



INAOE

Diseño robusto de circuitos digitales de alto desempeño bajo variaciones de proceso

por:

Greg Gustavo Rifka Hernández

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**

Diciembre 2018

Tonantzintla, Puebla

Supervisada por:

Dr. Víctor Hugo Champac Vilela, INAOE

Dr. Héctor Luis Villacorta Minaya, UPA

@INAOE 2018

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 escalamiento de la tecnología de semiconductores ha permitido que los circuitos integrados alcancen una mayor densidad de integración y velocidad de operación. Sin embargo, la reducción de las dimensiones también ha provocado que cualquier pequeña imperfección durante el proceso de fabricación afecte de forma más pronunciada a el rendimiento de los sistemas, dichas imperfecciones se pueden cuantificar como variaciones en los parámetros físicos de los transistores, lo que causa que las características de los circuitos integrados como el retardo se comporten como una variable aleatoria. El presente trabajo tiene como propósito presentar un conjunto de estrategias y herramientas que permitan a un diseñador de IC generar diseños robustos en los sistemas digitales a un en presencia de las variaciones en la manufactura.

Dentro de las herramientas presentadas se incluyen modelos semi-analíticos para predecir el valor de la desviación estándar del retardo esperada tanto en una compuerta como en un camino lógico, los modelos generados son examinados y utilizados para generar métricas que guíen a los diseñadores a evaluar la criticalidad de cada compuerta en un camino lógico con el fin de realizar contra medidas que disminuyan la variación de forma eficiente.

También durante el trabajo se presenta una herramienta CAD desarrollada que permite realizar la optimización matemática de un camino lógico de forma automática mediante un dimensionamiento de las compuertas en el camino.

Abstract

The continuous scaling in semiconductors technologies has allowed more dense and fastest integrated circuits, however the shrink in the physical feature length has make that any small imperfection in the manufacturing process affect more heavily in the general performance of the system, this imperfection can be seen like variation in the physical parameters of the electronic devices that conform the systems, this variation causes that the delay behave like a random variable. This present work has as objective give a set of tools and strategies that allows a designer make robust systems in presence of process variations.

One of the tools that is presented is a semi analytical model that predicts the value of the delay standard deviation in any logic gate or logic path, this model is used in the making of design metrics that guide designers to quantify the critically of a logic gate in a path, this is make in order to help the designer to do countermeasures in an efficient way.

Also in this work a CAD tool is presented, this tool was developed with the purpose of making an automatic optimization process for any logic path in presence of process variations.

Agradecimientos

Al **Concejo Nacional de Ciencia y Tecnología (CONACYT)** Por el apoyo económico otorgado a través de la beca para estudios de maestría.

Al **Instituto Nacional de Astrofísica Óptica Electrónica (INAOE)** por abrirme las puertas para realizar los estudios de posgrado.

A mi asesor el **Dr. Víctor Hugo Champac Vilela** por el conocimiento y formación académica que he recibido de su parte.

A mi asesor el **Dr. Luis Villacorta Minaya** por su apoyo incondicional en el proyecto que fue esta tesis a un en la distancia.

A mi novia **Adriana Monserratt García Bandala** por estar siempre a mi lado en todos mis proyectos de vida.

A mi hermana **Marla Erika Rifka Hernández** por ser mi motivación para dar un buen ejemplo y seguir superándome.

Así como el resto de mi familia y amigos.

*Dedicado en memoria de mi abuela
Juana de Dios Olga González Cruz.*

Índice general

Resumen	II
Abstract	III
Agradecimientos	IV
1. Introducción	1
1.1. Breve historia de la electrónica	1
1.2. El escalamiento tecnológico	3
1.3. Variaciones de proceso	6
1.3.1. Variaciones sistemáticas	7
1.3.2. Variaciones no sistemáticas	8
1.4. Esquinas de diseño	11
1.5. Análisis estadístico del retardo	13
1.6. Justificación	14
1.7. Organización de la tesis	15

2. Modelo de primer orden para análisis estadístico	16
2.1. Introducción	16
2.2. Parámetros Eléctricos	17
2.2.1. Resistencia de carga y descarga	17
2.2.2. Capacitancia de entrada	23
2.2.3. Auto-Capacitancia	25
2.3. Modelo estadístico del retardo de una compuerta lógica	29
2.3.1. Modelo del retardo	29
2.3.2. Modelo para la desviación estándar del retardo	32
2.3.3. Modelo compacto del retardo estadístico	39
2.3.4. Modelo del retardo estadístico considerando el efecto stack	42
2.3.5. Validaciones	51
2.4. Modelo estadístico del retardo en caminos lógicos	53
2.4.1. Validación	56
2.5. Conclusiones	58
3. Optimización de caminos lógicos	59
3.1. Modelo para la optimización de la desviación estándar del retardo de una compuerta lógica	60
3.2. Modelo para la optimización de la desviación estándar del retardo de un camino lógico	62

3.3. Optimización bajo restricción de área	65
3.3.1. Optimización de un camino lógico compuesto solo por inversores	66
3.3.2. Optimización de un camino lógico compuesto por compuertas de diferente tipo	75
3.4. Optimización sin aumentar la capacitancia de entrada del camino lógico	79
3.5. Conclusiones	84
4. Guías prácticas de diseño para optimización de un camino lógico	85
4.1. Comportamiento del retardo con respecto a los parámetros de diseño	85
4.1.1. Factor de escala de la compuerta	85
4.1.2. Capacitancia de carga	86
4.1.3. El Stack	87
4.2. Guías de diseño	87
4.2.1. Métrica para diferentes tipos de compuerta	90
4.2.2. Métrica considerando efectos adyacentes	91
4.2.3. Métrica con inclusión de covarianza	92
4.2.4. Métrica Completa	95
5. Herramienta CAD para la optimización de caminos lógicos considerando variaciones de proceso	96
5.1. Introducción	96

5.2. Descripción general de la herramienta CAD	97
5.2.1. Características	97
5.2.2. Diagrama de flujo general	100
5.2.3. Archivos de entrada y salida de la herramienta	101
5.3. Diseño de experimentos	104
5.3.1. Diseño Factorial	104
5.3.2. Caracterización de la librería digital	110
5.3.3. Computo del σ_D	113
5.4. Validación de la caracterización	115
5.4.1. Validación a nivel compuerta lógica	115
5.4.2. Validación en camino lógico básico	119
5.4.3. Validación en un camino lógico compuesto	120
5.5. Multiplicadores de LaGrange	121
5.5.1. Optimización de caminos lógicos bajo una variable de retardo	124
5.5.2. Optimización de caminos lógicos minimizando el área	134
5.6. Método Heurístico	135
5.7. Identificación de caminos lógicos	138
5.8. Conclusiones	140
6. Conclusiones	142

Apéndice A	144
A.1. Demostración de la Ecuación 2.17	144
A.2. Demostración de la Ecuación 2.101	147
A.3. Inferencia de la Ecuación 3.1	149
A.4. Demostración de la Ecuación 3.8	150
A.5. Cálculos sobre las sensibilidades	152
A.5.1. Demostración de las sensibilidades	152
A.5.2. Modelos simplificados	154
A.6. Gráficas auxiliares de validación para el modelo analítico a nivel compuerta	157
A.7. Transformación de Cholesky	159
Apéndice B	161
B.1. Obtención de los coeficientes de la métrica propuesta	161
B.1.1. Coeficiente a_1	161
B.1.2. Coeficiente a_2	163
Apéndice C	166
C.1. Aplicación de los multiplicadores de LaGrange	166
C.1.1. Ejemplo 1: Minimizar la distancia desde el origen a una recta dada	166
C.1.2. Aplicación en la optimización de caminos lógicos	168

Apéndice D	171
D.1. Archivo de entrada	171
D.2. Interfaz Primaria	173
D.2.1. Barra de Menú	174
D.2.2. Zona de trabajo	176
D.3. Interfaz de selectora de caminos	179
D.4. Interfaz de Optimización de Lagrange	181
D.4.1. Analizador del camino	182
D.4.2. Configuración y generación de optimización	183
D.4.3. Resultados de optimización	186
D.5. Interfaz del Método Heurístico	192
 Lista de figuras	 200
 Lista de tablas	 203

Capítulo 1

Introducción

1.1. Breve historia de la electrónica

El campo de la electrónica es quizá el área de estudio de mas rápido crecimiento en el conocimiento humano, sus orígenes se remontan a principios del siglo XX y su impacto moldeo a lo que hoy conocemos como la sociedad actual, los primeros pasos de dicha ciencia fueron dados en 1904 con el desarrollo de la válvula Fleming la cual es considerada como el primer dispositivo electrónico de la historia [6]. Rápidamente después de su descubrimiento la válvula Fleming fue modificada y adaptada para diferentes aplicaciones, dichos componentes derivados recibieron el nombre de válvulas termoiónicas debido a su principio de operación el cual se basa en la emisión de electrones en el vacío por parte de una terminal cuando esta recibe suficiente energía térmica, este control logrado por las válvulas termoiónicas sobre el flujo de electrones es lo que dio el nombre a la rama de la electrónica [6].

Las válvula termoiónicas fueron usadas ampliamente no solo para construir nuevos sistemas nunca antes vistos (como radares y sistemas de comunicación) si no que también permitieron mejorar las prestaciones de sistemas emergentes como fue el caso de las maquinas de computo [6].

Las primeras maquinas de computo se desarrollaron en base a dispositivos electromecánicos llamados reveladores [7], los cuales consisten principalmente en un brazo conductor móvil cuyas posiciones son controladas por una bobina que funciona como electro-imán, tal que el relevador implementa un conmutador eléctrico. Sin embargo debido a sus partes móviles los reveladores son propensos

a fallos bajo factores de actividad elevados, debido a que la válvula termoiónica puede hacer la tarea de conmutación sin recurrir a las partes móviles, estas remplazaron rápidamente a los reveladores en dichas aplicaciones dada su mayor confiabilidad y velocidad [7].

Con el paso del tiempo los requisitos del contexto histórico demandaban aplicaciones cada vez mas exigentes para los dispositivos electrónicos [7], claramente para cumplir las necesidades grandes investigaciones se llevaron acabo para buscar nuevos y mejores dispositivos electrónicos estas investigaciones tuvieron su mayor triunfo en 1947 cuando un equipo de físicos conducidos por el Dr. Shockley desarrollaron lo que hoy es el dispositivo electrónico mas importante, el transistor [6].

Un transistor recibe su nombre de la concatenación de la palabra inglesa tranfer resistor [6], lo cual puede traducirse como resistor de transferencia dicho nombre deja en claro su funcionamiento ya que un transistor es un dispositivo electrónico de al menos tres terminal que altera su resistencia eléctrica al recibir estímulos en su denominada terminal de control pudiendo el transistor entonces cumplir acciones de conmutación o amplificación en base a este efecto, los transistores son creados a partir de materiales semiconductores ya que en estos es relativamente fácil (comparado con los materiales aislantes y conductores) controlar su resistencia eléctrica por medio de procesos físicos, dependiendo del proceso físico utilizado para generar el efecto transistor se puede clasificar a dichos dispositivos en tres grandes grupos [8]:

- Transistor de punta de contacto.
- Transistor de unión bipolar (BJT)
- Transistor de efecto de campo (FET)

Si bien el primer tipo de transistor desarrollado fue el nombrado como punta de contacto, hoy en día es muy raro encontrarlo en aplicaciones practicas, siendo los FET's los encargados de dirigir el progreso de la electrónica actual [8]. Debido a que los transistores son bloques sólidos de material semiconductor modificado, estos son mas confiables que las válvulas termoiónicas las cuales deben encapsularse al vacío y someterse a un pre-calentamiento de su terminal para funcionar adecuadamente, con el paso del tiempo los transistores sustituyeron a los tubos de vacío en virtualmente todas las aplicaciones electrónicas [6].

Aun cuando los transistores fueron un paso en la dirección correcta en la mejora de los sistemas electrónicos, el numero de transistores requeridos para implementar

sistemas cada vez mas complejos crecía de una forma exponencial, este problema nombrado como *La tiranía de los números* limitaba los sistemas prácticos que podían implementarse, no por que los dispositivos no pudieran cumplir la tarea sino por que armar un sistema formado por mas de 10 000 transistores discretos era una tarea poco practica desde el punto de vista de la manufactura de los sistemas [7].

El ingeniero Jack Killby durante su trabajo en la empresa Texas Instruments visualizo una posible solución al examinar el proceso de fabricación de los transistores, el teorizo y demostró que un bloque de material semiconductor podía utilizarse no solo para implementar transistores individuales sino circuitos electrónicos completos, este tipo de circuitos en una sola pastilla de un material semiconductor se denomino circuito integrado (IC). El Dr. Robert Noyce toma la idea de Killby y la llevo al siguiente nivel con el desarrollo del proceso de fabricación planar el cual llevo a los IC al mercado comercial [6].

Aun con la llegada de los IC al mercado, los consumidores siempre solicitaban circuitos mas rápidos y eficientes, por lo que los fabricantes de IC's mejoraron año con año su proceso de fabricación con el final satisfacer dicha demanda, esta relación de oferta y demanda es mantenida aun hoy en dia y es lo que dio origen al mercado con mayor crecimiento en toda la historia humana [2].

1.2. El escalamiento tecnológico

Para lograr el proceso de mejora continua anteriormente mencionado los fabricantes de IC estudiaron y fabricaron diversas topologías eléctricas para implementar sus circuitos, casi todas estas primeras topologías tenían como bloque básico al transistor de unión bipolar [2], sin embargo el BJT tiene un consumo de potencia estático en la terminal de control cuando se utiliza como un conmutador, dicho efecto es dado por la propia física del BJT por lo que es inevitable durante el funcionamiento del transistor, este consumo de potencia se traduce en un calentamiento excesivo en los IC conforme aumentaba el numero de dispositivos dentro del mismo.

Para combatir dicho problema se desarrollo un nuevo tipo de transistor el cual tiene como característica principal un material aislante en su terminal de control, por lo que la potencia consumida de forma estática por dicha terminal es mínima, este transistor de compuerta aislada se denomino transistor de efecto de campo metal oxido y semiconductor o MOSFET por sus siglas en ingles, dicho nombre

es debido a los materiales originales utilizados en su fabricación, el MOSFET es el transistor que hoy se encuentra presente en la mayoría circuitos integrados desarrollados [2], aun cuando los materiales utilizados han cambiado desde su desarrollo original el principio de operación de dicho transistor es el mismo y su nombre a quedado invariante.

Aparte de su mejor desempeño como bloque de construcción para los circuitos digitales el MOSFET presenta la característica de escalabilidad lo cual se puede lograr siguiendo la ley de escalamiento bajo campo constante o también conocida como la ley de Dennard [2], dicha *ley* menciona que las características propias del dispositivo pueden mantenerse y a la vez mejorar a nivel global el desempeño de los IC si ciertos parámetros físicos de los transistores se escalan en un factor S , dichos parámetros son:

- Todas las dimensiones geométricas
- Voltajes aplicados
- Nivel de dopado

El nombre bajo campo constante hace referencia a que si sigue el escalamiento mencionado, los campos eléctricos aplicados a los transistores de un IC se mantienen invariantes aun después de aplicar el escalamiento, si bien en el pasado se propusieron otro tipo de escalamientos, el escalamiento bajo campo constante a sido la norma desde principios de 1980 hasta el surgimiento de las modernas tecnologías nanométricas [2]. Sin importar el tipo de escalamiento este se a utilizado para mejorar el desempeño de los circuitos integrados a lo largo del tiempo, ya que mientras mas transistores puedan ser empaquetados en un mismo circuito se pueden generar sistemas mas complejos. En 1965 el co-fundador de Intel Gordon Moore observo que el numero de transistores que pueden ser manufacturados en un IC comercial se ajusta muy bien a una curva lineal bajo una escala semilogarítmica, dicha gráfica se muestra en la Figura 1.1.

La Figura 1.1 se conoce como la ley de Moore [2] y como se observa en ella a partir de la década de los 80 a sido posible integrar en un solo IC mas de 100 mil transistores, dichos IC con tal elevado numero de transistores se conocen como sistemas de integración a muy alta escala (VLSI).

Así como el escalamiento a permitido aumentar el desempeño de los IC's, este también a traído efectos colaterales no deseados sobre todo en las modernas tecnologías nanométricas en las cuales un conjunto de fenómenos físicos

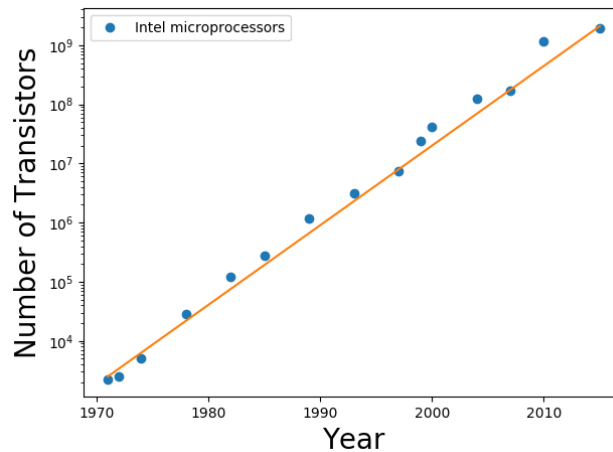


Figura 1.1: Ley de Moore

conocidos como efectos de canal corto modifican el comportamiento del transistor aun bajo la suposición de Dennard de mantener la escala en los campos eléctricos [2], estos efectos no deseados en menor o mayor medida han retrasado el escalamiento tecnológico para las modernas tecnologías [2], sin embargo no son la única causa de este retraso, conforme los transistores se vuelven mas pequeños la manufactura de los mismos se vuelve un proceso mas complejo y propenso a errores, esta complejidad hace que aun en los procesos mas modernos los valores de un circuito diseñado y uno fabricado varíen en cierta medida debido al tamaño extremadamente pequeño de los transistores estas variaciones por pequeñas que sean pueden alterar el comportamiento de los circuitos de manera importante [9].

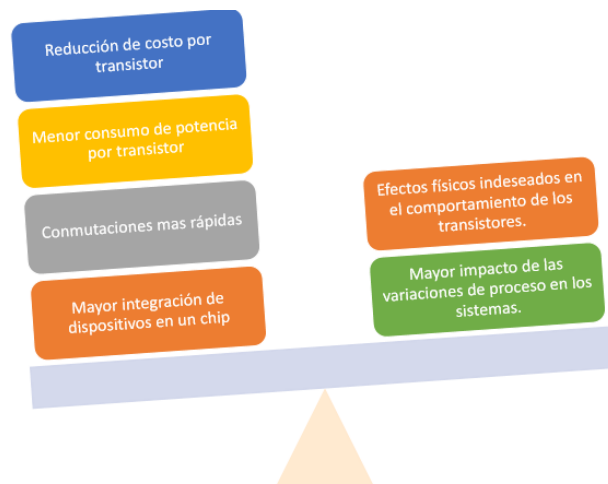


Figura 1.2: Trade-off del escalamiento tecnológico

1.3. Variaciones de proceso

En la industria de los semiconductores se denominan como variaciones de proceso a las fluctuaciones naturales que ocurren en los parámetros de los transistores cuando estos son fabricados. Si bien las variaciones de proceso siempre han estado presentes en las tecnologías nanométricas sus efectos se han intensificado ya que conforme se reducen las dimensiones de los dispositivos estos se hacen más sensibles a dichas variaciones [5]. Las variaciones de proceso son inducidas por una amplia gama de factores en cada uno de los diferentes pasos encontrados para llevar a cabo la manufactura de los IC. La alteración de los parámetros físicos de los dispositivos por dichas variaciones producen a la vez cambios en las características eléctricas de dichos dispositivos [5].

Un cambio en las características eléctricas de los dispositivos implica un cambio en el desempeño general de los sistemas formados por dichos dispositivos como puede verse en la Figura 1.3.

