerido y la cantidad de modificaciones requeridas para el segundo layout es bastante menor a que si no se realiza la etapa de análisis a alto nivel. Por otra parte, los sensores introducidos para monitorear los paths incorrectamente identificados podrían desconectarse del riel de alimentación y así penalizar únicamente en área el circuito realizado.

El circuito *C499* es el circuito con mayor cantidad de path críticos que deben ser monitoreados. Esto se debe a que es un circuito altamente balanceado, donde muchos caminos tienen un retardo similar al LCP y por tanto, aún si tienen alta correlación, deben ser monitoreados.

La tabla 5.3 muestra los resultados obtenidos al utilizar como umbral de selección valores mayores y menores a los de la tabla 5.2. Lógicamente, al aumentar el umbral de selección menos paths requieren ser monitoreados.

En las figuras 5.22 se muestran únicamente los paths identificados mediante

---

## 5.7. RESULTADOS EN CIRCUITOS ISCAS COMBINACIONALES

---

	$\varepsilon_R = 0.135\%$			
Circuito	PMP	PML	PNI	PII
C432	1518	1283	64	299
C499	2272	2120	0	152
C880	89	60	0	29
C1908	399	321	5	83
C2670	64	53	2	13

Tabla 5.2: Resultados en Circuitos ISCAS

	$\varepsilon_R = 0.0675\%$				$\varepsilon_R = 0.27\%$			
Circuito	PMP	PML	PNI	PII	PMP	PML	PNI	PII
C432	1707	1452	75	330	1302	1126	51	227
C499	2356	2193	0	163	2176	2087	0	89
C880	93	69	1	25	80	57	0	23
C1908	465	352	5	118	335	283	7	59
C2670	66	56	2	12	56	51	1	6

Tabla 5.3: Resultados en Circuitos ISCAS, Diferentes umbrales.

el análisis de layout y mediante el análisis estimado respecto al LCP (sin realizar la búsqueda de redundancia). Cada punto representa la desviación estándar de DR estimada vs la obtenida del análisis del layout. Como se puede observar, la estrategia seguida para la predicción de la varianza de DR resulta proporcionar valores cercanos a los reales. En la figura se muestran además los puntos correspondientes a los paths no identificados, en los cuales la desviación estándar de la diferencia de retardo es mayor que la estimada mediante la aproximación planteada. También se muestran los paths identificados incorrectamente respecto al LCP que corresponden a puntos cuyo valor de desviación estándar estimada es mayor que la del layout. Los paths correctamente detectados pueden estar en cualquiera de las dos regiones limitadas por la línea roja, ya que su detección también de-

pende del valor medio de la diferencia de retardos. Es decir, paths correctamente predecidos cuya varianza de DR real es mayor que la estimada tienen un retardo muy cercano al LCP, de manera que cumplen con el umbral de probabilidad aún con la menor varianza estimada.

### 5.7.2. Banda de guarda para tolerar paths no identificados

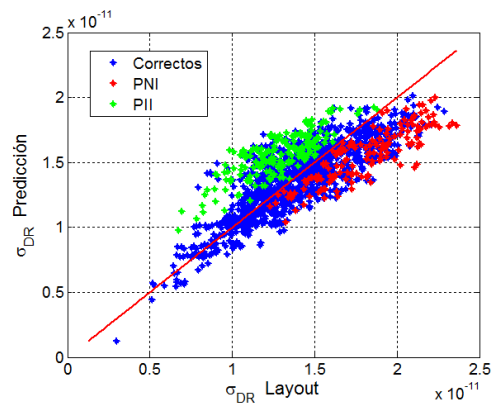
La tabla 5.4, muestra la banda de guarda total requerida en cada uno de los circuitos analizados, para tolerar el envejecimiento en los paths no detectados. En todos los casos, la reducción de la banda de guarda con respecto a un diseño convencional es significativa.