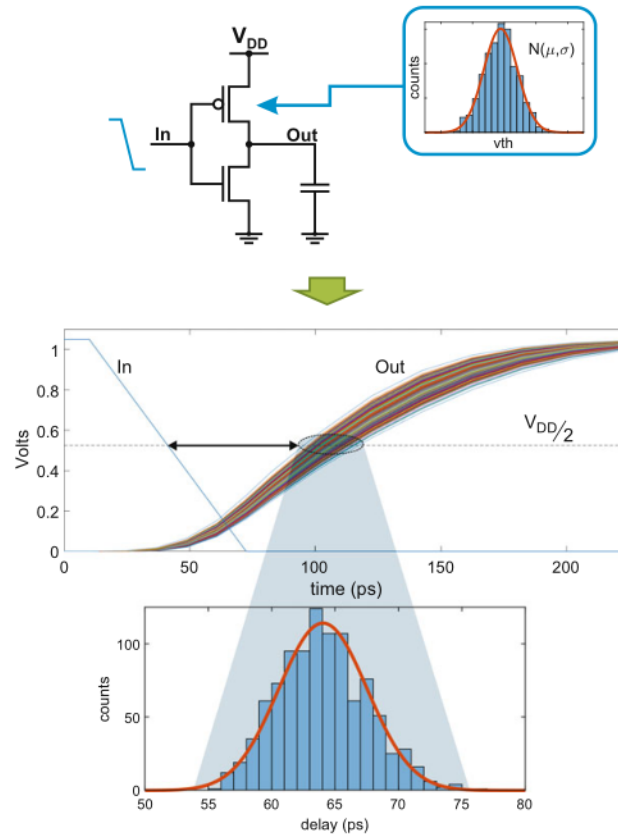


Figura 1.3: Impacto de las variaciones de proceso [17]

Para su estudio se divide a las variaciones de proceso en dos grupos:

- Sistemáticas.
- No Sistemáticas.

1.3.1. Variaciones sistemáticas

Las variaciones sistemáticas son las que ocurren de forma repetitiva durante la vida de un procesos de fabricación por lo que presentan un comportamiento que puede ser determinado y compensado cuando son identificadas, en otras palabras para un nodo tecnológico las variaciones sistemáticas estarán presentes siempre en las mismas proporciones independientemente del circuito que se fabrique, por lo cual se pueden identificar y compensar estudiando circuitos de prueba manufacturados en dicho nodo [15].

El tipo mas común de todas las variaciones sistemáticas son las que afectan a la geometría del Layout por medio del efecto conocido como LER (Line edge roughness) en el cual las geometrías rectangulares perfectas especificadas en la fase de diseño son implementadas por geometrías que incluyen curvaturas o cortes (ver Figura 1.4) que claramente modifican las características eléctricas deseadas, debido a que este efecto es sistemático los procesos de manufactura pueden efectuar contramedidas como el uso de de una técnica llamada OPC (Optical proximity correction) que mejora el desempeño del proceso de impresión fotolitográfica ante los efectos de LER [15].

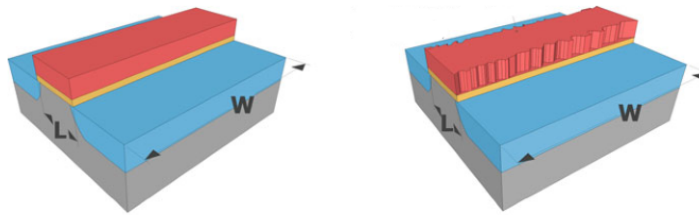


Figura 1.4: Efecto LER [5]

De forma similar el resto de variaciones sistemáticas son debidas a las limitaciones del proceso de fabricación de un nodo tecnológico, por lo general estas variaciones se pueden minimizar por medio de las llamadas *buenas practicas* durante el diseño del Layout, un ejemplo de una buena practica es

incluir transistores extras totalmente cortocircuitados (es decir sin función eléctrica) en los bordes de todas las cadena continuas de transistores que se deseen fabricar, los transistores extras se denominan informalmente como dummy's y tienen el propósito minimizar las variaciones sistemáticas del proceso de grabado el cual provoca desperfectos en los bordes de una cadena, debido a la naturaleza sistemática es de esperarse que al agregar los dummy's estos absorberán la mayor parte de la variación generada por el grabado aumentando la confiabilidad en la fabricación de los dispositivos importantes [15].

1.3.2. Variaciones no sistemáticas

Por su parte las variaciones no sistemáticas son caracterizadas por que no se presentan siempre con la misma proporción dentro de los circuitos fabricados por lo que para su estudio deben ser moldeadas por medio de métodos estadísticos, estas variaciones se clasifican en dos tipos dependiendo su comportamiento, globales y locales [5].

Variaciones globales

Las variaciones globales son aquellas que afectan en la misma medida a todos los transistores fabricados dentro de un circuito integrado específico y son moldeadas por un desplazamiento nominal del valor medio de los parámetros afectados por dichas variaciones, estas variaciones tiene como cualidad el no generar problemas de emparejamiento (Mismatch) ya que como su nombre lo dice actúan de forma global en todo el circuito integrado [15].

El origen de estas variaciones se debe a pequeños cambios en las condiciones iniciales de la fabricación para cada circuito integrado, dependiendo el origen de estos cambios se clasifican a las variaciones globales en los siguientes grupos [5].

- Variaciones dado a dado.
- Variaciones oblea a oblea.
- Variaciones lote a lote.
- Variaciones fábrica a fábrica.

Las variaciones dado a dado son causadas por una alta combinación de factores que se dan durante la fabricación de cada uno de los dados pertenecientes a una sola oblea, estos factores pueden tener un origen muy diverso que abarca desde las imperfecciones locales de la oblea hasta la no uniformidad del proceso de fabricación sobre toda la superficie de la oblea [15].

Por su parte las variaciones de oblea a oblea ocurren durante la manufactura de los circuitos en las diferentes obleas pertenecientes a un solo lote, las causas de estas variaciones son muy similares a las discutidas anteriormente [15].

Una variación lote a lote generalmente tiene una naturaleza más independiente al propio proceso de manufactura y es mas asociado con la preparación del mismo ya que no es raro que se de una situación donde entre la fabricación de dos lotes continuos se de un mantenimiento y/o limpieza a la maquinaria utilizada, lo cual altera las condiciones iniciales del siguiente lote [5].

Las variaciones fábrica a fábrica son generalmente más pronunciadas ya que se pueden deber a diferencias en las condiciones de los equipos y mascarillas utilizados por los diversos fabricantes.

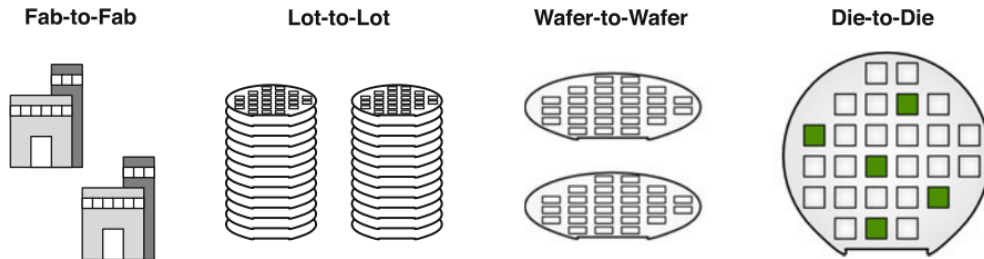


Figura 1.5: Variaciones Globales [5]

Variaciones Locales

Las variaciones locales o dentro del dado son un tipo de variación que afecta de diferente forma a cada uno de los transistores que componen un circuito integrado, debido a esta característica las variaciones locales si pueden generar problemas de emparejamiento entre dos o más transistores [16].

Las variaciones dentro locales son el tipo de variación dominante en las tecnologías nanométricas [5]. El origen de estas variaciones proviene de las diferentes inexactitudes ocurridas por los diferentes procesos utilizados durante

la manufactura del circuito a través de la superficie del dado [16].

Las variaciones locales se clasifican en dos grupos dependiendo de su comportamiento estadístico.

- Variaciones Correlacionadas
- Variaciones Independientes

Variaciones Correlacionadas

Las variaciones correlacionadas tienen como principal característica que el nivel de variación relativa entre dos transistores es función de la distancia entre ellos [15]. En este tipo de variaciones dos transistores cercanos tendrán prácticamente el mismo grado de variación mientras que para transistores alejados se presentarán variaciones que tienden a un carácter más aleatorio, la cantidad que mide el nivel de variación conjunta encontrado entre dos transistores con respecto a la distancia se denomina formalmente como coeficiente de correlación espacial (ρ) y toma valores en el rango $[-1, 1]$.

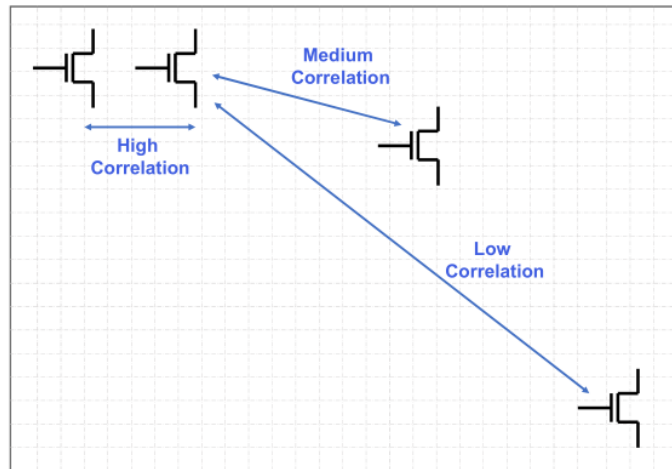


Figura 1.6: Variaciones Correlacionadas [17]

Los parámetros afectados en este tipo de variación son aquellos relacionados con la dimensión de los transistores, como lo son el ancho (W) y largo (L) del canal, así como también el espesor del óxido de la compuerta (T_{ox}).

Variaciones Independientes

Las variaciones independientes afectan a cada transistor de una forma totalmente aleatoria, son generadas principalmente por el proceso de dopado ya que la cantidad de impurezas en un transistor puede tener una fluctuación aleatoria en el número de dopantes implantados por unidad de área, como en las tecnologías nanométricas este numero de átomos dopantes puede ser de apenas unos cuantos cientos [16], cualquier pequeño exceso o falta de dopantes puede generar variaciones importantes en el voltaje de umbral (V_T) de los transistores, las variaciones independientes son encontradas con alta dispersión aún dentro de transistores cercanos por lo que este tipo de variación también se denomina como no correlacionada.

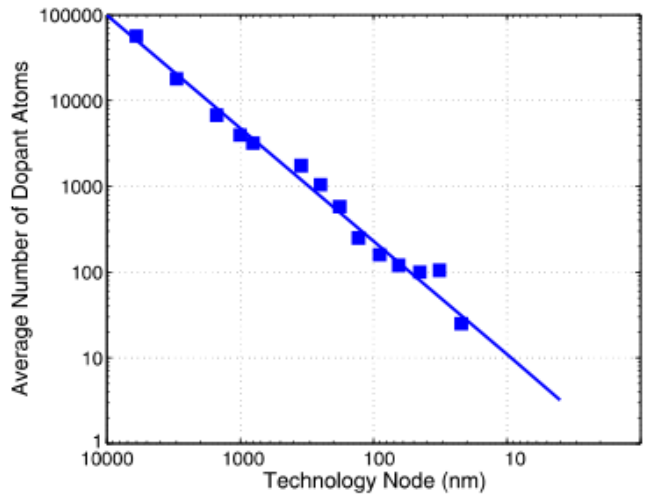


Figura 1.7: Numero de dopantes para diferentes nodos tecnológicos [16]

1.4. Esquinas de diseño

El análisis de esquinas de diseño es el método de estándar de validar el desempeño de los sistemas ante las variaciones ambientales y de proceso, dicho análisis consiste en examinar el circuito en las condiciones extremas en las que se planea operar/fabricar el circuito [5].

Posteriormente se realizan simulaciones eléctricas donde todos los transistores operan bajo las condiciones planteadas anteriormente, generalmente para diferenciar en este análisis se usa la nomenclatura T para especificar que un

transistor opera en condiciones típicas, F para transistores que operan a la velocidad de conmutación mas rápida esperada y S para los transistores con las velocidades de conmutación mas lentas esperadas. Posteriormente es común realizar el análisis combinando las redes nMOS y pMOS con los valores anteriores, tal que se llevaran acabo simulaciones con valores TT (nMOS y pMOS típicos), FS (nMOS rápido pMOS lento)....todas las posibles combinaciones de valores extremos (S y F) se conocen como esquinas de diseño y serán estos los puntos críticos de evaluación.

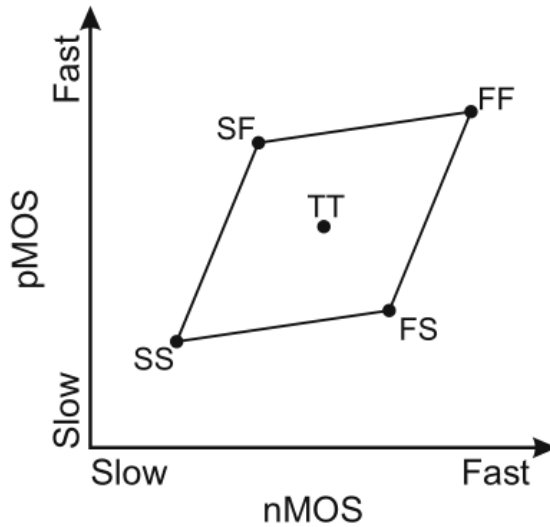


Figura 1.8: Esquinas de diseño [5]

El análisis por esquinas permite encontrar el comportamiento en peor caso del circuito evaluado bajo variaciones globales, sin embargo no considera variaciones del tipo local las cuales son importantes en las nuevas tecnologías, para cuantificar el efecto de las variaciones locales se debe realizar un análisis del tipo Monte Carlo en cada esquina de diseño lo cual es altamente costoso computacionalmente [15].

Generalmente los métodos de esquinas de diseño al ser pesimistas generan un sobre diseño ya que se esta asumiendo en la esquina SS que todos los transistores tienen las características degradadas [5], esto es una situación poco probable en un circuito fabricado, el sobre diseño viene dado ya que el método de esquinas busca asegurar una funcionalidad del 100% aun para este caso poco realista lo cual requiere generalmente un consumo de área y potencia muy alto [5].

El reducir este sobre diseño es la principal motivación de una nueva metodología conocida como análisis estadístico.

1.5. Análisis estadístico del retardo

El análisis estadístico del retardo se basa en modelar a las variaciones de proceso por medio de variables aleatorias (V.A) y estudiar los subsecuentes efectos en los circuitos por medio de procesos estocásticos [15].

Una variable aleatoria es una codificación numérica que se le asigna al resultado de un experimento aleatorio, estas variables pueden tomar valores de forma continua o discreta según la naturaleza del experimento aleatorio que codifican. El uso de variables aleatorias se puede usar para generalizar los resultados de experimentos aleatorios sin importar su naturaleza. Es común asociar a una variable aleatoria una distribución de probabilidad la cual se obtiene (comúnmente) por medios experimentales, dicha distribución permite conocer la probabilidad de que la variable aleatoria tome un valor específico en un experimento aleatorio. Cuando se combinan una o más variables aleatorias como entradas de una función matemática el resultado de la función será también una variable aleatoria en la que es posible calcular su distribución de probabilidad en función de las distribuciones individuales de las variables aleatorias que conforman la entrada de dicha función.

Las variaciones en los parámetros de fabricación (W, L, T_{ox}, V_T) pueden ser moldeadas mediante variables aleatorias que siguen una distribución de densidad de probabilidad del tipo normal con valores de σ y μ específicos del proceso de fabricación [5], para una compuerta lógica el retardo (D) es una función no lineal de los parámetros anteriormente mencionados, al ser esta una función no lineal el calcular su distribución de probabilidad directamente es un proceso complicado sin embargo este análisis se puede simplificar en gran medida si se identifica que los parámetros involucrados en el retardo están confinados a variaciones relativamente pequeñas, por lo cual se puede aproximar el retardo de propagación mediante el término de primer grado de la expansión de la serie de Taylor con el fin de linealizar la ecuación.

$$D = \mu_D + \sum_{i=1}^k \frac{\partial D}{\partial P_i} \Delta P_i \quad (1.1)$$

Donde

- μ_D : Representa el retardo evaluado en el valor nominal de los parámetros que presentan variación.

- P_i : Representa un parámetro fuente de variación.

Utilizando el operador segundo momento sobre la ecuación anterior se puede demostrar que:

$$\sigma_D^2 = \sum_{i=1}^n (S_{P_i}^D)^2 \sigma_{P_i}^2 \quad (1.2)$$

Donde

$$S_{P_i}^D = \frac{\partial D}{\partial P_i} \quad (1.3)$$

El modelo desarrollado en la ecuación 1.2 es la base del análisis estadístico los parámetros necesarios para computar el valor del σ_D pueden ser obtenidos por medio de simulaciones SPICE o en base a modelos analíticos ambas metodologías se discutirán durante el desarrollo de este trabajo.

La ventaja del análisis estadístico es que se ve al retardo no bajo un valor pesimista sino como una función de probabilidad que puede estudiarse para garantizar diseños con un grado de confiabilidad previamente escogido por el diseñador, aparte permite a los mismos diseñadores tener un acercamiento mayor al efecto de las variaciones de proceso en los circuitos digitales.

1.6. Justificación

El impacto de las variaciones de procesos ha causado que las características de los circuitos presenten desplazamientos con respecto a los valores de retardo a los que son diseñados causando pérdidas en el yield de los sistemas desarrollados. Estas variaciones de procesos presentan componentes aleatorias que dificultan el diseño de los circuitos y producen problemas de confiabilidad. Las técnicas existentes para hacer más robustos los circuitos integrados a este tipo de variaciones tales como el análisis de esquinas generan resultados pesimistas.

Para combatir lo anterior se presentan en este trabajo de forma independiente modelos matemáticos y herramientas de software que apoyen al diseñador de IC

a generar sistemas digitales óptimos bajo criterios estadísticos que reduzcan en general el sobre-diseño generado por los métodos pesimistas.

1.7. Organización de la tesis

En esta introducción se ha dado una primera perspectiva del efecto de las variaciones de proceso en los circuitos integrados digitales.

En el capítulo 2 se plantea un modelo semi-analítico para el estudio y análisis del efecto de las variaciones de proceso en el retardo tanto a nivel compuerta como a nivel camino lógico.

Posteriormente en el capítulo 3 se utiliza el modelo anteriormente desarrollado junto a técnicas de optimización con el fin de generar diseños robustos.

En el capítulo 4 se presentan una serie de ejemplos y ecuaciones que permiten obtener un entendimiento mas cualitativo de como utilizar las ecuaciones presentadas para llevar a cabo un dimensionamiento optimo de un camino lógico.

Por su parte el capítulo 5 se enfoca en el desarrollo y presentación de una herramienta CAD desarrollada para llevar a cabo el análisis estadístico en caminos lógicos, dicha herramienta es capaz de dimensionar directamente las compuertas pertenecientes a un camino para obtener valores óptimos en base a un criterio pre-seleccionado por el usuario. A lo largo del capítulo se explica no solo el funcionamiento externo de la herramienta sino también los algoritmos internos que esta implementa.

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

Capítulo 2

Modelo de primer orden para análisis estadístico

2.1. Introducción

El modelo eléctrico que aproxima el comportamiento de una compuerta lógica es mostrado en la Figura 2.1 [1], la cual muestra una terminal de entrada cargada por una capacitancia (C_{in}), que tiene origen a partir de las capacitancias de las compuertas de los transistores conectados a dicha terminal, el voltaje de esta terminal determina qué switch está conectado a la red, si el switch superior conduce este conectará la terminal de salida a la fuente de alimentación positiva conocida como V_{DD} , a través del resistor R_p la cual moldea el comportamiento de la red PullUp de los transistores pMOS, en caso contrario el resistor R_n moldeará el comportamiento de los transistores nMOS de la red PullDown la cual conectará la terminal de salida directamente a tierra, la terminal de salida es cargada por dos capacitancias, C_{self} la cual es una componente parásita que es asociada a la naturaleza de la misma compuerta lógica y una capacitancia de carga C_{FO} que representa la carga a la que es sometida la compuerta lógica, esta última capacitancia moldea tanto la carga producida por la etapa siguiente como las interconexiones.