Para el caso del circuito *C2670* al cambiar el umbral de decisión (tablas 5.2 y 5.3), prevalece un path no identificado el cual es siempre el que define la banda de guarda y por esto la cantidad requerida no se modifica. Tanto para el circuito *C1908* como para el *C432* al aumentar el umbral de decisión, alguno de los nuevos paths no identificados requiere de una mayor banda de guarda. En el caso del circuito *C880* para el umbral de 0.0675% alguno de los paths a monitorear no fue identificado en la estimación y por tanto se requiere una pequeña cantidad de banda de guarda para tolerar su envejecimiento. Para el circuito *C499* todos los paths que requieren ser monitoreados fueron identificados para cualquiera de los umbrales utilizados, por tanto no se requiere introducir la banda de guarda.

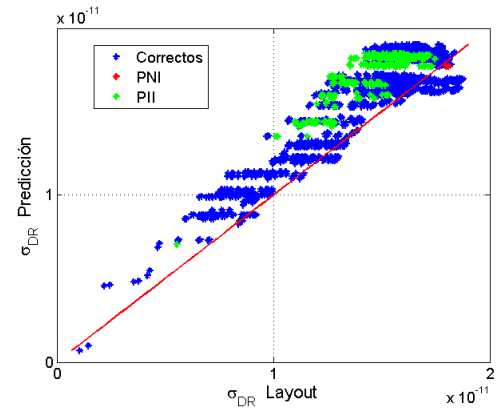
### 5.7.3. Compromiso entre paths no identificados e identificados incorrectamente

Es posible establecer un compromiso entre los paths identificados incorrectamente y los paths no identificados, cambiando el umbral de probabilidad utilizado en la predicción ( $\varepsilon_P$ ), el cual en los análisis anteriores se seleccionó igual al requerido ( $\varepsilon_R$ ) utilizado en el análisis del layout. Si el umbral de predicción se aumenta, menos paths son detectados y la cantidad de paths incorrectos debe disminuir,

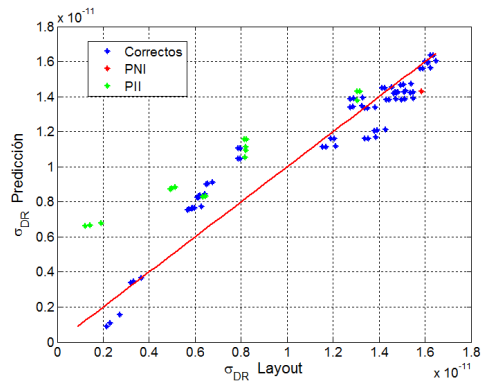
## 5.7. RESULTADOS EN CIRCUITOS ISCAS COMBINACIONALES



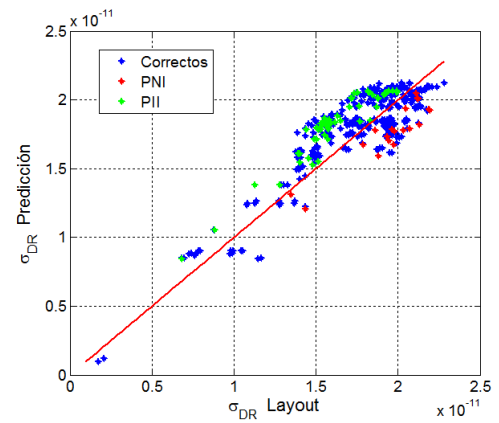
(a) Circuito *C432*



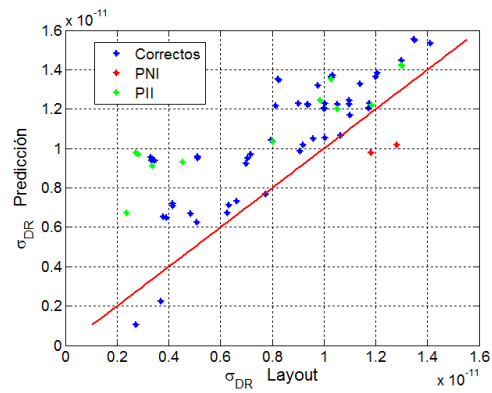
(b) Circuito *C499*



(c) Circuito *C880*



(d) Circuito *C1908*



(e) Circuito *C2670*

Figura 5.22: Desviación estándar de DR estimada vs real

---

## 5.7. RESULTADOS EN CIRCUITOS ISCAS COMBINACIONALES

---

Circuito	GB (ps)	GB (ps)	GB (ps)	GB (ps)
	Normal	Prop. $\varepsilon_R = 0.0675\%$	Prop. $\varepsilon_R = 0.135\%$	Prop. $\varepsilon_R = 0.27\%$
C432	23.788	6.454	9.345	16.703
C499	12.036	0	0	0
C880	13.963	0.616	0	0
C1908	21.699	1.461	1.461	7.068
C2670	25.119	3.179	3.179	3.179