Este modelo puede describir cualquier compuerta lógica a través de las cuatro cantidades C_{in} , C_{self} , R_p y R_n , es común que los diseñadores de circuitos integrados tengan una familia de compuertas lógicas con la misma topología y función lógica pero diferentes características eléctricas, esto es logrado escalando el ancho del canal de todos los transistores en la compuerta bajo la misma proporción [3], el

factor de escala se denomina como K y es uno de los principales parámetros de diseño.

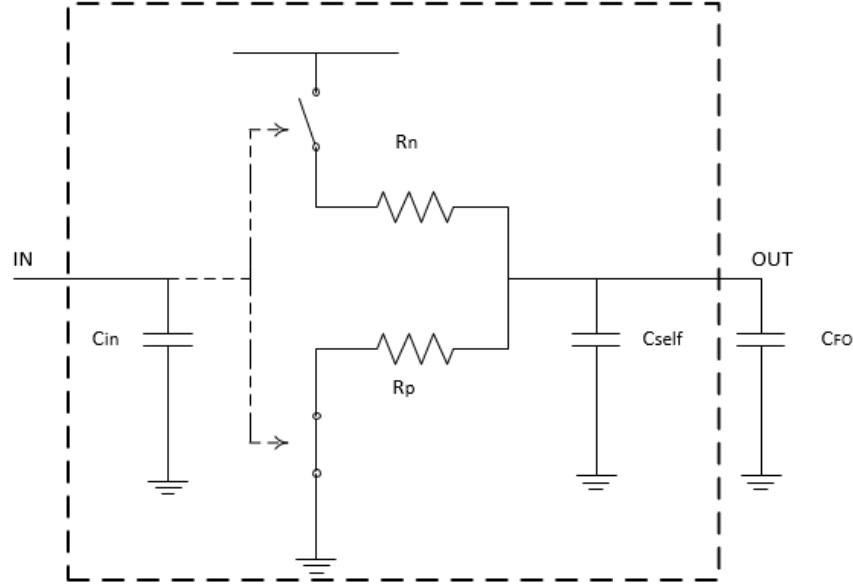


Figura 2.1: Modelo eléctrico de una compuerta lógica

En las siguientes secciones se desarrollará un modelo analítico para cada uno de los parámetros del modelo eléctrico y su dependencia con el factor K .

2.2. Parámetros Eléctricos

2.2.1. Resistencia de carga y descarga

De la Figura 2.1 se observa que la compuerta lógica cumple la función de cargar o descargar el nodo de salida por medio del resistor activo correspondiente, el retardo claramente es proporcional al tiempo de carga o descarga respectivo, cabe notar que en todo momento solo un resistor está activo en este modelo debido a la naturaleza complementaria de las redes PullDown y PullUp. Es una práctica de diseño común tener en cada resistor el mismo valor con el fin de tener tiempos de carga y descarga iguales [3]. La naturaleza de los resistores es debida a la resistencia encontrada en el canal de los transistores que forman las compuertas lógicas el cual depende principalmente de la carga acumulada en el canal y su

geometría, se conoce a partir de la teoría básica que el MOSFET tiene tres zonas de operación y que la función que rige la corriente en el canal varía dependiendo de estas zonas por lo que esta resistencia es no lineal en naturaleza, sin embargo para análisis de primer orden se puede definir esta resistencia por medio de un modelo lineal invariante en el tiempo (LTI), el cual puede ser encontrado por medio del valor promedio de la resistencia del canal durante la conmutación, la expresión resultante se puede encontrar en la literatura [3] para los modelos de transistores más sencillos y corresponde con la ecuación 2.1.

$$R_{DS} = \text{mean} \left(\frac{V_{DS}}{I_{DS}} \right) = \frac{V_{DD}}{\mu C_{ox} \frac{W}{L} (V_{DD} - V_T)^\alpha} \quad (2.1)$$

La compuerta más sencilla de diseñar y estudiar es el inversor CMOS, el cual consta únicamente de dos transistores, su esquemático es mostrado en la Figura 2.2 en dicho inversor las resistencias de red R_n y R_p corresponden con la resistencia de canal de cada transistor respectivo.

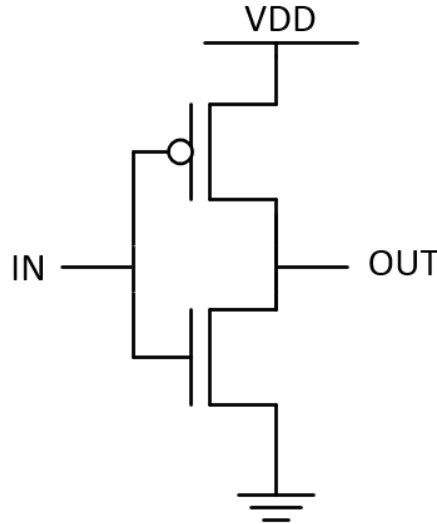


Figura 2.2: Inversor CMOS

Es una práctica común de diseño digital mantener el largo de canal en su dimensión mínima para favorecer la velocidad de conmutación, a la longitud de canal mínima que puede manufacturarse para una tecnología específica se le denominará por medio del símbolo L_{min} .

También por las consideraciones de tener tiempos de carga y descargas simétricos, el inversor es comúnmente diseñado tal que ambos transistores presenten la misma

resistencia de canal, sin embargo debido a que la movilidad de los portadores de carga es normalmente menor en los transistores pMOS por la naturaleza del silicio, estos presentan una mayor resistencia comparando con un dispositivo nMOS de las mismas dimensiones. Tal que si se quiere tener en el inversor tiempos de carga y descarga simétricos, se debe escalar el transistor pMOS para que se presente la misma resistencia de canal que en el transistor nMOS, tomando como referencia la relación de movilidad $\frac{\mu_n}{\mu_p} = 2$ se procede a establecer la condición de simetría para el inversor como:

$$R_n = R_p \rightarrow W_p = 2W_n \quad (2.2)$$

En el diseño digital es común expresar el ancho de los transistores por medio de un factor de escalamiento K , si se considera al inversor simétrico de dimensiones mínimas como el inversor con dimensiones $W_n = W_{min}$, $W_p = 2W_{min}$. Entonces se define al inversor simétrico escalado K -veces como el inversor con dimensiones $W_n = KW_{min}$ y $W_p = 2KW_{min}$.

La relación de los factores de dimensionamiento base de los transistores nMOS y pMOS para una compuerta lógica se denominan como el factor de diseño de la compuerta (FD), bajo esta definición el inversor CMOS tiene un FD de 1:2.

Expandir el análisis a otras compuertas lógicas es directo ya que estas son diseñadas para tener la misma resistencia que el inversor esto es logrado seleccionando el FD de forma adecuada.

Por ejemplo considerando la compuerta NAND2 que se muestra en la Figura 2.3 en esta compuerta la red PullDown está conformada por dos nMOS en serie, considerando que dos transistores en serie presentan una resistencia total igual a la suma de cada componente entonces cada transistor deberá tener un ancho de canal del doble de valor comparado con el transistor nMOS del inversor con el fin de cumplir la condición de diseño, por otra lado la red PullUp tiene dos pMOS en paralelo, la resistencia de la red se toma como el peor de los casos de conducción lo cual se da en esta conmutación cuando solo un transistor conduce está es la misma situación encontrada en la red PullUp del inversor, por lo tanto la NAND2 deberá tener un FD de 2:2.

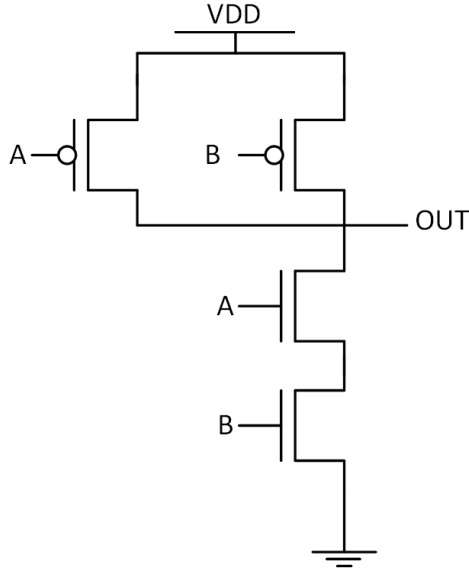


Figura 2.3: Esquemático de la compuerta NAND2

Debido a que en una compuerta lógica todos los transistores pertenecientes a una red de conmutación tienen las mismas dimensiones [3], una vez conocido el factor de diseño para una compuerta lógica se puede conocer el ancho de canal de los transistores que conforman cada red para cualquier factor de escala.

$$W_n = FD_n \times K \times W_{min} \quad (2.3)$$

$$W_p = FD_p \times K \times W_{min} \quad (2.4)$$

El análisis anterior se puede extender fácilmente a cualquier compuerta lógica para obtener su factor de diseño adecuado, en la Tabla 2.1 se muestran los factores de diseño para las compuertas digitales básicas.

En realidad, debido diversos efectos físicos la resistencia de dos o más transistores en serie es ligeramente menor que la suma de sus resistencias [12], por lo que para tener tiempos de carga y descargas idénticos este factor debe ser considerado, sin embargo el agregar este factor aumenta la complejidad matemática por lo que no será incluido en este análisis de primer orden, aun así se puede utilizar un simulador eléctrico como SPICE para definir de forma mas precisa los valores de FD para cualquier topología.

Compuerta	Factor de diseño (FD)
Inversor	1:2
NAND2	2:2
NAND3	3:2
NAND4	4:2
NOR2	1:4
NOR3	1:6
NOR4	1:8

Tabla 2.1: Factor de dimensionamiento en las compuertas digitales básicas.

Validación

Para comprobar la exactitud de las ecuaciones desarrolladas se hará uso del simulador eléctrico SPICE junto con la tecnología de 65nm, los parámetros de los transistores pertenecientes a dicha tecnología se muestran en la Tabla 2.2.

Dichos parámetros fueron obtenidos directamente del documento de modelado de los transistores por parte del proceso de Global Foundries a menos que se especifique lo contrario de forma explícita.

Parámetro	Símbolo	Valor	Unidad
Ancho de canal mínimo dibujable	W_{min}	120	nm
Largo de canal mínimo dibujable	L_{min}	65	nm
Voltaje umbral sin polarización nMOS	V_{T0_n}	442.39	mV
Voltaje umbral sin polarización pMOS	V_{T0_p}	-430.80	mV
Espesor del oxido	T_{ox}	2.39	nm
Movilidad de los portadores de carga (N) ¹	μ_n	3.30E-3	$\frac{m^2}{V \times s}$
Movilidad de los portadores de carga (P) ¹	μ_p	1.65E-3	$\frac{m^2}{V \times s}$
Permeabilidad relativa del oxido	κ_{ox}	3.9	-
Voltaje de Alimentación	V_{DD}	1.2	V
Efecto de saturación de velocidad	α	1.25	-
Capacitancia de difusión por unidad de área ¹	C_{j0}	0.703	$\frac{mF}{m^2}$
Capacitancia de difusión por unidad de perimetro ¹	C_{jsw}	155.25	$\frac{pF}{m}$

¹Obtenido por simulación.

Tabla 2.2: Parámetros para la tecnología de 65 nm

Para comparar la resistencia predicha por el modelo analítico con el simulador eléctrico se propone el circuito de prueba mostrado en la Figura 2.4, en él cual la resistencia equivalente del transistor podrá ser encontrada por medio del simulador SPICE simplemente llevando a cabo un análisis transitorio durante el tiempo de conmutación y posteriormente promediando como lo dicta la ecuación 2.1, el análisis será llevado desde el transistor de dimensiones mínimas hasta el transistor escalado 10 veces, por lo que se realizarán un total de 10 simulaciones.

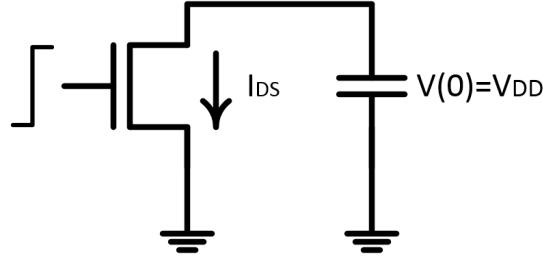


Figura 2.4: Circuito de prueba para obtener la resistencia equivalente del transistor

Por su parte para calcular la resistencia a partir del modelo analítico se requiere sustituir en la ecuación 2.1 los parámetros pertinentes de la Tabla 2.2. Comparando los resultados de dicha ecuación con los resultados de la simulación propuesta se forma la Tabla 2.3.

Escala (K)	SPICE	Modelo Desarrollado	Unidades
1	14.40	17.76	k Ω
2	7.97	8.88	k Ω
3	5.52	5.92	k Ω
4	4.14	4.44	k Ω
5	3.31	3.35	k Ω
6	2.75	2.96	k Ω
7	2.36	2.53	k Ω
8	2.06	2.22	k Ω
9	1.83	1.97	k Ω
10	1.64	1.77	k Ω

Tabla 2.3: Comparación del parámetro eléctrico de la resistencia de canal entre el modelo desarrollado y SPICE

De la tabla anterior se puede encontrar que

$$\text{Error relativo promedio} = 9,5 \% \quad (2.5)$$

2.2.2. Capacitancia de entrada

La capacitancia de entrada en las compuertas lógicas se debe a las propias capacitancias vistas en las compuertas de los transistores conectados en dicha terminal de entrada, la capacitancia de compuerta C_G de un transistor es generada debido a la separación de la compuerta y el cuerpo del semiconductor por el material dieléctrico lo que naturalmente forma un capacitor de placas paralelas, el objetivo de esta capacitancia es atraer la mayor carga posible del cuerpo del semiconductor al canal por lo que es una capacitancia intrínseca del transistor, si bien esta capacitancia al igual que la resistencia puede tener un comportamiento no lineal se puede aproximar un valor LTI por medio de la siguiente expresión [2]:

$$C_G = \frac{k_{ox}\epsilon_0}{T_{ox}} \times \text{Área} = C_{ox}(WL) \quad (2.6)$$

La capacitancia de entrada de una compuerta CMOS básica puede calcularse en base a la siguiente ecuación.

$$C_{in} = C_{G_N} + C_{G_P} \quad (2.7)$$

Esto debido a que en las compuertas básicas siempre existen dos transistores conectados a cada terminal de entrada, un transistor pMOS y uno nMOS.

Con base a los criterios de dimensionamiento discutidos en la sección anterior se tiene para una compuerta lógica:

$$C_{in} = C_{ox}((W_n L_{min}) + (W_p L_{min})) \quad (2.8)$$

Expresando la ecuación anterior en términos del factor de dimensionamiento de la compuerta lógica.

$$C_{in} = K \times W_{min} L_{min} C_{ox} \times (FD_n + FD_p) \quad (2.9)$$

Los factores de dimensionamiento pueden encontrarse por medio de la Tabla 2.1 para las compuertas digitales básicas.

Validación

La capacitancia de entrada también puede ser obtenida por simulaciones realizadas por medio de SPICE, para esto basta solicitar al simulador la medición de la capacitancia del nodo respectivo, posteriormente se promedia la capacitancia medida durante la conmutación. Por su parte la ecuación analítica puede encontrarse en 2.9 en ésta ecuación podemos ver que depende del factor de diseño propio de la topología, este factor puede encontrarse en la Tabla 2.1 para las compuertas lógicas básicas. Para llevar a cabo la comparación se escoge como topología de prueba el inversor escalado, su esquemático puede ser visto en la Figura 2.1 y su respectivo factor de dimensionamiento en la Tabla 2.1, utilizando estos datos junto a los parametros pertinentes de la Tabla 2.2 se realiza la comparación.

Escala (K)	SPICE	Modelo Desarrollado	Unidades
1	0.469	0.311	fF
2	0.797	0.623	fF
3	1.124	0.934	fF
4	1.426	1.246	fF
5	1.741	1.558	fF
6	2.055	1.869	fF
7	2.369	2.180	fF
8	2.683	2.492	fF
9	2.997	2.804	fF
10	3.312	3.115	fF

Tabla 2.4: Comparación del parámetro eléctrico de la capacitancia de entrada para un inversor entre el modelo desarrollado y SPICE

De la tabla anterior se puede encontrar que

$$\text{Error relativo promedio} = 13 \% \quad (2.10)$$

2.2.3. Auto-Capacitancia

La Auto-capacitancia (C_{self}) de una compuerta lógica es la capacitancia vista en la terminal de salida cuando no existe carga en dicha terminal, su naturaleza es debida (principalmente) a las capacitancias de drenaje C_D de los transistores nMOS y pMOS conectados en la terminal de salida [2]. La capacitancia de drenaje de un transistor puede ser encontrada por la siguiente ecuación [2].

$$C_D = C_{GD} + C_{DB} \quad (2.11)$$

De la ecuación anterior se observa que la capacitancia de drenaje tiene dos componentes una debido al acople capacitivo entre la compuerta y el drenaje C_{GD} y otra componente debida a la capacitancia que presenta la zona de difusión hacia el cuerpo del semiconductor C_{DB} . Para el primer término se puede suponer que durante la conmutación la capacitancia total de compuerta en los transistores se divide equitativamente entre el drenaje y la fuente [2].

$$C_{GD} = \frac{1}{2}C_G \quad (2.12)$$

Por otro lado la capacitancia de difusión existe debido a que en la estructura del MOSFET siempre se crean diodos parásitos entre el cuerpo del semiconductor y las regiones de difusión debido a que estas zonas siempre tienen dopado opuesto, en las compuertas lógicas básicas se diseñan tal que estos diodos siempre se encuentren polarizados en inversa [2], de la teoría básica de dispositivos semiconductores se sabe que cuando existe dicha polarización inversa se crea una región de agotamiento dentro del diodo la cual actúa físicamente como un dieléctrico cuyo ancho varía con respecto al voltaje aplicado, por lo que es posible moldear al diodo como un capacitor controlado por voltaje en este punto de operación, la ecuación que rige este fenómeno se conoce como la ecuación del varactor y es una ecuación no lineal [4].

Sin embargo debido a que el MOSFET es una estructura tridimensional la capacitancia de agotamiento es formada por dos componentes, una encontrada por las interfaces laterales de las caras de la zonas de difusión con el cuerpo del semiconductor y otra debido a la cara inferior de la zona de difusión que nuevamente hace contacto con el cuerpo del semiconductor, estas componentes son dependientes del nivel de dopada y la profundidad alcanzada por las impurezas dentro del cuerpo del semiconductor [3], sin embargo es común

caracterizar para un proceso de fabricación específico medidas más superficiales que dependen únicamente del área y el perímetro de la vista superior de una zona de difusión, con estas consideraciones claramente la capacitancia de difusión de un transistor dependerá del Layout en el que se dibuje el transistor.

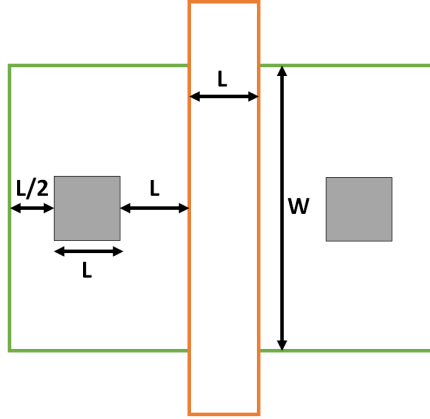


Figura 2.5: Layout sencillo de un MOSFET

Si se toma como referencia el Layout de la Figura 2.5 las medidas de área y perímetro vendrán dadas por las siguientes relaciones.

$$P_D = P_S = 2W + 2(2,5L) = 2W + 5L \quad (2.13)$$

$$A_D = A_S = W \times 2,5L \quad (2.14)$$

Con base a las dimensiones del transistor es común escribir la capacitancia de drenaje sin polarización por medio de la siguiente expresión [3]:

$$C_{DB0} = C_{J0} \times A_D + C_{JSW} \times P_D \quad (2.15)$$

Donde C_{J0} y C_{SW} son las capacitancias sin polarización de la zona activa por unidad de área y perímetro respectivamente. Recordemos que el modelo completo de la capacitancia del drenaje corresponde con la ecuación del varactor sin embargo para evitar las complejidades en el modelo que surgen debido a las no linealidades se escoge utilizar la capacitancia de peor caso para formar el modelo LTI. El valor máximo que toma la ecuación del varactor es su punto sin polarización debido a que en esta situación la separación de las zonas N y P es mínima, con base a esto el modelo LTI corresponde con la ecuación 2.15.

Con los modelos de las ecuaciones 2.12 y 2.15 ya se puede analizar el valor de C_{self} para una compuerta lógica, la cual corresponde con la suma de las capacitancias de drenaje de los transistores conectado en el nodo de salida, es decir:

$$C_{self} = \frac{1}{2}(\mathfrak{o}_n C_{G_n} + \mathfrak{o}_p C_{G_p}) + (\mathfrak{o}_n C_{DB0_n} + \mathfrak{o}_p C_{DB0_p}) \quad (2.16)$$

Donde

- \mathfrak{o}_n : Representa el número de transistores nMOS conectados a la terminal de salida.
- \mathfrak{o}_p : Representa el número de transistores pMOS conectados a la terminal de salida

Esta cantidad claramente depende de la topología propia de la compuerta, los valores específicos para las compuertas básicas se pueden encontrar en la Tabla 2.5.

Compuerta	Transistores conectados en la terminal de salida $\mathfrak{o}_n:\mathfrak{o}_p$
Inversor	1:1
NAND2	1:2
NAND3	1:3
NAND4	1:4
NOR2	2:1
NOR3	3:1
NOR4	4:1

Tabla 2.5: Número de transistores en el nodo de salida para las topologías básicas.

Si se examina a detalle la ecuación expresada anteriormente se encontrará que esta sigue un modelo lineal con respecto al factor de escalamiento K , esto queda demostrado en el Apéndice A.1.

$$C_{self} = AK(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + B(\mathfrak{o}_n + \mathfrak{o}_p) \quad (2.17)$$

Donde

$$A = \left(\frac{1}{2} L_{min} C_{ox} + 2,5 C_{j0} L_{min} + 2 C_{jsw} \right) W_{min} \quad (2.18)$$

$$B = 5 C_{jsw} L_{min} \quad (2.19)$$

Por lo general los parámetros C_{j0} y C_{jsw} no están disponibles abiertamente para una tecnología dada, sin embargo se puede seguir la ecuación 2.17 y ajustar los parámetros A y B por medio de un simulador eléctrico.

Validación

Este parámetro puede ser examinado de forma similar a la capacitancia de entrada midiendo ahora la capacitancia de salida del inversor sin carga alguna, la metodología es idéntica a la de la sección anterior, la expresión analítica se puede encontrar en la ecuación 2.17, la comparación entre la ecuación y simulación se muestra a continuación.

Escala (K)	SPICE	Modelo Desarrollado	Unidades
1	0.373	0.399	fF
2	0.706	0.704	fF
3	1.012	1.010	fF
4	1.317	1.315	fF
5	1.622	1.620	fF
6	1.927	1.926	fF
7	2.231	2.232	fF
8	2.537	2.537	fF
9	2.842	2.843	fF
10	3.147	3.143	fF

Tabla 2.6: Comparación del parámetro eléctrico de la auto-capacitancia para un inversor entre el modelo desarrollado y SPICE

De la tabla anterior se puede encontrar que

$$\text{Error relativo promedio} = 0,5 \% \quad (2.20)$$

Cabe destacar que los parámetros C_{j0} , C_{jsw} fueron obtenidos por simulación, lo que ejemplifica su gran ajuste. De los resultados anteriores vemos que $C_{in} \approx C_{self}$, lo cual es común en tecnologías nanométricas [2].

2.3. Modelo estadístico del retardo de una compuerta lógica

Una vez que es conocido la forma analítica de los parámetros eléctricos del modelo propuesto se procede a calcular el retardo de la compuerta lógica predicho por el modelo mostrado en la Figura 2.1, es de destacar que se quiere encontrar tanto el retardo nominal como la desviación estándar debida a las posibles variaciones de proceso, se tomará como referencia la tecnología de fabricación de 65nm de Global Foundries para cualquier cálculo numérico mostrado.

2.3.1. Modelo del retardo

Como las compuertas lógicas son diseñadas para obtener tiempos de carga y descarga simétricos en el nodo de salida, basta con analizar solo una conmutación para obtener el retardo nominal, tomando la conmutación de descarga se puede ver que el problema de hallar el retardo se reduce a analizar una simple red RC. Para realizar el análisis sobre la red PullDown se asume que el nodo de salida, esta cargado para los tiempos $t < 0$ al valor V_{DD} . La solución de la red RC se muestra a continuación:

$$V_{out} = V_{DD} \exp\left(\frac{-t}{R_n C_L}\right) \quad (2.21)$$

Donde:

$$C_L = C_{FO} + C_{self} \quad (2.22)$$

Para encontrar el retardo de la red basta con encontrar el tiempo de propagación de alto a bajo (t_{PHL}) del nodo de salida [3], el cual corresponde con buscar el

tiempo en el cual la salida toma el valor $\frac{V_{DD}}{2}$ durante su descarga [3], al aplicar esta situación a la ecuación 2.21 se obtiene.

$$\frac{V_{DD}}{2} = V_{DD} \exp\left(\frac{-D}{R_n C_L}\right) \quad (2.23)$$

Por lo tanto

$$D = \ln(2) R_n (C_{FO} + C_{self}) \quad (2.24)$$

Como la resistencia de una compuerta lógica se puede escribir en término de R_{DS} con base a la ecuación 2.1 se puede entonces expresar el retardo como:

$$D = \ln(2) \times \frac{V_{DD}}{\mu_{n,p} C_{ox} \frac{W}{L} (V_{DD} - V_T)^\alpha} \times (C_{FO} + C_{self}) \quad (2.25)$$

El resultado anterior es independiente de la red de conmutación activa, cuando existe más de un transistor en serie como en la compuerta NAND2 la resistencia efectiva que actúa corresponde a la resistencia de todos los transistores en la red, el número de transistores en serie dentro de una compuerta lógica se denominará como χ esta variable se introducirá mas adelante en la ecuación anterior.

Retardo en compuertas formadas por transistores en serie

El modelo de la Figura 2.1 y la ecuación del retardo calculada en 2.25 niega toda capacitancia interna de las redes PullDown y PullUp, ya que la única capacitancia considerada es la propia carga de la compuerta, para considerar estas capacitancias internas se debe modificar el modelo propuesto en la Figura 2.1, ya que se debe ver la topología de una compuerta lógica no como una única red RC, sino como una red distribuida donde existen capacitancias en los nodos internos formados por la unión de transistores en la red, para ejemplificar esto se muestra la Figura 2.6.

Claramente en este modelo el retardo ya no solo depende de la topología, escala y carga, sino también de la carga eléctrica almacenada en los nodos internos, la cual puede depender en gran medida la historia anterior a una conmutación [2].

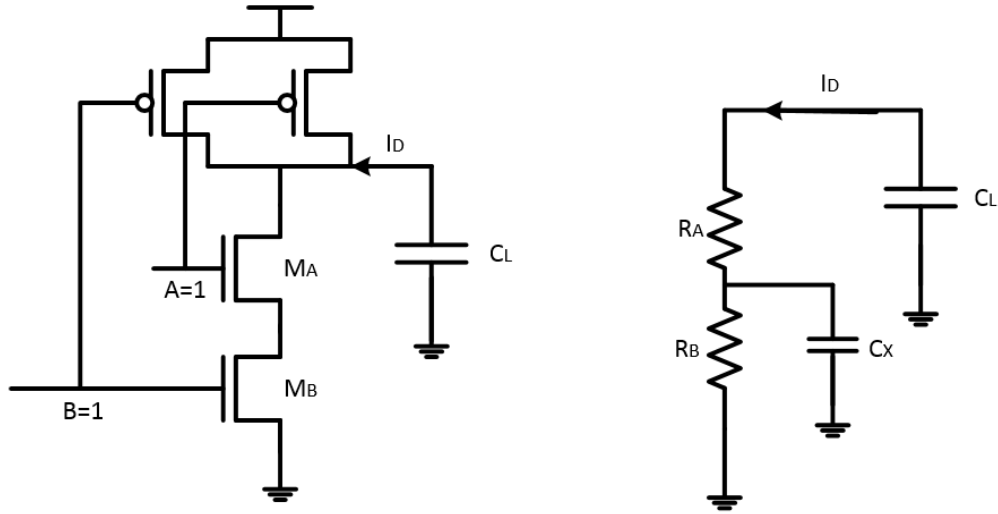


Figura 2.6: Modelo distribuido de una compuerta lógica

Por ejemplo si se tiene en un instante de tiempo el vector de entrada $[A = 1, B = 0]$, tal que el transistor M_2 forma un camino de conducción desde V_{DD} para cargar la capacitancia C_X , si la condición anterior se mantiene hasta llegar al estado estable $(V_{DD} - V_T)$ y en algún momento posterior se cambia el vector de entrada al valor $[A = 1, B = 1]$ el cual activa el transistor M_1 mismo que para producir la transición lógica en la salida no solo deberá descargar a C_L sino también a la carga almacenada en C_X , se puede aproximar el retardo de este caso de conmutación como:

$$D = \ln(2)(2R_{DS})(C_{FO} + C_{self}) + \ln(2)(R_{DS})(C_X) \quad (2.26)$$

Esta aproximación se conoce como aproximación de Elmore [2], cabe aclarar que en una red de conmutación en serie todos los transistores tienen la misma resistencia (idealmente), la suma de todas las resistencias en la red serie siempre deberá sumar a la resistencia de un inversor de la misma escala por las condiciones de diseño discutidas en la sección anterior.

Por su parte si ahora se considera la condición en la cual se tiene el vector de entrada $[A = 0, B = 1]$, el transistor M_2 no se encuentra originalmente conectado por lo cual no se almacena carga en C_X , implicando que el retardo una vez llevada acabo la conmutación se reduce a la siguiente ecuación.

$$D = \ln(2)(2R_{DS})(C_{FO} + C_{self}) \quad (2.27)$$

La capacitancia interna de una red de conmutación se forma comúnmente de la conexión entre el Source y el Drain de dos transistores idénticos, como dichas zonas son generalmente simétricas se pueden considerar que:

$$C_X = C_D + C_S \approx 2C_D = 2AK(2FD) + 2B \quad (2.28)$$

Donde los parámetros A, B se pueden encontrar por las ecuaciones 2.18, 2.19 respectivamente.

En general si se quiere tomar el retardo de peor caso de una estructura en serie deberá usarse la ecuación mostrada a continuación.

$$D = \ln(2)(\chi \times R_{DS})(C_{FO} + C_{self}) + \sum_{k=1}^{\chi} (\chi - k) \times \ln(2)(R_{DS})(C_x) \quad (2.29)$$

El parámetro χ representa el numero de transistores conectados en serie, por lo tanto el valor tomado en una evaluación depende de la topología de la compuerta analizada.

2.3.2. Modelo para la desviación estándar del retardo

Una vez conocida la forma analítica del retardo se procede a encontrar la forma analítica de la desviación estándar del retardo (σ_D) en las compuertas lógicas, para calcular dicha expresión se hará uso del modelo estadístico del retardo [5], el cual requiere primeramente calcular la sensibilidad del retardo con respecto a la/las variable(s) de proceso de interés. Para la tecnología CMOS planar las variables de proceso que más afectan al desempeño de las compuertas lógicas son las variaciones de V_T , W , L y T_{ox} [5] y serán estas las variaciones examinadas. Se empezara el análisis de la desviación estándar mostrando el cálculo de la sensibilidad del retardo con respecto al parámetro V_T con el fin de mostrar la metodología, posteriormente se ampliara directamente el resultado al resto de variaciones mencionadas (los cálculos necesarios para estas variaciones se encuentran en el Apéndice A.5.1).

En esta sección no se consideran los efectos de apilación de transistores en serie, las ecuaciones encontradas aquí serán posteriormente modificadas en una sección siguiente para incluir estos efectos.

Sensibilidad del retardo con respecto al V_T

De la teoría básica del modelo estadístico del retardo es conocido que la sensibilidad con respecto a un parámetro específico se define de la siguiente forma [5]:

$$S_{V_T}^D = \frac{\partial D}{\partial V_T} \quad (2.30)$$

Utilizando la ecuación 2.24.

$$\frac{\partial D}{\partial V_T} = \frac{\partial (\ln(2)R_{n,p}(C_L))}{\partial V_T} = \ln(2)C_L \times \frac{\partial(R_{n,p})}{\partial V_T} \quad (2.31)$$

Donde C_L representa toda la carga vista por la compuerta ($C_{FO} + C_{self}$), mientras que $R_{n,p}$ representa la resistencia de la red de conmutación activa.

Recordando que la expresión para $R_{n,p}$ puede encontrarse en la ecuación 2.1, tal que sustituyendo los parámetros adecuados:

$$R_{n,p} = \frac{V_{DD}}{\mu C_{ox} \frac{W_{n,p}}{L_{min}} (V_{DD} - |V_T|)^\alpha} \quad (2.32)$$

Cabe destacar que la expresión anterior solo es conveniente cuando la red esta formada por un solo transistor (o una conexión de transistores en paralelo), ya que esta abstrae toda la resistencia de la red en un único elemento, al derivar dicha expresión se obtiene.

$$\frac{\partial R_{n,p}}{\partial V_T} = \frac{V_{DD}}{\mu C_{ox} \frac{W_{n,p}}{L_{min}}} \frac{\partial \left(\frac{1}{(V_{DD} - V_T)^\alpha} \right)}{\partial V_T} \quad (2.33)$$

$$\frac{\partial R_{n,p}}{\partial V_T} = \frac{V_{DD}}{\mu C_{ox} \frac{W_{n,p}}{L_{min}} (V_{DD} - |V_T|)^\alpha} \times \left(\frac{\alpha}{V_{DD} - |V_T|} \right) \quad (2.34)$$

Para simplificar cálculos posteriores se escribirá la expresión encontrada sustituyendo las relaciones $C_{ox} = \frac{\epsilon_{ox}}{T_{ox}}$ y $W_{n,p} = FD \times K \times W_{min}$, tal que aplicando estos cambios y al reacomodar términos:

$$\frac{\partial R_{n,p}}{\partial V_T} = \frac{1}{K \times FD} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min} T_{ox}}{\mu \epsilon_{ox} W_{min}} \times \left(\frac{\alpha}{V_{DD} - |V_T|} \right) \quad (2.35)$$

Sustituyendo la expresión anterior en la ecuación 2.31 para encontrar la sensibilidad del retardo con respecto al parámetro V_T .

$$S_{V_T}^D = \frac{C_L}{K \times FD} \times \frac{\ln(2)V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min} T_{ox}}{\mu \epsilon_{ox} W_{min}} \times \left(\frac{\alpha}{V_{DD} - |V_T|} \right) \quad (2.36)$$

La desviación estándar del retardo debida a este parámetro se podrá encontrar entonces por medio del modelo estadístico del retardo el cual menciona que:

$$\sigma_{D_{V_T}}^2 = \left(S_{V_T}^D \times \sigma_{V_T} \right)^2 \quad (2.37)$$

Por lo tanto

$$\sigma_{D_{V_T}} = S_{V_T}^D \times \sigma_{V_T} \quad (2.38)$$

Si se utiliza la versión simplificada de la expresión anterior es de notar que se deben eliminar todo signo negativo en la magnitud total de la sensibilidad y/o la desviación estándar ya que la expresión original relaciona términos cuadráticos los cuales absorben el signo.

Por su parte la desviación estándar propia de un parámetro se puede aproximar a través de la hoja de caracterización del proceso de fabricación simplemente dividiendo entre tres el valor de variación máxima.

$$\sigma_{V_T} = \frac{\% \text{Variación Máxima} \times V_T}{3} \quad (2.39)$$

Sin embargo esto último no será del todo correcta en la práctica para el parámetro V_T debido al efecto conocido como fluctuación aleatoria de dopantes o RDF por sus siglas en ingles [5], este efecto modifica el valor del σ_{V_T} predicho anteriormente de manera inversamente proporcional al área del transistor considerado, dicho esto se puede agregar un término extra a la ecuación 2.39 para moldear efectos de RDF [5]:

$$\sigma_{V_T} = \sigma_{V_{T0}} \times \sqrt{\frac{W_{min}L_{min}}{W \times L_{min}}} \quad (2.40)$$

Donde $\sigma_{V_{T0}}$ corresponde con el término calculado en 2.39. Expresando la ecuación anterior en términos de la escala y factor de diseño de la compuerta lógica, se obtiene:

$$\sigma_{V_T} = \sigma_{V_{T0}} \times \sqrt{\frac{1}{K \times FD}} \quad (2.41)$$

Cabe destacar que la inclusión del efecto RDF proviene de trabajar con una tecnología nanométrica, estos efectos sobre la modificación de la desviación estándar de un parámetro específico son características puramente del proceso de fabricación y no de la física de los transistores por lo que no se pueden moldear por medio de parámetros eléctricos.

Expresiones para todas las variaciones de proceso consideradas

Claramente podemos ver que la metodología utilizada para el análisis del $\sigma_{D_{V_T}}$ puede utilizarse para examinar otro tipo de variaciones de proceso, por lo que al generalizar el análisis de la desviación estándar para todas las variaciones de interés se obtienen las siguientes expresiones:

$$S_L^D = \frac{\ln(2)C_L}{K \times FD} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{T_{ox}}{\mu\epsilon_{ox}W_{min}} \quad (2.42)$$

$$S_{T_{ox}}^D = \frac{\ln(2)C_L}{K \times FD} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min}}{\mu\epsilon_{ox}W_{min}} \quad (2.43)$$

$$S_W^D = -\frac{\ln(2)C_L}{K^2 \times FD^2} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min}T_{ox}}{\mu\epsilon_{ox}W_{min}} \times \frac{1}{W_{min}} \quad (2.44)$$

La demostración de estas expresiones se encuentra anexa en el Apéndice [A.5.1](#), para el cálculo de dichas expresiones se supuso que la capacitancia de carga es mucho mayor que la auto-capacitancia, lo cual es una situación común en los diseños digitales.

La desviación estándar de cada parámetro estudiado se pueden obtener conociendo las variaciones máximas del proceso de fabricación y utilizando las siguientes expresiones.

$$\sigma_W = \frac{\% \text{Variación Máxima} \times W_{min}}{3} \quad (2.45)$$

$$\sigma_L = \frac{\% \text{Variación Máxima} \times L_{min}}{3} \quad (2.46)$$

$$\sigma_{T_{ox}} = \frac{\% \text{Variación Máxima} \times T_{ox}}{3} \quad (2.47)$$

En base al modelo estadístico del retardo se expresa el impacto de cada parámetro en la desviación estándar de retardo actuando de forma aislada como:

$$\sigma_{D_{V_T}} = S_{V_T}^D \times \sigma_{V_T} \quad (2.48)$$

$$\sigma_{D_W} = S_W^D \times \sigma_W \quad (2.49)$$

$$\sigma_{D_L} = S_L^D \times \sigma_L \quad (2.50)$$

$$\sigma_{D_{T_{ox}}} = S_{T_{ox}}^D \times \sigma_{T_{ox}} \quad (2.51)$$

Si se utilizan estas ecuaciones se debe ignorar cualquier signo negativo proveniente de la sensibilidad por lo comentado en los resultados obtenidos de la expresión [2.38](#).

La desviación estándar total del retardo generada por cada una de las variaciones de proceso actuando al mismo tiempo se puede encontrar usando el modelo estadístico del retardo [5].

$$\sigma_D = \sqrt{\sigma_{D_{V_T}}^2 + \sigma_{D_W}^2 + \sigma_{D_L}^2 + \sigma_{D_{T_{ox}}}^2} \quad (2.52)$$

Generalización de la desviación estándar para estructuras en serie

Como se menciona anteriormente el modelo de σ_D desarrollado moldea a toda la red de conmutación como un único elemento resistivo, lo cual es correcto si la red esta conformada por un único transistor o una combinación exclusivamente paralela de dichos elementos (ya que en el modelo el calculo del σ_D es marcada para el peor caso), si existe una conexión en serie como en la red PullDown de la compuerta NAND2 (ver Figura 2.3) es claro que la resistencia se encuentra distribuida en dos elementos idénticos (si conmutan simultaneamente).