Tabla 5.4: Banda de guarda requerida para tolerar envejecimiento

mientras que la cantidad de paths no seleccionados aumenta, ya que menos paths cumplen con el umbral de decisión de la predicción. Si el umbral de predicción se disminuye, los resultados son los opuestos:

$$\begin{aligned}
 &\text{si } \varepsilon_P \geq \varepsilon_R \quad PII \downarrow \quad PNI \uparrow \\
 &\text{si } \varepsilon_P \leq \varepsilon_R \quad PII \uparrow \quad PNI \downarrow
 \end{aligned}
 \tag{5.20}$$

La tabla 5.5 muestra los resultados para los circuitos ISCAS analizados, para casos donde se usó en la predicción un umbral mayor y menor al “requerido”. En la tabla 5.2 se mostraron los resultados, para el caso donde  $\varepsilon_P = \varepsilon_R$ .

Al utilizar un umbral menor en la predicción al requerido, aunque se reduce el conjunto de paths no identificado, la penalización en área que implica el conjunto incorrectamente identificado podría ser significativa. Por otra parte, utilizar un umbral de predicción mayor al requerido reduce la penalización en área de estos paths. La banda de guarda introducida puede ser menor a la mostrada, si se combina la inserción de sensores con la penalización en el período de reloj. De esta forma, se insertarían sensores en los paths que requieran mayor banda de guarda y podrían dejarse sin monitorear aquellos que tengan un menor impacto.

	$\varepsilon_R = 0.135\%$							
	$\varepsilon_P = 0.5 \cdot \varepsilon_R$				$\varepsilon_P = 2 \cdot \varepsilon_R$			
Circuito	PMP	PNI	PII	$\Delta GB$ (ps)	PMP	PNI	PII	$\Delta GB$ (ps)
C432	1707	26	450	6.454	1302	135	154	16.702
C499	2356	0	236	0	2176	12	68	0.869
C880	93	0	33	0	80	1	21	0.461
C1908	465	4	148	1.461	335	19	33	7.068
C2670	66	2	15	3.179	56	2	5	3.179

Tabla 5.5: Compromiso entre paths no identificados e identificados incorrectamente

# Capítulo 6

## Conclusión y trabajo futuro

La identificación de los caminos críticos que deben ser monitoreados en un circuito adaptivo es de vital importancia. En el presente trabajo, se analizó una metodología para la selección de caminos críticos en un circuito digital considerando envejecimiento por NBTI y variaciones de proceso. Como parte de la metodología, se busca minimizar el impacto de las modificaciones requeridas en etapas posteriores a la elaboración del primer layout, realizando un cálculo estimado a alto nivel de la probabilidad de que un path requiera ser monitoreado. Esto permite predecir un conjunto tentativo de paths a monitorear sobre el cual se puede tomar acciones anticipadas sobre la inserción de los sensores de retardo requeridos.

La metodología propuesta fue programada en C++ y aplicada sobre circuitos ISCAS combinacionales. Para aplicar el programa realizado, se deben pre-caracterizar el tiempo de retardo y el slew rate a la salida en función de la capacitancia de carga y el slew rate a la entrada de cada una de las compuertas lógicas presentes en los circuitos a analizar. Como archivos de entrada, se requiere una descripción de las entradas y salidas principales del circuito, así como la información de la topología a nivel compuerta.