Cuando los transistores se implementan codo a codo como lo es dentro de una compuerta lógica estos muestran un fenómeno llamado correlación total [5], en el cual las variables W, L, T_{ox} de todos los transistores se pueden tratar como una única variable (en general la correlación total es un caso especial de la correlación espacial con un coeficiente de correlación $\rho = 1$, la correlación espacial se discutirá a detalle en la sección dedicada a caminos lógicos), por su parte la variable V_T no muestra fenómenos de correlación a este ultimo tipo de variables se le denomina como variables independientes.

Se puede analizar el σ_D producido por variables correlacionadas por medio del modelo estadístico del retardo, examinando variaciones únicamente en W para el caso de dos transistores en serie, se puede escribir:

$$\sigma_{D_W}^2 = \sigma_{D_{W_i}}^2 + \sigma_{D_{W_j}}^2 + 2\text{COV}(D_i, D_j) \quad (2.53)$$

Expandiendo

$$\sigma_{D_W}^2 = (S_{W_i}^D \times \sigma_W)^2 + (S_{W_j}^D \times \sigma_W)^2 + 2(\rho \times S_{W_i}^D \times S_{W_j}^D \times \sigma_W \times \sigma_W) \quad (2.54)$$

Como la correlación es total implica que $\rho = 1$, por su parte ya que los transistores son parte de la misma compuerta lógica implica que estos ven la misma carga y

posen las mismas dimensiones por lo cual usando la ecuación 2.44 se puede ver que:

$$S_{W_i}^D = S_{W_j}^D = S_W^D \quad (2.55)$$

Sustituyendo el resultado anterior en la ecuación 2.54

$$\sigma_{D_W}^2 = 4(S_W^D \times \sigma_W)^2 \quad (2.56)$$

Claramente esta condición se repite si se toma cualquier otra variable que presente correlación total, sin embargo esta situación no se presenta para la variable independiente del V_T , ya que el termino de covarianza es cero, escribiendo esto en forma de ecuación.

$$\sigma_{D_{V_T}}^2 = (S_{V_{T_i}}^D \times \sigma_{V_T})^2 + (S_{V_{T_j}}^D \times \sigma_{V_T})^2 \quad (2.57)$$

Como nuevamente todos los transistores presentan la misma carga y escala dentro de una compuerta, se puede usar la situación.

$$S_{V_{T_i}}^D = S_{V_{T_j}}^D \quad (2.58)$$

Tal que

$$\sigma_{D_{V_T}}^2 = 2(S_{V_{T_i}}^D \times \sigma_{V_T})^2 \quad (2.59)$$

Por inspección es fácil ver que en una red formada por χ transistores en serie las variables correlacionadas incrementan su contribución en la varianza de forma proporcional al cuadrado del numero de transistores en serie, mientras que las variables independientes incrementan solo de forma lineal, expresando esto en la ecuación 2.52 se obtiene una versión mas general de la desviación estándar de una compuerta lógica.

$$\sigma_D = \sqrt{\chi \left(\sigma_{D_{V_T}}^2 \right) + \chi^2 \left(\sigma_{D_W}^2 + \sigma_{D_L}^2 + \sigma_{D_{T_{ox}}}^2 \right)} \quad (2.60)$$

La ecuación anterior toma en cuenta los efectos de correlación espacial existentes entre los transistores de una compuerta lógica, pero niega las capacitancias internas de las estructuras en serie como la vista en la Figura 2.6, estos efectos son directos de considerar si se utiliza el modelo de Elmore.

2.3.3. Modelo compacto del retardo estadístico

Las ecuaciones anteriores forman todo el marco teórico del modelo propuesto, sin embargo todas las expresiones están en términos de los parámetros tecnológicas de los transistores, lo que hace las ecuaciones anteriores algo voluminosas, para evitar esto se puede expresar la ecuación del retardo nominal de una compuerta lógica vista en la ecuación 2.25 de una forma más compacta al agrupar términos de manera inteligente con el fin de mostrar la ecuación con una dependencia más clara de los parámetros de diseño como son la carga (C_L) y la geometría de los transistores (K), este modelo es mostrado a continuación:

$$D = \frac{\kappa}{\beta_0} \times \frac{C_L}{K} \times \frac{1}{FD} \quad (2.61)$$

Donde:

$$\beta_0 = \frac{\mu\epsilon_{ox}W_{min}}{T_{ox}L_{min}} \quad (2.62)$$

$$\kappa = \ln(2) \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \quad (2.63)$$

Los términos mostrados en las ecuaciones 2.62 y 2.63 fueron escogidos con el fin de separar lo más posible los parámetros dependientes del proceso de fabricación de aquellos basados en la alimentación.

Cabe recalcar que el término 2.62 al depender de la movilidad varia según el tipo de red de conmutación activa.

La ecuación 2.61 es correcta siempre que para una compuerta, la carga o descarga del nodo de salida dependa únicamente de un transistor nMOS o pMOS, cuando

se tienen redes en serie de transistores el retardo total de la compuerta se puede ver como la suma de cada componente individual, tal que se puede demostrar fácilmente que si se tienen χ transistores en serie, entonces.

$$D = \left(\frac{\kappa}{\beta_0} \times \frac{C_L}{K} \times \frac{\chi}{FD} \right) + \sum_{k=1}^{\chi} \left(\frac{\kappa}{\beta_0} \times \frac{C_X}{K} \times \frac{\chi - k}{FD} \right) \quad (2.64)$$

Si se evalúa en los valores medios de los parametros sujetos a variaciones se obtiene la definición del retardo nominal de la compuerta.

$$\mu_D = \left(\frac{\kappa}{\beta_0} \times \frac{C_L}{K} \times \frac{\chi}{FD} \right) + \sum_{k=1}^{\chi} \left(\frac{\kappa}{\beta_0} \times \frac{C_X}{K} \times \frac{\chi - k}{FD} \right) \Big|_{\Delta P_i=0, \forall i} \quad (2.65)$$

A lo largo de este documento se sobreentenderá que siempre que se use el símbolo μ_D implica que la evaluación deberá darse en el valor medio de todos los parámetros que conforman la ecuación.

De forma similar se puede expresar las sensibilidades del retardo a las variaciones de proceso de una forma compacta, utilizando las variables definidas anteriormente, dichas expresiones a la vez se pueden sustituir dentro de las ecuaciones del modelo estadístico, para encontrar el valor de la desviación estándar del retardo producida por cada parámetro en una forma más amena.

$$\sigma_{D_{V_T}} = \frac{\kappa}{\beta_0} \times \frac{C_L}{(K \times FD)^{1.5}} \times \delta_{V_T} \quad (2.66)$$

$$\sigma_{D_W} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K^2 \times FD^2} \times \delta_{D_W} \quad (2.67)$$

$$\sigma_{D_L} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K \times FD} \times \delta_L \quad (2.68)$$

$$\sigma_{D_{T_{ox}}} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K \times FD} \times \delta_{T_{ox}} \quad (2.69)$$

Donde:

$$\delta_{V_T} = \left(\frac{\alpha}{V_{DD} - |V_T|} \times \sigma_{V_{T0}} \right) \quad (2.70)$$

$$\delta_W = \left(\frac{1}{W_{min}} \times \sigma_W \right) \quad (2.71)$$

$$\delta_L = \left(\frac{1}{L_{min}} \times \sigma_L \right) \quad (2.72)$$

$$\delta_{T_{ox}} = \left(\frac{1}{T_{ox}} \times \sigma_{T_{ox}} \right) \quad (2.73)$$

La demostración de dichas ecuaciones se encuentra en el Apéndice [A.5.2](#). Es de destacar que en la expresión [2.67](#) se eliminó el signo negativo a propósito por los resultados discutidos de la ecuación [2.38](#).

Utilizando la ecuaciones anteriores para definir la desviación estándar total del retardo debida al efecto de cada variación de proceso actuando de forma simultanea se tiene:

$$\sigma_D = \left(\frac{\kappa}{\beta_0} \frac{C_L}{K} \right) \sqrt{\left(\frac{\delta_{V_T}}{\sqrt{K} \times FD^{1,5}} \right)^2 + \left(\frac{\delta_W}{K \times FD^2} \right)^2 + \left(\frac{\delta_L}{FD} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD} \right)^2} \quad (2.74)$$

La derivación de esta expresión se obtiene directamente al sustituir las ecuaciones [2.66](#) a [2.69](#) en [2.52](#), dicha expresión es el resultado principal del modelo de la Figura [2.1](#) ya que nos permite conocer la desviación estándar total de una compuerta lógica, al conocer simplemente su carga (C_L), escala (K) y topología (FD), el resto de parámetros del modelo son constantes tecnológicas que sirven para ajustar la expresión a diferentes procesos de fabricación (Se asume directamente que el efecto RDF esta presente por lo que es un modelo más adecuado para tecnologías nanométricas).

Si se consideran los efectos de serie en las compuertas, la ecuación anterior es modificada a:

$$\sigma_D = \left(\frac{\kappa C_L}{\beta_0 K} \right) \sqrt{\chi \left(\frac{\delta_{V_T}}{\sqrt{K} F D^{1.5}} \right)^2 + \chi^2 \left(\left(\frac{\delta_W}{K F D^2} \right)^2 + \left(\frac{\delta_L}{F D} \right)^2 + \left(\frac{\delta_{T_{ox}}}{F D} \right)^2 \right)} + \left(\frac{\kappa C_X}{\beta_0 K} \right) \sum_{k=1}^{\chi} \left(\sqrt{(\chi - k) \left(\frac{\delta_{V_T}}{\sqrt{K} F D^{1.5}} \right)^2 + (\chi - k)^2 \left(\left(\frac{\delta_W}{K F D^2} \right)^2 + \left(\frac{\delta_L}{F D} \right)^2 + \left(\frac{\delta_{T_{ox}}}{F D} \right)^2 \right)} \right) \quad (2.75)$$

Se recuerda que en la ecuación anterior el parámetro χ hace referencia al numero de transistores en serie dentro de la red, para las compuertas básicas solo existen transistores en serie en la red PullDown de las compuertas NAND y en la red PullUp de las compuertas NOR, en ambos casos el numero de transistores en serie corresponde con el numero de entradas de la compuerta.

2.3.4. Modelo del retardo estadístico considerando el efecto stack

Una apilación de transistores en serie se denomina como Stack, en dichas estructuras existen consideraciones que afectan al retardo estadístico aparte de la capacitancias agregadas por los nodos intermedios y la correlación espacial discutidas anteriormente, los fenómenos a discutir hacen que la resistencia de cada transistor en una red de conmutación varíe de transistor a transistor generando comportamientos que no son modelados en las ecuaciones anteriores.

Entre los principales efectos que afectan este comportamiento son:

- Efecto cuerpo
- Diferencias en los punto de operación de cada transistor

Efecto cuerpo

El efecto cuerpo es un cambio en el voltaje umbral del transistor generada por una diferencia de potencial entre la fuente y el sustrato ya que cuando un potencial superior de 0[V] es aplicado entre estas terminales la región de agotamiento de

los diodos parásitos se agranda dificultando el encendido del transistor (es decir aumenta el valor efectivo del voltaje umbral) ya que el potencial a vencer visto por V_{GS} aumenta para generar la inversión [10], matemáticamente esta fenómeno se moldea como:

$$V_T = V_{T0} + \gamma(\sqrt{2\phi + V_{SB}} - \sqrt{2\phi}) \quad (2.76)$$

Donde V_{T0} es el voltaje umbral sin polarización, γ se denomina coeficiente de cuerpo y es un parámetro tecnológico mientras que ϕ es el potencial de superficie el cual es un parámetro físico de cada material semiconductor. Así como la carga almacenada en los nodos intermedios depende del historial de conmutación, también el efecto cuerpo ya que los niveles de voltaje de los nodos intermedios dependen de esta historia. En un Stack de transistores siempre que exista un camino de conducción de V_{DD} a un nodo intermedio del Stack dicho nodo se cargara hasta un valor máximo $V_{DD} - V_T$ dicho valor se toma basado en la teoría de transistores de paso [3], como en general se toma el sustrato del chip como la tierra del circuito este potencial en los nodos intermedios sera el mismo potencial que produce el efecto cuerpo.

Si se examina la estructura mostrada en la Figura 2.7 por medio de simulaciones eléctricas con el fin de analizar la evolución del potencial de los nodos intermedios bajo consideraciones de conmutación de peor caso se podrá encontrar la Figura 2.8.

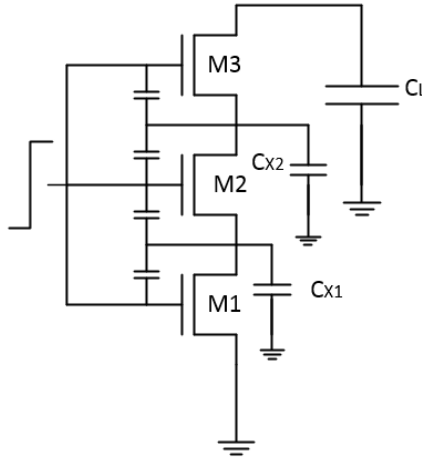


Figura 2.7: Stack de tres transistores nMOS mostrando todas las capacitancias parásitas de la red

Debido al aumento en el voltaje umbral dado por el efecto cuerpo se puede ver

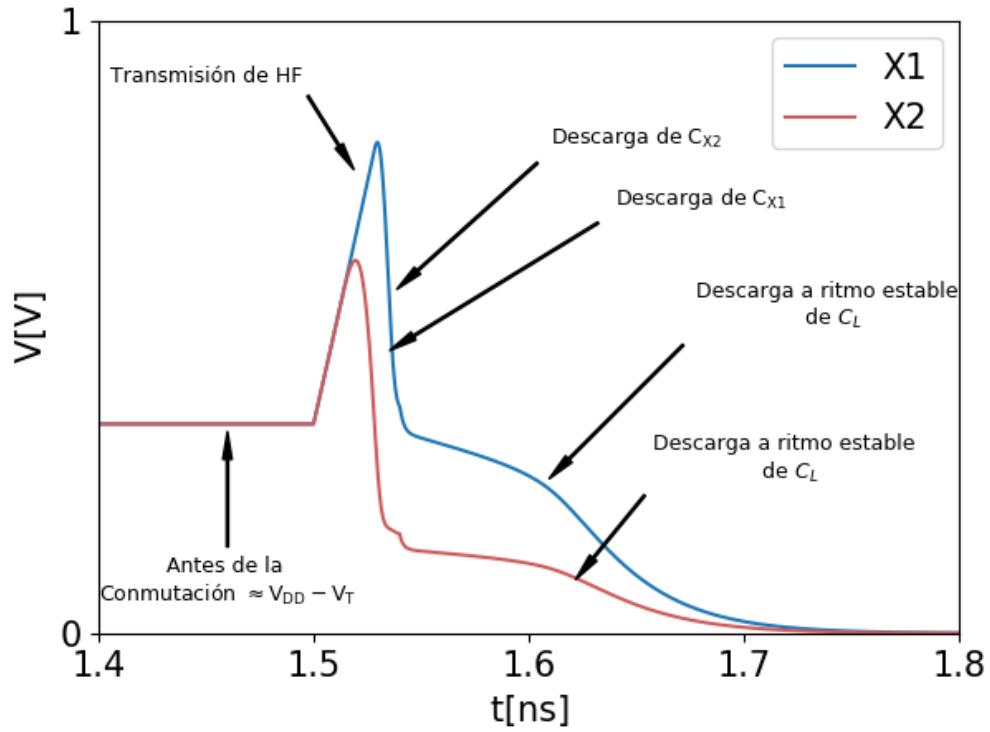


Figura 2.8: Evolución temporal de los nodos intermedios

por la ecuación 2.1 que la resistencia del MOSFET aumentara con respecto a lo predicho si solo se usa el valor nominal de V_T .

Diferencias en los punto de operación de cada transistor

Siguiendo el ejemplo de la red mostrada en la Figura 2.7 y los resultados discutidos se puede tomar para $t = t_0$ los valores

$$V_{OUT} = V_{DD} \quad (2.77)$$

$$V_{X1} = V_{DD} - V_{T2} \quad (2.78)$$

$$V_{X2} = V_{DD} - V_{T3} \quad (2.79)$$

Si se calcula el valor del V_{DS} para cada transistor para este instante, con lo que se encuentra:

$$V_{DS3} = (V_{DD}) - (V_{DD} - V_{T3}) = V_{T3} \quad (2.80)$$

$$V_{DS2}(0) = (V_{DD} - V_{T3}) - (V_{DD} - V_{T2}) = V_{T3} - V_{T2} \approx 0[V] \quad (2.81)$$

$$V_{DS1}(0) = V_{DD} - V_{T2} \quad (2.82)$$

Claramente los puntos de operación de cada transistor son diferentes entre si y en conjunto todos son diferentes a la situación encontrada por la Figura 2.4, dicha Figura fue la base para calcular la ecuación 2.1 lo que hace a esta ultima poco adecuada para moldear redes en Stack.

Visualizando por medio de simulaciones la evolución de los valores de V_{DS} para cada transistor en la red se genera la Figura 2.9.

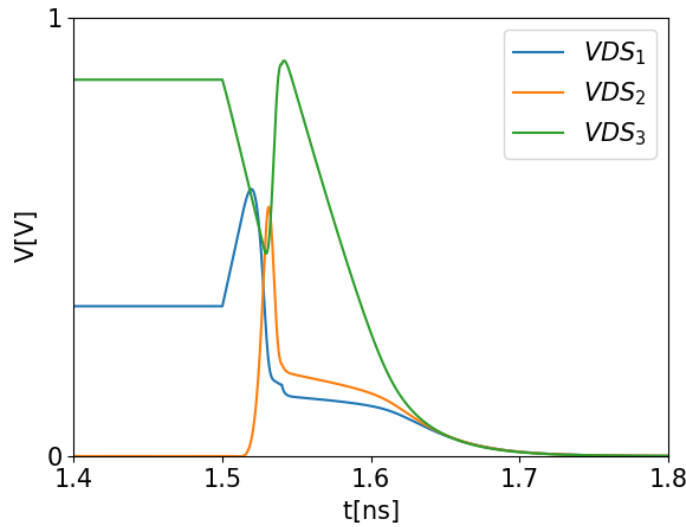


Figura 2.9: Evolución temporal de los valores de V_{DS}

Corrección del modelo

Cabe destacar que calcularla resistencia promedio para cada transistor en esta estructura por medio de la solución directa de la red en Stack tiene poco valor practico para los modelos analíticos debido a la complejidad del calculo que provee una ecuación engorrosa con poco insight, una alternativa empírica surge

al reutilizar el modelo de la ecuación 2.1 pero ponderando la participación en la red de cada transistor por medio de un coeficiente de corrección pre-caracterizado para la tecnología ξ .

La base de este modelo se encuentra en que la resistencia promedio en cada transistor de un Stack sera diferente en magnitud a la del resto de transistores, pero se comportara con la misma tendencia que se estudio en el modelo dado por la ecuación 2.1.

Para calcular la resistencia promedio del canal por medio de simulaciones eléctricas se puede usar la siguiente ecuación.

$$R_{SPICE} = \frac{\int_{t_0}^{t_f} V_{DS} dt}{\int_{t_0}^{t_f} I_{DS} dt} \quad (2.83)$$

Dicha ecuación tiene su origen en el teorema del valor medio, cabe destacar que t_0 marca el punto de inicio de la conmutación, mientras que t_f marca el final de la misma.

Para que el valor de R_{SPICE} calculado corresponda con el peor caso de conmutación es necesario pre-cargar los nodos internos, esto se puede hacer aplicando una serie de vectores a la compuerta bajo análisis antes de ejecutar la conmutación a examinar.

Por ejemplo para la estructura NAND2 mostrada en la Figura 2.10

Se debe aplicar la siguiente secuencia de vectores para generar la pre-carga del nodo intermedio.

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}_{AltaImpedancia} \rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{Pre-carga} \rightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}_{AltaImpedancia} \rightarrow \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{Conmutación} \quad (2.84)$$

La evolución temporal de los nodos una vez aplicados los vectores correspondientes se muestra a continuación.

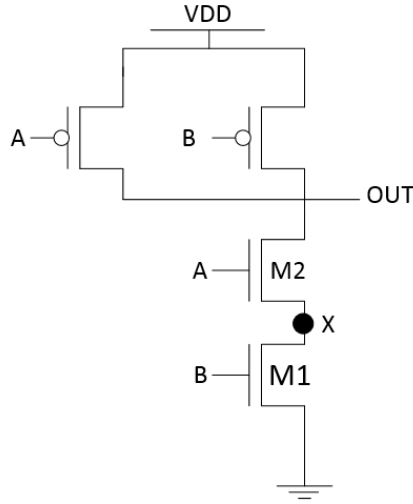


Figura 2.10: Estructura NAND2 y su nodo intermedio

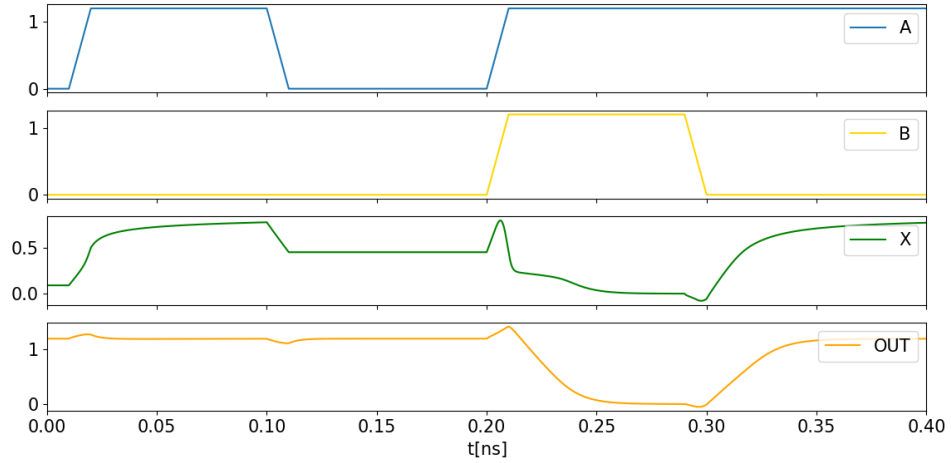


Figura 2.11: Formas de onda en una estructura NAND2

El ejemplo de la diferencia de la resistencia predicha por el modelo analítico y el valor promedio de la simulación eléctrica se muestra en la Tabla 2.7.

Definiendo el factor de corrección como:

$$\xi = \frac{R_{SPICE}}{R_{MA}} \quad (2.85)$$

De esta forma la resistencia del Stack de dos transistores se puede expresar como:

Transistor	$R_{MA}[\text{k}\Omega]$	$R_{SPICE}[\text{k}\Omega]$	Error
M1	8.9	5.3	40.44 %
M2	8.9	10.9	-22.92%
Total	17.6	16.2	7.95 %

Tabla 2.7: Comparación de la resistencia de canal predicha por el modelo analítico y SPICE

$$R_T = \xi_1 R_{MA_1} + \xi_2 R_{MA_2} \quad (2.86)$$

Si se expande el análisis anterior a una red de χ transistores en Stack se encuentra que todos los transistores encontrados en el cuerpo de la red (exceptuando los extremos) tendrán prácticamente las mismas condiciones eléctricas (ver ecuación 2.81) lo que genera un comportamiento distinguido en el cual transistores mas cercanos al suministro de alimentación de la red tendrán una resistencia menor en promedio (ya que poseen caminos mas cortos de descarga) que los transistores superiores, por su parte el transistor mas cercano al nodo de salida siempre tendrá una resistencia mucho mayor debido a la actuación del efecto cuerpo y su punto de polarización, de forma similar el transistor conectado a suministro de salida no sigue el comportamiento monotónico debido a su punto de operación (ver ecuación 2.82).

Calculando los valores de ξ para las diferentes estructuras NAND.

Estructura	Transistor	$R_{SPICE}[\text{k}\Omega]$	$R_{MA}[\text{k}\Omega]$	ξ
NAND2	M1	5.30	8.80	0.59
	M2	10.94	8.80	1.23
NAND3	M1	3.30	5.92	0.56
	M2	2.77	5.92	0.47
	M3	9.38	5.92	1.58
NAND4	M1	2.23	4.44	0.50
	M2	1.79	4.44	0.40
	M3	2.76	4.44	0.62
	M4	8.60	4.44	1.93

Tabla 2.8: Valor de ξ para las topologías NAND

El parámetro ξ altera la expresión del retardo nominal ya que la contribución de cada transistor de la red esta ponderada, se puede visualizar del análisis del retardo nominal que al sustituir la contribución anterior la ecuación del retardo nominal de una compuerta lógica sera:

$$\mu_D = \left(\frac{\kappa}{\beta_0} \times \frac{C_L}{K} \times \frac{\sum_{i=1}^X (\xi_i)}{FD} \right) + \sum_{k=1}^X \left(\frac{\kappa}{\beta_0} \times \frac{C_X}{K} \times \frac{\sum_{i=1}^{X-k} (\xi_i)}{FD} \right) \quad (2.87)$$

En la ecuación anterior el factor ξ_1 representa la participación del transistor con su fuente conectado al suministro de alimentación (GND en redes nMOS y V_{DD} para las redes pMOS).

Así como la expresión del retardo nominal las nuevas contribución afectan al σ_D , sin embargo estas contribuciones afectan de diferente forma a las variaciones independientes de las correlacionadas.

Por ejemplo analizando variaciones solo en el W para una compuerta en Stack de dos transistores.

$$\sigma_{D_W}^2 = (\xi_1 S_{W_1}^D \sigma_W)^2 + (\xi_2 S_{W_2}^D \sigma_W)^2 + 2\rho \times \xi_1 S_{W_1}^D \sigma_W \times \xi_2 S_{W_2}^D \sigma_W \quad (2.88)$$

Como los transistores ven la misma situación eléctrica

$$S_{W_1}^D = S_{W_2}^D = S_W^D \quad (2.89)$$

Por la cercanía de los transistores en la celda

$$\rho = 1 \quad (2.90)$$

Tal que:

$$\sigma_{D_W}^2 = (S_W^D)^2 (\sigma_W)^2 \times (\xi_1^2 + \xi_2^2 + 2\xi_1 \xi_2) \quad (2.91)$$

Se puede interpretar el factor ξ_i como la ponderación de la contribución del transistor i .

Para el caso de variaciones independientes se puede demostrar

$$\sigma_{D_{V_T}}^2 = (S_{V_T}^D)^2 (\sigma_W)^2 \times (\xi_1^2 + \xi_2^2) \quad (2.92)$$

Combinando los resultados anteriores se puede expresar el σ_D como:

$$\begin{aligned} \sigma_D = & \left(\frac{\kappa}{\beta_0} \right) \times \\ & \left\{ \frac{C_L}{K} \times \sqrt{(\xi_1^2 + \xi_2^2) \left(\frac{\delta_{V_T}}{\sqrt{K}FD^{1,5}} \right)^2 + (\xi_1^2 + \xi_2^2 + 2\xi_1\xi_2) \left(\left(\frac{\delta_W}{KFD^2} \right)^2 + \left(\frac{\delta_L}{FD} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD} \right)^2 \right)} \right. \\ & \left. + \frac{C_X}{K} \times \left(\sqrt{(\xi_1)^2 \left(\frac{\delta_{V_T}}{\sqrt{K}FD^{1,5}} \right)^2 + (\xi_1^2) \left(\left(\frac{\delta_W}{KFD^2} \right)^2 + \left(\frac{\delta_L}{FD} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD} \right)^2 \right)} \right) \right\} \end{aligned} \quad (2.93)$$

Expandiendo los resultados a una estructura de Stack de cualquier numero de transistores.

$$\begin{aligned} \sigma_D = & \left(\frac{\kappa}{\beta_0} \right) \times \left\{ \frac{C_L}{K} \sqrt{\left(\sum_{i=1}^{\chi} \xi_i^2 \right) \left(\frac{\delta_{V_T}}{\sqrt{K}FD^{1,5}} \right)^2 + \left(\sum_{i=1}^{\chi} \xi_i^2 + 2 \sum_{i=1}^{\chi} \sum_{j=i+1}^{\chi} \xi_i^2 \xi_j^2 \right) \left(\left(\frac{\delta_W}{KFD^2} \right)^2 + \left(\frac{\delta_L}{FD} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD} \right)^2 \right)} \right. \\ & \left. \frac{C_X}{K} \times \sum_{k=1}^{\chi} \left(\sqrt{\left(\sum_{i=1}^{\chi-k} \xi_i^2 \right) \left(\frac{\delta_{V_T}}{\sqrt{K}FD^{1,5}} \right)^2 + \left(\sum_{i=1}^{\chi-k} \xi_i^2 + 2 \sum_{i=1}^{\chi-k} \sum_{j=i+1}^{\chi-k} \xi_i \xi_j \right) \left(\left(\frac{\delta_W}{KFD^2} \right)^2 + \left(\frac{\delta_L}{FD} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD} \right)^2 \right)} \right) \right\} \end{aligned} \quad (2.94)$$

Si se supone en la ecuación anterior que $C_L \gg C_X$ lo cual es una situación común en los diseños digitales típicos entonces la ecuación 2.94 se puede simplificar a la siguiente expresión.

$$\sigma_D = \left(\frac{\kappa}{\beta_0} \right) \times \left\{ \frac{C_L}{K} \sqrt{(\Lambda) \left(\frac{\delta_{V_T}}{\sqrt{K}FD^{1,5}} \right)^2 + (\Upsilon) \left(\left(\frac{\delta_W}{KFD^2} \right)^2 + \left(\frac{\delta_L}{FD} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD} \right)^2 \right)} \right\} \quad (2.95)$$

Donde:

$$\Lambda = \sum_{x=1}^{\chi} \xi_x^2 \quad (2.96)$$

$$\Upsilon = \sum_{x=1}^{\chi} \xi_x^2 + 2 \sum_{x=1}^n \sum_{x_2=x+1}^{\chi} \xi_x \xi_{x_2} \quad (2.97)$$

2.3.5. Validaciones

Si se toma como referencia la tecnología de fabricación de 65nm de Global Foundries la cual presenta en los parámetros discutidos variaciones máximas del 15% en las variaciones de proceso estudiadas se tiene en conjunto con los parámetros de la tecnología información suficiente para computar la desviación estándar del retardo, bajo esta misma suposición de variación se puede usar SPICE para realizar simulaciones del tipo MonteCarlo con el fin de encontrar el valor de σ_D de una manera precisa, los resultados de dicho experimento se muestran en la Figura 2.12 para el inversor.

Validaciones para otro tipo de compuertas se pueden encontrar en el Apéndice A.6.

Analizando los diversos parámetros del modelo se encontró que la principal fuente de error es proveniente del término C_{self} ya que este término recordemos sigue la ecuación del varactor pero, para el modelo analítico se supuso un término LTI, se puede demostrar que el error del modelo analítico es proporcional al porcentaje de participación de C_{self} en C_L , por lo tanto el modelo solo es adecuado mientras se garantice la situación expresada donde $C_{FO} \gg C_{self}$, para mostrar lo anterior de forma gráfica se hace uso de la Figura 2.13

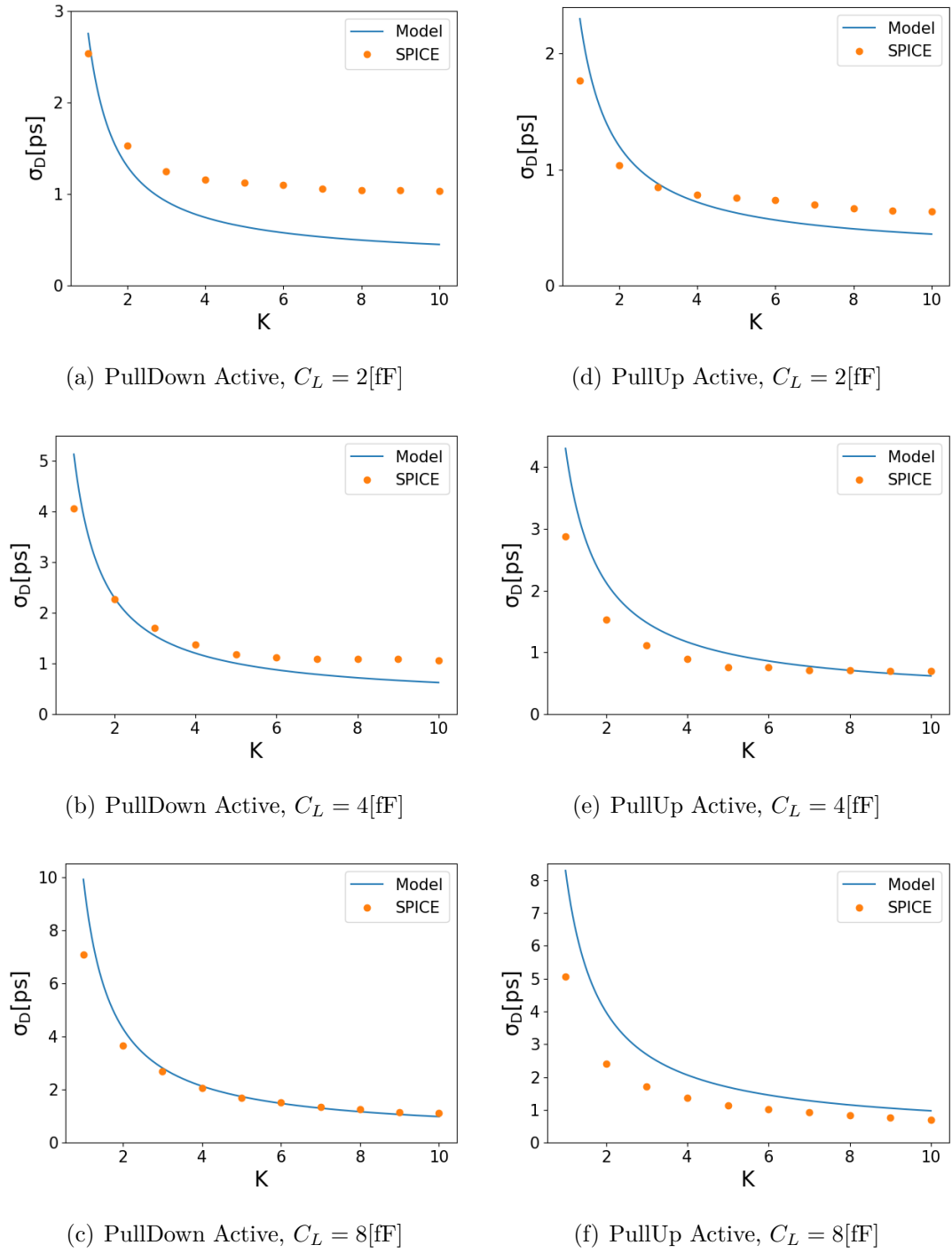


Figura 2.12: Comparación entre el modelo analítico del σ_D y SPICE, en el Inversor

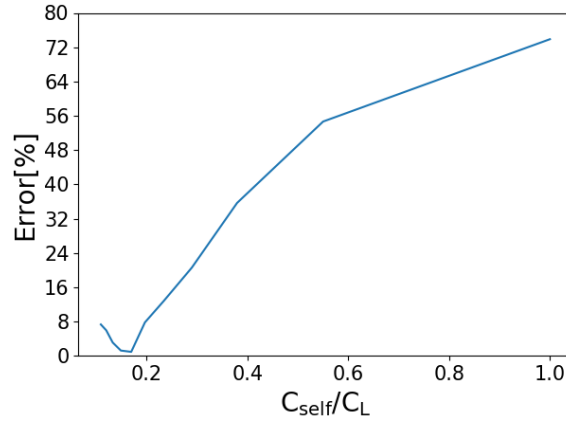


Figura 2.13: Error en función de C_{self} en el σ_D para el Modelo Analítico, probado en un inversor con $K=2$

2.4. Modelo estadístico del retardo en caminos lógicos

Si bien anteriormente se proporcionaron las ecuaciones para estudiar el retardo estadístico en cualquier compuerta lógica, en la práctica los diseñadores están más interesados en estudiar caminos lógicos los cuales son solamente conexiones en cascada de compuertas como lo es mostrado en la Figura 2.14.

El análisis de caminos lógicos como es de esperarse está basado en la aplicación de las ecuaciones desarrolladas combinadas nuevamente con el modelo estadístico del retardo [6].

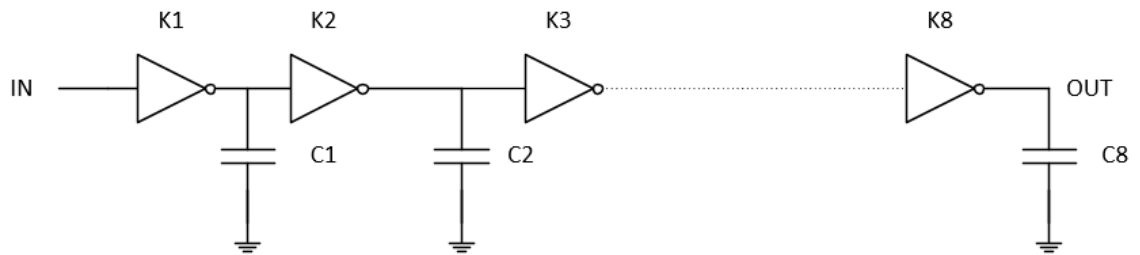


Figura 2.14: Ejemplo de camino lógico

En los caminos lógicos la capacitancia de carga total C_L en cada nodo puede hallarse sumando la auto capacitancia de la compuerta cuya salida esté conectada a dicho nodo, más la capacitancia de entrada de la compuerta siguiente, más la capacitancia extrínseca marcada en el esquemático la cual generalmente es

utilizada para moldear efectos de interconexión y/o cargas en el nodo dadas por compuertas que están fuera del camino.

Con base al modelo estadístico del retardo se puede calcular el retardo nominal de un camino lógico simplemente como la suma de los retardos de cada compuerta en el camino [6], es decir:

$$\mu_P = \sum_{i=1}^n \mu_{D_i} \quad (2.98)$$

Con el modelo analítico desarrollado se puede expresar el retardo de cada etapa por medio de la ecuación 2.87.

La desviación estándar de un camino lógico por su parte es calculada mediante la suma estadística, la cual es propuesta en el modelo estadístico del retardo, la estructura matemática encontrada en dicho modelo se da en la siguiente ecuación.

$$\sigma_P^2 = \sum_{i=1}^n \sigma_{D_i}^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{COV}(D_i, D_j) \quad (2.99)$$

Donde claramente los términos σ_{D_i} pueden ser encontrados con base a la ecuación 2.95, por su parte el segundo termino se denomina como covarianza y se expresa en el modelo estadístico del retardo como la suma de productos de los σ_{D_P} individuales de cada compuerta ponderadas por medio de un coeficiente de correlación ρ .

$$\text{COV}(D_i, D_j) = \sum_{k=1}^n \rho_{i,j} \times (S_{P_k,i}^D \sigma_{P_k,i}) \times (S_{P_k,j}^D \sigma_{P_k,j}) \quad (2.100)$$

El coeficiente de correlación es una medida que determina el grado con el cual dos variables correlacionadas tienden a variar conjuntamente [13], generalmente dicho coeficiente tiene un comportamiento decreciente con respecto a la distancia de separación entre las compuertas bajo análisis [5].