Mediante experimentos, se observó una dependencia entre los valores de var-

---

ianza y covarianza entre paths, los cuales definen la varianza de la diferencia de retardos. Además se observó que existe también una relación entre el valor de correlación equivalente para el cálculo de la varianza de un path y la correlación equivalente para el cálculo de la covarianza entre paths. En base a esta observación y experimentos realizados, se concluyó que asumir un valor de correlación global entre todas las compuertas del circuito permitiría realizar una estimación menos pesimista de la varianza de la diferencia de retardos. Además, utilizar un valor de correlación global de cero proporciona un límite superior de la varianza de la diferencia de retardos bajo la anterior consideración.

La cobertura de los paths que no son detectados en el análisis estimado puede asegurarse introduciendo una banda de guarda en el período de operación inicial, esta banda de guarda es menor que la que se requeriría en un circuito convencional.

Es posible generar un compromiso entre la cantidad de paths incorrectamente seleccionados en el análisis estimado y la cantidad de paths no detectados mediante la modificación del umbral de decisión utilizado en el análisis estimado. De esta forma, al utilizar un valor de umbral de predicción mayor al requerido se puede reducir la penalización en área que resulta de incluir sensores en aquellos paths que no lo requieran. Aunque el conjunto de paths no identificado se incrementa, sería posible identificar aquellos paths que tengan un mayor impacto en la banda de guarda e insertar sensores en ellos, mientras que para tolerar el envejecimiento de otros que tengan un menor impacto, puede utilizarse la banda de guarda.

Como trabajo futuro es importante buscar estrategias que permitan realizar una comparación entre los NCP y el LCP más realista pero que aún proporcionen un caso conservativo de selección en el análisis. Una de estas estrategias es determinar la mínima degradación del LCP y de esta forma no asumir que éste no presenta efectos de envejecimiento. Otra posibilidad sería buscar una estimación de la correlación entre el grado de envejecimiento de dos paths.



# Lista de Figuras

1.1. Tasa de Fallas representada en la Curva de tina de baño. [2] . . .	3
1.2. Degradación del rendimiento e incumplimiento del tiempo de vida requerido. [4] . . . . .	5
1.3. Tendencias en la curva de tina de baño . . . . .	6
1.4. Mecanismos de Hot Carriers . . . . .	7
1.5. TDDBD . . . . .	8
1.6. Estructura lógica de un circuito digital . . . . .	8
1.7. Margen de Seguridad para garantizar la correcta operación del cir- cuito . . . . .	9
2.1. Condición de polarización donde el efecto de NBTI se presenta . .	14
2.2. Interface del silicio y óxido de compuerta . . . . .	15
2.3. Tuneleo de huecos y rompimiento de enlaces $S_i - H$ en la interface (Reacción). . . . .	16
2.4. Difusión de los átomos de hidrógeno en el óxido (Difusión). . . .	16
2.5. Pasivación de enlaces con átomos de hidrógeno que se recuperan (recuperación) . . . . .	17
2.6. Dependencia del NBTI con el campo eléctrico y temperatura. [6] .	18
2.7. Desviación del voltaje umbral como consecuencia del NBTI. [5] . .	19
2.8. Dependencia del NBTI con la frecuencia de operación. [8] . . . .	20
2.9. LTDM como borde superior de la degradación del $V_{th}$ [13] . . . .	21
2.10. Curvas obtenidas al aplicar el LTDM . . . . .	22

2.11. Variación del retardo de subida de una compuerta inversora . . . . .	24
2.12. Degradación del retardo para diferentes conjuntos de probabilidad en las entradas principales de un circuito digital. [13] . . . . .	25
2.13. Degradación del retardo diferente para dos caminos del circuito C17. [13] . . . . .	26
2.14. Error de operación debido al envejecimiento . . . . .	26
3.1. Curvas de movilidad vs campo eléctrico en silicio forzado [18] . . . . .	30
3.2. Conjunto de pareto para disminuir efecto de NBTI y consumo de potencia simultáneamente durante el período de standby [19] . . . . .	32
3.3. Técnica de reemplazo de compuertas para controlar el estado de algunos nodos internos . . . . .	33
3.4. Sobre dimensionamiento de las compuertas . . . . .	34
3.5. Condiciones de estrés en una compuerta inversora . . . . .	35
3.6. Condiciones de estrés en una compuerta NOR de dos entradas . . . . .	36
3.7. Ordenamiento de los nodos de entrada para disminuir el envejecimiento en una compuerta NOR3 . . . . .	36
3.8. Tipos de sensores . . . . .	39
3.9. Sensor de predicción de error . . . . .	40
3.10. Funcionamiento del sensor de predicción de error . . . . .	41
3.11. Sensor de detección de error . . . . .	42
3.12. Funcionamiento del sensor de detección de error . . . . .	43
3.13. Corrección de error en un microprocesador [29]. . . . .	43
3.14. Funcionamiento del sensor de detección de error . . . . .	44
4.1. Clasificación de las variaciones de proceso . . . . .	47
4.2. Caracterización del retardo de propagación de una compuerta digital	49
4.3. Matriz de TPHL para SR vs CL en un inversor . . . . .	50
4.4. Matriz de tphl vs CL vs DXW para una compuerta inversora . . . . .	51
4.5. Correlación espacial entre compuertas . . . . .	52

---

4.6. Degradación de la frecuencia en función del tiempo. [14] . . . . .	54
4.7. Desviación del $V_{th}$ debido a NBTI para diferentes valores de desviación en el $V_{th}$ inicial . . . . .	54
4.8. Sensibilidad de la degradación por NBTI a variaciones de proceso en el $V_{th}$ . . . . .	55
4.9. Cambio del factor de envejecimiento con el tiempo. . . . .	57
4.10. Dos paths a analizar . . . . .	62
5.1. Diagrama de flujo de la metodología propuesta . . . . .	66
5.2. Detección de paths en un circuito prueba . . . . .	67
5.3. Probabilidades de degradación para máximo envejecimiento del path [21] . . . . .	68
5.4. Valor de $\alpha$ en las compuertas con transición de bajo a alto en la salida . . . . .	69
5.5. Probabilidades de degradación para máximo envejecimiento del path	69
5.6. Criterio de prefiltrado de paths . . . . .	70
5.7. Región de integración . . . . .	71
5.8. Probabilidad de $D_{NCP} \geq D_{LCP}$ . . . . .	73
5.9. Probabilidad de $D_{NCP} \geq D_{LCP}$ . . . . .	73
5.10. Probabilidad de $D_{NCP} \geq D_{LCP}$ . . . . .	74
5.11. Dos paths con correlación estructural . . . . .	74
5.12. Tres paths en un circuito digital . . . . .	76
5.13. Valores de $\sigma_{NCP}^2 + \sigma_{LCP}^2$ y $2COV(NCP, LCP)$ de los paths críticos de circuitos ISCAS analizados . . . . .	82
5.14. Valores de $peq$ para el cálculo de la varianza de los NCP y su covarianza con el LCP . . . . .	84
5.15. Dos paths con $M = 20$ y $N = Variable$ Inversores . . . . .	85
5.16. Correlación entre paths vs Correlación espacial global . . . . .	85
5.17. $\sigma_{DR}$ vs $\rho_g$ . . . . .	86

---

5.18. $\sigma_{DR}$ vs Path depth . . . . .	87
5.19. Dos paths de dos inversores . . . . .	89
5.20. Definición de la banda de guarda para envejecimiento . . . . .	91
5.21. LTDM y Modelo Ajustado . . . . .	92
5.22. Desviación estándar de DR estimada vs real . . . . .	96

# Lista de Tablas

- 5.1. Path topológicos y prefiltrados en cada circuito . . . . . 93
- 5.2. Resultados en Circuitos ISCAS . . . . . 94
- 5.3. Resultados en Circuitos ISCAS, Diferentes umbrales. . . . . 94
- 5.4. Banda de guarda requerida para tolerar envejecimiento . . . . . 97
- 5.5. Compromiso entre paths no identificados e identificados incorrec-  
tamente . . . . . 98