Como muestra el Apéndice A.2 la ecuación 2.100 se puede expresar en función de los términos definidos en la sección anterior, tal que:

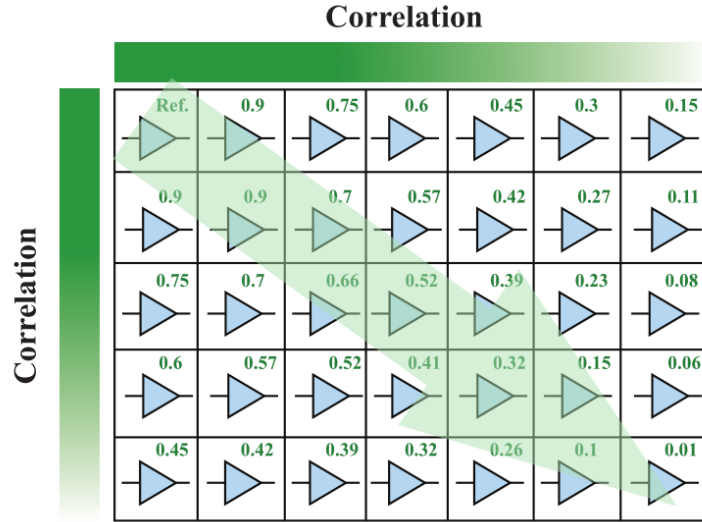


Figura 2.15: Comportamiento de la correlación entre compuertas [17]

$$\text{COV}(D_i, D_j) = \left[\left(\frac{\kappa^2}{\beta_{0i}\beta_{0j}} \right) \left(\frac{C_{L_i}}{K_i} \frac{C_{L_j}}{K_j} \right) \right] \times \left[\rho_{V_{T_{i,j}}} \left(\frac{\sqrt{\Lambda_i \Lambda_j} \delta_{V_T^2}}{\sqrt{(K_i)(K_j)(FD_i FD_j)^{1,5}}} \right) + \frac{\sqrt{\Upsilon_i \Upsilon_j}}{FD_i FD_j} \left(\rho_{W_{i,j}} \left(\frac{\delta_W^2}{K_i K_j FD_i FD_j} \right) + \rho_{L_{i,j}} \delta_L^2 + \rho_{T_{ox,i,j}} \delta_{T_{ox}}^2 \right) \right] \quad (2.101)$$

Donde

$$\Lambda_x = \sum_{x=1}^{\chi} \xi_x^2 \quad (2.102)$$

$$\Upsilon_x = \sum_{x=1}^{\chi} \xi_x^2 + 2 \sum_{x=1}^n \sum_{x_2=x+1}^{\chi} \xi_x \xi_{x_2} \quad (2.103)$$

Recordando que las variaciones en el parámetro V_T son independientes [6], lo que implica que:

$$\rho_{V_{T_{i,j}}} = 0 \quad (2.104)$$

Por su parte las variaciones de W, L y T_{ox} son del tipo correlacionadas, y se puede asumir que las correlaciones entre estos parámetros se distribuyen de la misma forma [6]:

$$\rho_{W_{i,j}} = \rho_{L_{i,j}} = \rho_{T_{ox_{i,j}}} = \rho_{i,j} \quad (2.105)$$

Existen en la literatura multitud de modelos de correlación espacial, siendo el modelo exponencial [6] el modelo más común utilizado para el análisis estadístico, dicho modelo se muestra a continuación.

$$\rho_{i,j} = \exp\left(\frac{-d_{i,j}}{CD}\right) \quad (2.106)$$

Donde $d_{i,j}$ representa la distancia entre las compuertas i y j , por su parte CD es la llamada distancia de correlación, la cual es específica del proceso de fabricación y corresponde con la mitad de la longitud del dado [5].

Bajo estas consideraciones la ecuación 2.101 se puede simplificar:

$$\text{COV}(D_i, D_j) = \rho_{i,j} \times \frac{\kappa^2}{\beta_{0i}\beta_{0j}} \times \frac{C_{L_i}C_{L_j}\sqrt{\Upsilon_i\Upsilon_j}}{FD_iFD_j} \left(\frac{\delta_W^2}{K_i^2K_j^2FD_iFD_j} + \frac{\delta_L^2}{K_iK_j} + \frac{\delta_{T_{ox}}^2}{K_iK_j} \right) \quad (2.107)$$

2.4.1. Validación

Claramente los cálculos realizados para computar al retardo estadístico requieren el uso de las ecuaciones anteriormente desarrolladas, el unico termino externo es la covarianza especificada por la ecuación 2.107. Para validar la ecuación 2.107 se deben utilizar nuevamente simulaciones eléctricas sin embargo SPICE solo es capaz de generar variables aleatorias independientes. Es posible construir un conjunto de variables aleatorias correlacionadas a partir de un conjunto de variables aleatorias independientes utilizando la transformación de Cholesky. Una explicación detallada de la transformación de Cholesky se da en el Apéndice A.7.

La topología usada para la validación de la covarianza se muestra en la Figura 2.16.

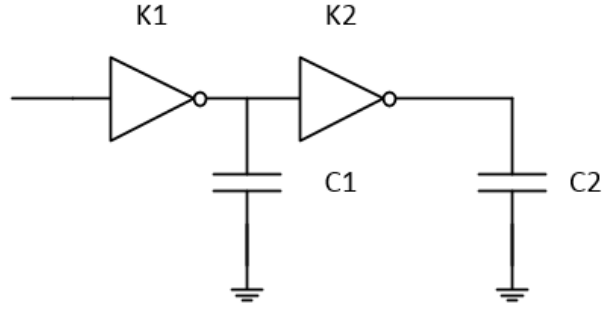


Figura 2.16: Camino de validación de la covarianza

Tomando como condición $\rho = 1$ se procede a realizar un análisis de cortes bajo las variables K_1, K_2, C_1, C_2 de todo el espacio de trabajo de la covarianza, los resultados se muestran a continuación en la Figura 2.17.

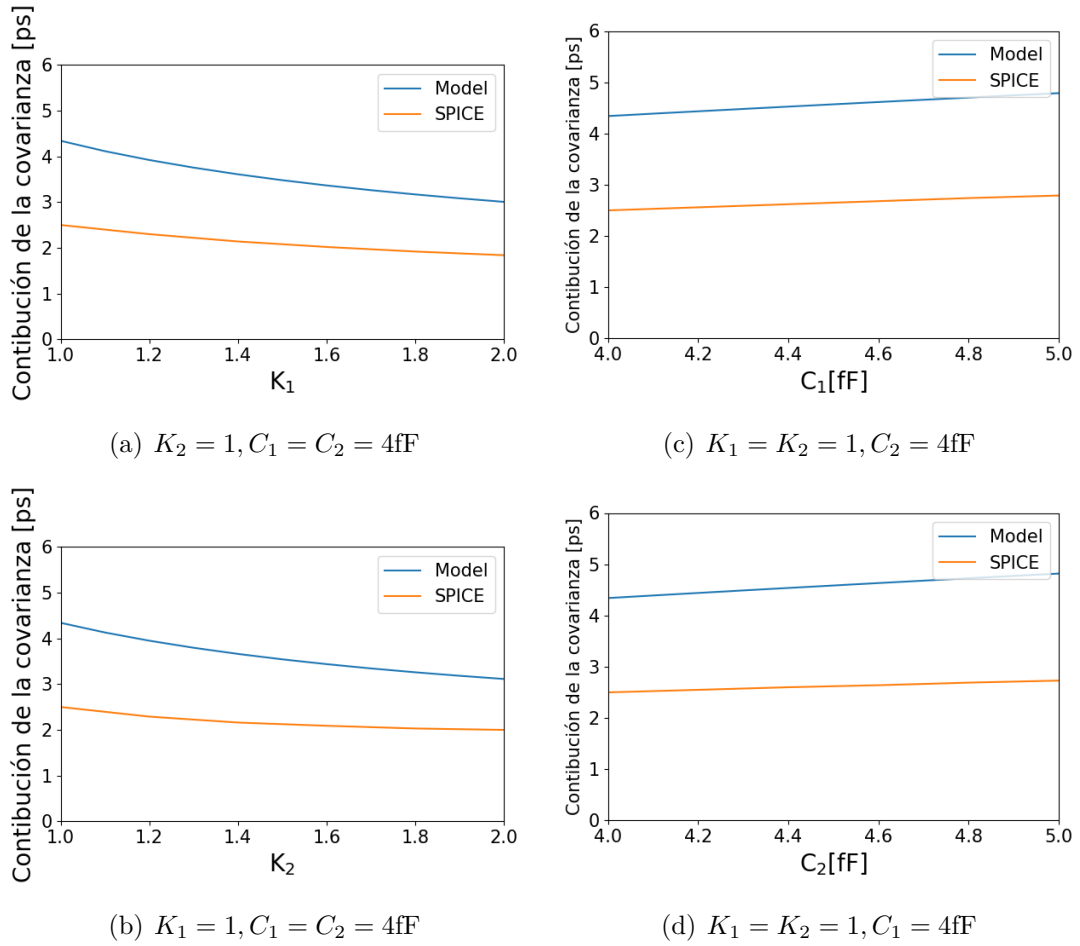


Figura 2.17: Validación de la covarianza

Si bien existe un error entre el modelo propuesto y los resultados obtenidos por medio de SPICE (apoyado por la transformada de Cholesky) es importante destacar que la tendencias de ambas gráficas son prácticamente idénticas.

2.5. Conclusiones

El uso de modelos analíticos es una herramienta que no debe pasar desapercibida a un diseñador de circuitos integrados, ya que estos permiten tener una visión interna del fundamento matemático implicado en los sistemas, si bien es posible utilizar modelos más precisos por medio de herramientas computacionales estos modelos funcionan típicamente como cajas negras por lo que su uso en el análisis y estudio es limitado, por su parte los modelos analíticos carecen de la precisión de dichos modelos computacionales pero permiten una libertad al diseñador a la hora de realizar tareas de análisis, los modelos propuestos en este capítulo buscan brindar un apoyo en dichas tareas de análisis con respecto al efecto de las variaciones de proceso encontradas en los sistemas VLSI, las cuales debido al constante escalamiento tecnológico son cada vez más necesarias de examinar durante el flujo de diseño para garantizar el correcto funcionamiento de los IC's.

Para el desarrollo del modelo analítico propuesto se ha tomado como base que una compuerta lógica puede ser vista como una sencilla red RC conmutada (Figura 2.1) donde los parámetros eléctricos de dicha red se expresan en función de los parámetros de los transistores que conforman a la compuerta, directamente el modelo RC puede usarse para computar el retardo de la compuerta.

Al aplicar en el retardo de la compuerta operaciones diferenciales se puede encontrar la sensibilidad de dicho retardo a cambios en los parámetros de los transistores en este trabajo se consideran como parámetros de interés al conjunto W, L, T_{ox} y V_T por ser estos los parámetros que definen principalmente el comportamiento de los transistores, logrado el cálculo de las sensibilidades se aplico el modelo estadístico del retardo para plantear las ecuaciones finales del modelo analítico las cuales se evaluaron tanto a nivel compuerta como a nivel camino lógico.

Capítulo 3

Optimización de caminos lógicos

Muchas veces los diseñadores de circuitos integrados están interesados en realizar diseños robustos frente a variaciones de proceso sobretodo en tecnologías nanométricas las cuales llegan a ser muy sensibles a estas variaciones [5], de las ecuaciones definidas en el capitulo anterior se vio que las variaciones de proceso producen una desviación estándar en el valor del retardo por lo que pueden afectar a no cumplir un criterio de velocidad en los circuitos.

Podemos ver en la ecuación 2.95 que el σ_D para una compuerta con topología definida es inversamente proporcional a la escala de la compuerta y directamente proporcional a la carga de la misma, evaluar la ecuación para diferentes escalas y cargas nos proporciona una familia de curva las cuales siguen la misma tendencia como se observa en la Figura 3.1.

Sin embargo la desviación estándar de la compuerta también es dependiente de la topología debido al factor FD el cual a la vez depende del tipo de conmutación (red activa), los valores de dicho parámetro podrán encontrar en la Tabla 2.1, para mostrar esta dependencia topológica se incluyen las siguientes figuras:

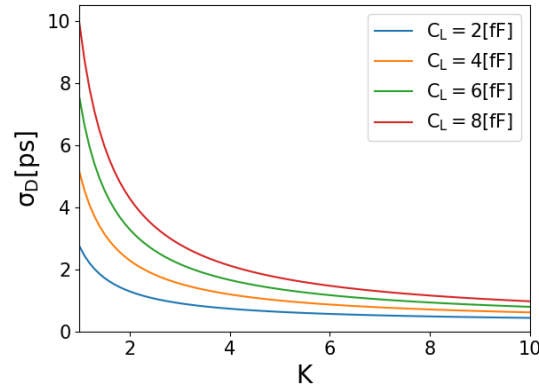


Figura 3.1: Familia de curvas de la desviación estándar en un inversor obtenidas por el modelo analítico propuesto

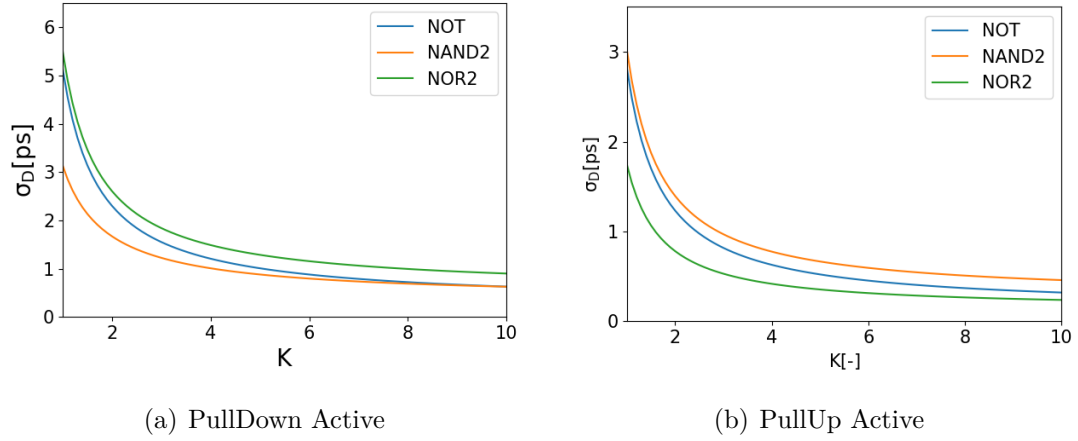


Figura 3.2: Comparación del σ_D en las compuertas básicas con $C_L = 4[\text{fF}]$

3.1. Modelo para la optimización de la desviación estándar del retardo de una compuerta lógica

En las Figuras anteriores claramente la pendiente de la gráfica indican el ritmo de cambio de σ_D con respecto a la escala de la compuerta, es conocido del cálculo diferencial que la pendiente puede ser encontrada por medio de la derivada, un valor grande en dicha derivada indica que el valor de σ_D se reduce significativamente con respecto a un pequeño aumento de escala en la compuerta, por el contrario un valor bajo en la derivada produce una reducción pequeña para el mismo aumento de área, por lo tanto si se quiere reducir la

desviación estándar del camino gastando la menor área posible la derivada puede servir como una guía de diseño valiosa, dicho cálculo se puede realizar por medio de la ecuación 2.95, el cálculo de la misma se muestra en el Apéndice A.3.

$$\frac{\partial \sigma_D}{\partial K} = - \frac{\Lambda \sigma_{D_{V_T}} \frac{\partial \sigma_{D_{V_T}}}{\partial K} + \Upsilon \left(\sigma_{D_W} \frac{\partial \sigma_{D_W}}{\partial K} + \sigma_{D_L} \frac{\partial \sigma_{D_L}}{\partial K} + \sigma_{D_{T_{ox}}} \frac{\partial \sigma_{D_{T_{ox}}}}{\partial K} \right)}{\sqrt{\Lambda \left(\sigma_{D_{V_T}} \right)^2 + \Upsilon \left(\left(\sigma_{D_W} \right)^2 + \left(\sigma_{D_L} \right)^2 + \left(\sigma_{D_{T_{ox}}} \right)^2 \right)}} \quad (3.1)$$

Donde

$$\frac{\partial \sigma_{D_{V_T}}}{\partial K} = -1,5 \frac{\kappa}{\beta_0} \frac{C_L \delta_{V_T}}{FD^{1,5}} \left(\frac{1}{K^{2,5}} \right) \quad (3.2)$$

$$\frac{\partial \sigma_{D_W}}{\partial K} = -2 \frac{\kappa}{\beta_0} \frac{C_L \delta_W}{FD^2} \left(\frac{1}{K^3} \right) \quad (3.3)$$

$$\frac{\partial \sigma_{D_L}}{\partial K} = - \frac{\kappa}{\beta_0} \frac{\delta_L C_L}{FD} \left(\frac{1}{K^2} \right) \quad (3.4)$$

$$\frac{\partial \sigma_{D_{T_{ox}}}}{\partial K} = - \frac{\kappa}{\beta_0} \frac{\delta_{T_{ox}} C_L}{FD} \left(\frac{1}{K^2} \right) \quad (3.5)$$

Por su parte los valores propios del σ_D ya fueron calculados anteriormente en las ecuaciones 2.66, 2.69.

Para evaluar la precisión obtenida entre la ecuación anterior, nuevamente se hace uso del simulador SPICE, los resultados se muestran a continuación.

Esta y otras comprobaciones se pueden inferir por medio de la Figura 2.12, así como en el Apéndice A.6 para otras topologías.

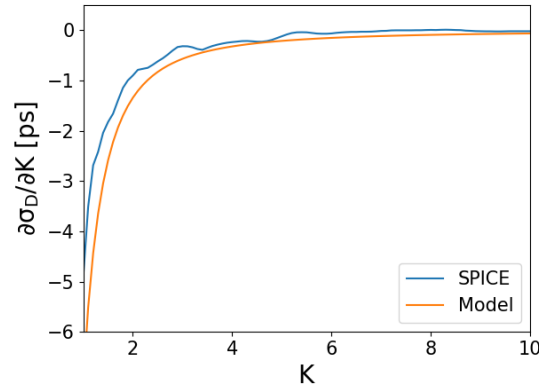


Figura 3.3: Comparación entre la derivada del σ_D del modelo analítico y SPICE, en un inversor con $C_L = 4[\text{fF}]$

Es común utilizar la ecuación 3.1 como métrica en los diseños VLSI al evaluar la derivada para cada compuerta en un camino lógico siendo la compuerta con el valor mas negativo el mejor candidato a dimensionar ya que esta compuerta será la que reducirá en mayor forma su desviación estándar para un incremento de escala dado, si bien se pueden usar modelos mas complejos y exactos de la desviación estándar en una compuerta para realizar el cálculo anterior, el uso la forma analítica definida en 3.1 permite observar directamente el impacto de cada variable de diseño dentro la expresión, mientras que los modelos mas exactos son por lo general oscuros al diseñador.

3.2. Modelo para la optimización de la desviación estándar del retardo de un camino lógico

Si bien la metodología propuesta anteriormente se puede usar como una primera guía de diseño, esta se basa en suponer que la desviación estándar del camino se reduce en la mayor medida si se aumenta el tamaño de las compuertas que agregan la mayor contribución independiente, sin embargo el estudio de la ecuación 2.99 muestra que la metodología anterior no es de todo cierta por dos principales razones.

- No toma en cuenta las correlaciones espaciales.

- En un camino lógico la compuerta i sirve como carga a la compuerta $i - 1$, esto quiere decir que si la compuerta i aumenta en escala esto produce un aumento en la carga de $i - 1$, lo que se traduce en un aumento proporcional en la desviación estándar de dicha compuerta.

En general la derivada de la varianza de un camino lógico con respecto a la escala de una etapa se puede obtener al diferenciar la ecuación 2.99, el resultado se muestra a continuación.

$$\frac{\partial \sigma_P^2}{\partial K_i} = 2\sigma_{D_i} \frac{\partial \sigma_{D_i}}{\partial K_i} + 2\sigma_{D_{i-1}} \frac{\partial \sigma_{D_{i-1}}}{\partial K_i} + 2 \sum_{j \neq i}^x \frac{\partial \text{COV}(D_i, D_j)}{\partial K_i} \quad (3.6)$$

La ecuación 3.6 es independiente del modelo analítico por lo cual puede usarse como métrica incluso en modelos mas complejos, sin embargo esta ecuación es una aproximación ya que muchos fenómenos que ocurren en caminos reales no están presentes, por ejemplo aparte de la correlación espacial en un camino real existe un segundo tipo de correlación dada por la pendiente de la señal de entrada proveniente de la etapa previa, esto es así por que el retardo de una compuerta lógica en realidad no solo depende de la carga y capacidad de conducción sino de como esta es conducida por su etapa anterior este fenómeno genera una correlación fuerte entre compuertas adyacentes [5], sin embargo este fenómeno es difícil de moldear por modelos de primer orden.