# Bibliografía

- [1] Manfred Dietrich, Joachim Haase, “Process Variations and Probabilistic Integrated Circuit Design”, ISBN 978-1-4419-6620-9, Springer New York Dordrecht Heidelberg London
- [2] Alvin W. Strong, Ernest Y. Wu, Rolf-Peter Vollertsen, Jordi Suñé, Giuseppe La Rosa, Stewart E. Rauch III, Timothy D. Sullivan. “Reliability wearout mechanisms in advanced CMOS technologies”. IEEE Press Series on Microelectronic Systems. 2009.
- [3] Seyab, Nor Zaidi Haron, Said Hamdioui, “CMOS scaling impacts on reliability, What do we understand?” Delft University of Technology, Computer Engineering Laboratory.
- [4] M. A. Alam, ”A broad overview of reliability of semiconductor MOSFET” Purdue University, W. Lafayette, IN. disponible digitalmente en: <https://nanohub.org/resources/5863/download/2008.11.11-ece612-l20.pdf>
- [5] Alam, M. A., Kuffluoglu, H., Varghese, D. & Mahapatra, S. (2007). “A comprehensive model for pmos nbtj degradation: Recent progress”, Microelectronics Reliability Vol. 47(No.6): 853-862.
- [6] Mahapatra, S.; Bharath Kumar, P.; Alam, M.A., “Investigation and modeling of interface and bulk trap generation during negative bias temperature instability of p-MOSFETs,” Electron Devices, IEEE Transactions on , vol.51, no.9, pp.1371,1379, Sept. 2004

- [7] S. Rangan, N. Mielke, E. C. C. Yeh, “Universal recovery behavior of negative bias temperature instability,” IEDM, pp. 341-344, 2003
- [8] Chen, G.; Chuah, K. Y.; Li, M. -F; Chan, D. S H; Ang, C.H.; Zheng, J.Z.; Jin, Y.; Kwong, D.L., “Dynamic NBTI of PMOS transistors and its impact on device lifetime,” Reliability Physics Symposium Proceedings, 2003. 41st Annual. 2003 IEEE International , vol., no., pp.196,202, 30 March-4 April 2003
- [9] B. C. Paul, K. Kang, H. Kufluoglu, M. A. Alam, and K. Roy. “Impact of NBTI on the temporal performance degradation of digital circuit” , IEEE Electronic Device Letters, 26(8):560-562, August 2005.
- [10] Robert Entner, “Modeling and Simulation of Negative Bias Temperature Instability” , Dissertation, Wien, im April 2007.
- [11] Bhardwaj, S.; Wenping Wang; Vattikonda, R.; Yu Cao; Vrudhula, S., “Predictive Modeling of the NBTI Effect for Reliable Design” Custom Integrated Circuits Conference, 2006. CICC '06. IEEE , vol., no., pp.189,192, 10-13 Sept. 2006
- [12] “Solid state circuit technologies”, Edited by Jacobus W. Swart, ISBN 978-953-307-045-2, 2010.
- [13] Wenping Wang; Shengqi Yang; Bhardwaj, S.; Vrudhula, S.; Liu, T.; Yu Cao, “The Impact of NBTI Effect on Combinational Circuit: Modeling, Simulation, and Analysis,” Very Large Scale Integration (VLSI) Systems, IEEE Transactions on , vol.18, no.2, pp.173,183, Feb. 2010
- [14] Wenping Wang; Reddy, V.; Yang, B.; Balakrishnan, V.; Krishnan, Srikanth; Yu Cao, “Statistical prediction of circuit aging under process variations,” Custom Integrated Circuits Conference, 2008. CICC 2008. IEEE , vol., no., pp.13,16, 21-24 Sept. 2008

- [15] Song Jin; Yinhe Han; Huawei Li; Xiaowei Li, “ $P^2$ CLRAF: An Pre- and Post-Silicon Cooperated Circuit Lifetime Reliability Analysis Framework,” Test Symposium (ATS), 2010 19th IEEE Asian , vol., no., pp.117,120, 1-4 Dec. 2010
- [16] Lorenz, D.; Georgakos, G.; Schlichtmann, U., “Aging analysis of circuit timing considering NBTI and HCI” On-Line Testing Symposium, 2009. IOLTS 2009. 15th IEEE International , vol., no., pp.3,8, 24-26 June 2009
- [17] M.S. Khan and S. Hamdioui, “NBTI-Aware Nanoscaled Circuit Delay Assessment and Mitigation”, Proc. 3rd HiPEAC Workshop on Design for Reliability, Heraklion, Greece, January, 2011.
- [18] A.E. Islam, et al. “On the Possibility of Degradation Free Field Effect Transist”, Applied Physics Letter, 2008
- [19] Yu Wang; Xiaoming Chen; Wenping Wang; Balakrishnan, V.; Yu Cao; Yuan Xie; Huazhong Yang, “On the efficacy of input Vector Control to mitigate NBTI effects and leakage power,” Quality of Electronic Design, 2009. ISQED 2009. Quality Electronic Design , vol., no., pp.19,26, 16-18 March 2009
- [20] Kumar, S.V.; Kim, C.H.; Sapatnekar, S.S., “NBTI-Aware Synthesis of Digital Circuits,” Design Automation Conference, 2007. DAC '07. 44th ACM/IEEE , vol., no., pp.370,375, 4-8 June 2007
- [21] Wenping Wang; Zile Wei; Shengqi Yang; Yu Cao, “An efficient method to identify critical gates under circuit aging,” Computer-Aided Design, 2007. ICCAD 2007. IEEE/ACM International Conference on , vol., no., pp.735,740, 4-8 Nov. 2007
- [22] Yu Wang; Xiaoming Chen; Wenping Wang; Yu Cao; Yuan Xie; Huazhong Yang, “Gate replacement techniques for simultaneous leakage and aging op-



- timization,” Design, Automation & Test in Europe Conference & Exhibition, 2009. DATE '09. , vol., no., pp.328,333, 20-24 April 2009
- [23] Yu Wang; Xiaoming Chen; Wenping Wang; Yu Cao; Yuan Xie; Huazhong Yang, “Leakage Power and Circuit Aging Cooptimization by Gate Replacement Techniques,” Very Large Scale Integration (VLSI) Systems, IEEE Transactions on , vol.19, no.4, pp.615,628, April 2011
- [24] S. Kiamehr, F. Firouzi, and M. Tahoori, “Input and transistor reordering for nbti and hci reduction in complex cmos gates,” in Proceedings 01 the great lakes symposium on VLSI. ACM, 2012, pp. 201-206.
- [25] H. Luo, Y. Wang, K. He, R. Luo, H. Yang, and Y. Xie, “A novel gate-level nbti delay degradation model with stacking effect,” Integrated Circuit and System Design. Power and Timing Modeling, Optimization and Simulation, pp. 160-170, 2007.
- [26] Kai-Chiang Wu; Marculescu, D., “Joint logic restructuring and pin reordering against NBTI-induced performance degradation,” Design, Automation & Test in Europe Conference & Exhibition, 2009. DATE '09. , vol., no., pp.75,80, 20-24 April 2009
- [27] Kiamehr, Saman; Firouzi, Farshad; Tahoori, Mehdi.B, “Aging-aware timing analysis considering combined effects of NBTI and PBTI,” Quality Electronic Design (ISQED), 2013 14th International Symposium on , vol., no., pp.53,59, 4-6 March 2013
- [28] Hashimoto, M., “Run-time adaptive performance compensation using on-chip sensors,” Design Automation Conference (ASP-DAC), 2011 16th Asia and South Pacific , vol., no., pp.285,290, 25-28 Jan. 2011
- [29] Bowman, K.; Tschanz, J.; Wilkerson, C.; Shih-Lien Lu; Karnik, T.; De, V.; Borkar, S., “Circuit techniques for dynamic variation tolerance,” Design Au-

- tomation Conference, 2009. DAC '09. 46th ACM/IEEE , vol., no., pp.4,7, 26-31 July 2009
- [30] Das, S.; Blaauw, D, “Adaptive design for nanometer technology,” Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on , vol., no., pp.77,80, 24-27 May 2009 doi: 10.1109/ISCAS.2009.5117689
- [31] Agarwal, M.; Paul, B.C.; Ming Zhang; Mitra, S, “Circuit Failure Prediction and Its Application to Transistor Aging,” VLSI Test Symposium, 2007. 25th IEEE, vol., no., pp.277,286, 6-10 May 2007
- [32] Ernst, D.; Nam Sung Kim; Das, S.; Pant, S.; Rao, R.; Pham, T.; Ziesler, C.; Blaauw, D; Austin, T.; Flautner, K.; Mudge, T., “Razor: a low-power pipeline based on circuit-level timing speculation” Microarchitecture, 2003. MICRO-36. Proceedings. 36th Annual IEEE/ACM International Symposium on , vol., no., pp.7,18, 3-5 Dec. 2003
- [33] Das, S.; Roberts, D.; Seokwoo Lee; Pant, S.; Blaauw, D; Austin, T.; Flautner, K.; Mudge, T., “A self-tuning DVS processor using delay-error detection and correction,” Solid-State Circuits, IEEE Journal of , vol.41, no.4, pp.792,804, April 2006
- [34] Vazquez, J.C., Champac, V., Semião, J., Teixeira, I.C., Santos, M.B., Teixeira, J.P., ”Process Variations-Aware Statistical Analysis Framework for Aging Sensors Insertion”, Journal of Electronic Testing, 2013.
- [35] Ghosh, S.; Bhunia, S.; Roy, K., “CRISTA: A New Paradigm for Low-Power, Variation-Tolerant, and Adaptive Circuit Synthesis Using Critical Path Isolation” Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on , vol.26, no.11, pp.1947,1956, Nov. 2007

- [36] Jinjun Xiong; Zolotov, V.; Lei He, “Robust Extraction of Spatial Correlation,” *Computer-Aided Design of Integrated Circuits and Systems*, IEEE Transactions on , vol.26, no.4, pp.619,631, April 2007.
- [37] Lizheng Zhang; Yuheng Hu; Chen, C.C., “Statistical timing analysis with path reconvergence and spatial correlations,” *Design, Automation and Test in Europe*, 2006. DATE '06. Proceedings , vol.1, no., pp.5 pp.,, 6-10 March 2006
- [38] Blaauw, D; Chopra, K.; Srivastava, A.; Scheffer, L., “Statistical Timing Analysis: From Basic Principles to State of the Art,” *Computer-Aided Design of Integrated Circuits and Systems*, IEEE Transactions on , vol.27, no.4, pp.589,607, April 2008
- [39] Kyung-Ki Kim; Wei Wang; Ken Choi, “On-Chip Aging Sensor Circuits for Reliable Nanometer MOSFET Digital Circuits”, *Circuits and Systems II: Express Briefs*, IEEE Transactions on , vol.57, no.10, pp.798,802, Oct. 2010
- [40] Amin, C.S.; Menezes, N.; Killpack, K.; Dartu, F.; Choudhury, U.; Hakim, N.; Ismail, Y.I., “Statistical static timing analysis: how simple can we get?,” *Design Automation Conference*, 2005. Proceedings. 42nd , vol., no., pp.652,657, 13-17 June 2005
- [41] Orshansky, M.; Bandyopadhyay, A., “Fast statistical timing analysis handling arbitrary delay correlations,” *Design Automation Conference*, 2004. Proceedings. 41st , vol., no., pp.337,342, 7-11 July 2004
- [42] Merrett, M.; Yangang Wang; Zwolinski, M.; Maharatna, K.; Alioto, M., “Design metrics for RTL level estimation of delay variability due to intradie (random) variations,” *Circuits and Systems (ISCAS)*, Proceedings of 2010 IEEE International Symposium on , vol., no., pp.2498,2501, May 30 2010-June 2 2010

- [43] Galarza-Medina, F.J., “Detection of Small-Delay Defects in Nanometer Technologies Using Inter-Path Correlation”. Master Thesis, INAOE, México, 2011
- [44] Galarza-Medina, F.J.; Garcia-Gervacio, J.L.; Champac, V.; Orailoglu, A., “Small-delay defects detection under process variation using Inter-Path Correlation,” VLSI Test Symposium (VTS), 2012 IEEE 30th , vol., no., pp.127,132, 23-25 April 2011
- [45] Montgomery, Douglas C; “Probabilidad y estadística aplicadas a la ingeniería”; Limusa; 2<sup>a</sup> Ed; México 2008.