Con lo anteriormente discutido es fácil notar que en la ecuación 3.6 los siguientes efectos no están considerados

- La contribución de la covarianza de señal producida entre la compuerta i e $i + 1$, debida a que la compuerta i funciona directamente como el driver de $i - 1$.
- La contribución de la covarianza de señal producida entre la compuerta i e $i - 1$, la cual se genera debido a que el aumento de K_i aumenta la carga de la compuerta $i - 1$, lo que hace que la pendiente de la señal de entrada disminuya.

Analizando la ecuación 3.6 se observa que el primer término diferencial corresponde con el uso de la ecuación 3.1, por su parte el término de la covarianza se obtiene simplemente derivando parcialmente la expresión 2.107 con respecto a la escala de la compuerta i .

$$\frac{\partial \text{COV}(D_i, D_j)}{\partial K_i} = -\rho_{i,j} \times \frac{\kappa^2}{\beta_{0i}\beta_{0j}} \times \frac{C_{L_i}C_{L_j}\sqrt{\Upsilon_i\Upsilon_j}}{FD_iFD_j} \left(\frac{2\delta_W^2}{K_i^3K_j^2FD_iF_j} + \frac{\delta_L^2}{K_i^2K_j} + \frac{\delta_{T_{ox}}^2}{K_i^2K_j} \right) \quad (3.7)$$

La ecuación anterior es correcta para la mayoría de los casos encontrados sin embargo falla en una situación especial cuando se tienen compuertas adyacentes ya que en esta situación puede surgir un nuevo termino debido a que la carga de la compuerta con menor posición en la cadena se vuelve una función de la escala encontrada en la compuerta posterior esto agrega un termino extra a la ecuación anterior en el caso $j = i - 1$ sin embargo dicho no sera considerado para mantener las ecuaciones lo mas compactas posibles, también se considero para el calculo de esta ecuación la condición $C_{FO} \gg C_{self}$.

Por su parte para el término diferencial de la etapa anterior se puede calcular a partir de la expresión 2.95, dicho calculo se encuentra en el Apéndice A.4.

$$\begin{aligned} \frac{\partial \sigma_{D_{i-1}}}{\partial K_i} &= (FD_{i_n} + FD_{i_p}) \left(\frac{\kappa W_{min} L_{min} C_{ox}}{\beta_{0i-1} K_{i-1}} \right) \\ &\sqrt{\Lambda_{i-1} \left(\frac{\delta_{V_T}}{\sqrt{K_{i-1}} \times FD_{i-1}^{1.5}} \right)^2 + \Upsilon_{i-1} \left(\left(\frac{\delta_W}{K_{i-1} \times FD_{i-1}^2} \right)^2 + \left(\frac{\delta_L}{FD_{i-1}} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD_{i-1}} \right)^2 \right)} \end{aligned} \quad (3.8)$$

Cabe destacar que no todas las compuertas gastan la misma cantidad de área para un mismo incremento en su escala ya que su dimensionamiento base y número de transistores dentro de la compuerta son muy variantes dependiendo la topología, con base a los factores de diseño definidos en la Tabla 2.1 se puede expresar el área de una compuerta lógica en función de su escala por la siguiente expresión.

$$\text{Área de compuerta} = K \times (\#n\text{MOS} \times FD_n + \#p\text{MOS} \times FD_p) [W_{min} L_{min}] \quad (3.9)$$

El número de transistores nMOS y pMOS que conforman una compuerta es muy dependiente de la topología, sin embargo para las compuertas básicas (Inversores, NAND's y NOR's) el número de nMOS y pMOS siempre es el mismo y corresponde con el número de entradas de la compuerta.

Una alternativa al uso de las métricas es usar el conjunto de ecuaciones mostradas en esta sección con las cuales es posible armar un problema típico de optimización para un camino lógico específico, en el cual se puede escoger como

función objetivo minimizar el área (3.9), la desviación estándar (2.99) o el retardo nominal (2.98) de dicho camino, tomando como variables de decisión las escalas requeridas para cada compuerta en el camino lógico se puede llevar a cabo un dimensionamiento automático, para obtener resultados dentro de un rango de diseño factible es posible utilizar una de las ecuaciones anteriores diferente a la función objetivo seleccionada como una función de restricción, si bien dicha optimización se puede llevar a cabo utilizando los modelos matemáticos completos de una tecnología específica dicho cálculo será ralentizado por la amplia complejidad de dichos modelos volviéndolo en la mayoría de los casos inviable, por su parte los modelos discutidos en este capítulo son lo suficientemente simples como para usarse en un tiempo razonable aún con circuitos grandes, el costo a pagar es una reducción de la exactitud de los cálculos la cual no debe tomarse a la ligera.

3.3. Optimización bajo restricción de área

En esta sección se utilizará el modelo analítico para la selección óptima del tamaño de las compuertas que conforman los caminos lógicos, la forma en la que se va a aplicar esta selección es por medio de métricas que permitan cuantificar el impacto de cada una de las compuertas en el camino lógico. Se realizarán dos tipos de optimizaciones diferentes para mostrar la versatilidad del modelo propuesto en ambas se usará el método iterativo discutido en la sección anterior. Se validarán los resultados del modelo analítico con respecto al uso directo de la tecnología de 65nm en el simulador eléctrico SPICE. En la primera de las optimizaciones se trabajará con un camino lógico compuesto únicamente por inversores y se dará una explicación detallada del proceso de optimización para las situaciones:

1. Sin considerar la correlación espacial y un dimensionamiento de compuertas continuo.
2. Sin considerar la correlación espacial y un dimensionamiento de compuertas discretizado.
3. Considerando correlación espacial y un dimensionamiento de compuertas continuo.

La segunda optimización se realizará para un camino lógico genérico solamente analizándolo para la situación tres ya que es la más general.

3.3.1. Optimización de un camino lógico compuesto solo por inversores

El camino a optimizar se muestra en la Figura 3.4.

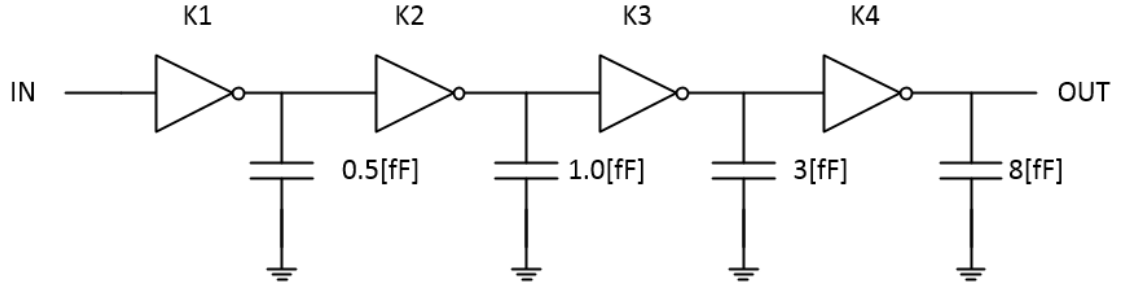


Figura 3.4: Ejemplo de camino lógico

Sera considerado que la señal aplicada en el nodo de entrada conmuta de la forma $0 \rightarrow 1$ con este dato y por medio de la Tabla 2.1 se infiere que:

$$FD_1 = FD_3 = 1 \quad (3.10)$$

$$FD_2 = FD_4 = 2 \quad (3.11)$$

Con base a está ultima situación se revela que el σ_P es una función que depende de la señal de entrada aplicada, en estos ejemplos solo se optimizara la transición marcada sin considerar como esto afecta al σ_P de la otra transición, los valores expresados anteriormente serán comunes a todos los tipos de optimización que se llevarán acabo.

Optimización sin correlación espacial y aumento continuo en el tamaño de las compuertas logicas

Para simplificar los cálculos se empezara por definir el proceso de optimización sin considerar la correlación espacial del camino en cuestión.

Anteriormente se había definido una formula para calcular el área de una compuerta lógica en la ecuación 2.14, claramente el área que consume un camino

lógico es simplemente la suma del área de cada compuerta que compone a dicho camino.

$$A_P = \sum_{i=1}^n A_i \quad (3.12)$$

Comúnmente los diseñadores de IC's están interesados en minimizar el valor del σ_P para un total de área disponible (A_C), esta situación plantea directamente una restricción de diseño que se pueden expresar por medio de la siguiente ecuación.

$$A_P \leq A_C \quad (3.13)$$

Cuando se tiene una topología lógica el área disponible puede utilizarse con el fin de incrementar el desempeño global del camino lo cual se logra aumentando de forma selectiva la escala de las compuertas, tal que la solución del problema de optimización es básicamente encontrar la forma de *repartir* el área disponible a las diferentes compuertas del camino con el fin de conseguir el menor valor de σ_P , para lograr lo anterior se puede evaluar la ecuación 3.6 para cada compuerta ya que esta indica directamente la cantidad de impacto que generará sobre el camino aumentar la escala de la compuerta evaluada, en este contexto se denominará como métrica de la compuerta (M) al resultado de evaluar la ecuación 3.6 sobre una compuerta específica, si al evaluar todas las compuertas se escoge aquella con métrica mas negativa con el fin de incrementar la escala de dicha compuerta en una factor ΔK se obtendrá de esta forma la mayor cantidad de reducción en el σ_P , si se repite este proceso de forma iterativa hasta alcanzar la restricción impuesta por la ecuación 3.13 se habrá utilizado el área para dimensionar el camino de forma que se alcance la mayor reducción del σ_P con el área disponible.

El diagrama de flujo que describe el proceso anterior se muestra en la Figura 3.5.

El parámetro ΔK se denomina paso de iteración y su valor define tanto la complejidad del computo requerido como la precisión de los resultados, conociendo que el rango normal de escalas en las compuertas digitales abarca desde $K = 1$ hasta $K = 20$ se escoge el valor $\Delta K = 0,2$ con el fin de dividir el rango en aproximadamente 100 puntos.

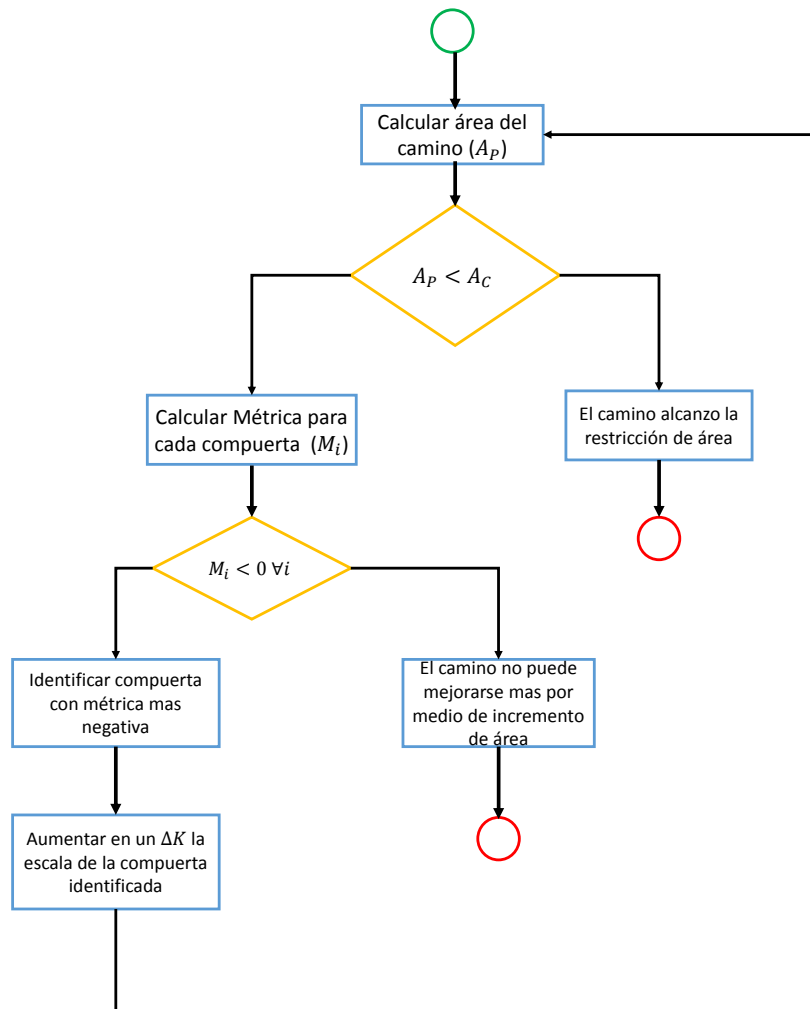


Figura 3.5: Diagrama de flujo del proceso de optimización

Antes de realizar la optimización, se recomienda estudiar el espacio de trabajo el cual se define como el conjunto de todos los diseños factibles (tamaños de compuertas) que se pueden obtener de la topología, el espacio de trabajo se gráfica con respecto al área del camino en la Figura 3.6 en la cual se supuso la condición $A_C = 27$.

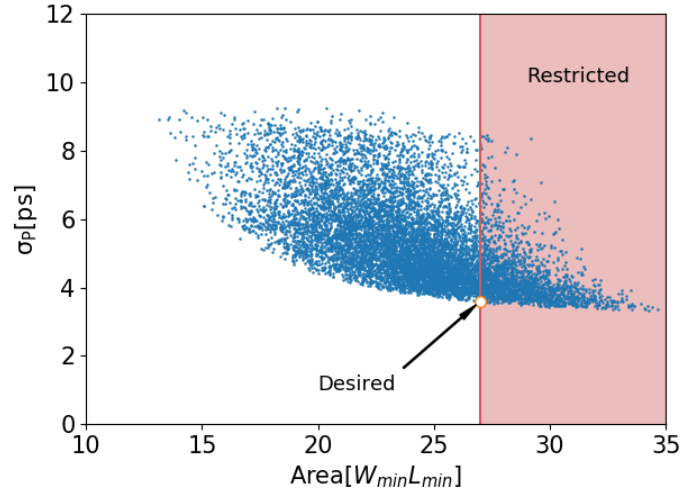


Figura 3.6: Espacio de optimización para el camino de prueba con $A_C = 27$

Los resultados así como algunas iteraciones intermedias del proceso de optimización se muestran en la Tabla 3.1.

Iteración	G1		G2		G3		G4		σ_P [ps]	A_P [$W_{min}L_{min}$]
	K	M	K	M	K	M	K	M		
1	1.0	-3.5	1.0	-3.2	1.0	-13	1.0	-22	9.75	12.0
2	1.0	-3.5	1.0	-3.2	1.0	-13	1.2	-14	8.33	12.6
3	1.0	-3.5	1.0	-3.2	1.0	-14	1.4	-8.0	7.56	13.2
.
25	1.6	-0.9	1.6	-1.0	2.4	-0.9	3.2	-0.9	3.56	26.4
26	1.6	-1.0	1.8	-0.7	2.4	-1.1	3.2	-1.0	3.53	27.0

Tabla 3.1: Proceso de optimización iterativo

Gratificando algunos resultados para tener una mejor visión del comportamiento.

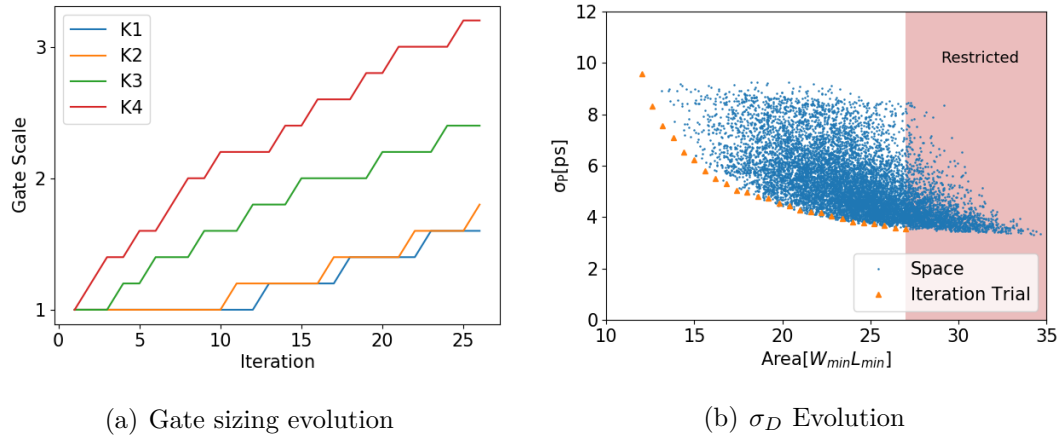


Figura 3.7: Resultados de optimización base

Claramente por observación se identifica que el valor encontrado es el mínimo posible que puede conseguirse de la función σ_P , respetando la restricción dada.

Para validar que los resultados anteriormente obtenidos son correctos se proceden a utilizar la tecnología de 65nm con la cual se puede realizar nuevamente la optimización pero con los modelo completos de transistores (a un alto costo computacional) con el fin de comparar resultados.

Utilizando el software SPICE se encuentran los tamaños óptimos que minimizan el valor del σ_P para el camino evaluado, la comparación de estos resultados con el modelo analítico se muestra en la Tabla 3.2

Parámetro	Modelo Analítico	SPICE
K1	1.6	1.7
K2	1.8	1.6
K3	2.4	2.7
K4	3.2	3.0
σ_P [ps]		
Valor real en tecnología	2.70	2.65

Tabla 3.2: Comparación de valor optimo bajo restricción de área sin considerar correlación espacial

A continuación se calcula el error entre los valores obtenidos del σ_P por ambos modelos.

$$\text{Error relativo del } \sigma_D = 1,85 \% \quad (3.14)$$

A un cuando el modelo analítico predecía un valor de σ_P diferente que el valor obtenido por SPICE para las mismas dimensiones, el comportamiento de ambos modelos siguen la misma tendencia, por lo cual aun al tener un error en la magnitud del σ_P predicho el minimizar con base al modelo genera una buena reducción en el valor del σ_P real.

Optimización bajo tamaños discretizados de las compuertas lógicas sin considerar correlación espacial

Muchas veces los tamaños requeridos para implementar el circuito óptimo no serán realizables debido a que generalmente un diseño digital se sintetiza por medio de un conjunto de bloques de construcción llamados celdas estándar los cuales son implementaciones físicas de las compuertas lógicas (o incluso bloques digitales más complejos) de dimensiones fijas [2].

La tecnología de 65nm utilizada tiene un granulado de celdas para las compuertas básicas tal que se abarcan las dimensiones X1, X2, X3, X4, X8, X12, X16 y X20 en esta nomenclatura el numero que sigue a la letra X representa el factor de escala de la compuerta, claramente al tener valores limitados en las dimensiones de las compuertas implica que para poder implementar un circuito optimizado se tendrá en muchos casos que redondear los valores obtenidos para usar las celdas más cercanas. Para caminos críticos el diseño Full-Custom permite tener un control total sobre el tamaño directo de los transistores en el camino.

Para acercar mas la metodología anterior a un diseño basado en celdas estándar se puede utilizar un paso de iteración ΔK como un numero entero, repitiendo el ejemplo anterior bajo esta consideración se obtienen los siguientes resultados.

Iteración	G1		G2		G3		G4		σ_P [ps]	A_P [$W_{min}L_{min}$]
	K	M	K	M	K	M	K	M		
1	1.0	-3.5	1.0	-3.2	1.0	-13	1.0	-22	9.57	12.0
2	1.0	-3.5	1.0	-3.2	1.0	-14	2.0	-2.3	6.49	15.0
3	1.0	-3.5	1.0	-4.2	2.0	-1.3	2.0	-4.6	5.07	18
4	1.0	-3.5	1.0	-4.2	2.0	-1.6	3.0	-1.5	4.22	21
5	1.0	-6.2	2.0	-0.3	2.0	-3.5	3.0	-2.1	4.04	24
6	2.0	-0.5	2.0	-0.4	2.0	-1.6	3.0	-0.8	3.76	27

Tabla 3.3: Proceso de optimización iterativo con paso discreto

Claramente los valores de escala devueltos son directamente sintetizables por medio de las celdas estándar discutidas anteriormente.

Para comparar los resultados obtenidos entre la versión base y discreta de la optimización se muestra la Tabla 3.4 y la Figura 3.8.

Parámetro	Caso base	Caso discreto
K1	1.6	2.0
K2	1.8	2.0
K3	2.4	2.0
K4	3.2	3.0
σ_P [ps]	3.53	3.76
# Iteraciones	26	6

Tabla 3.4: Comparación entre la optimización base y discreta

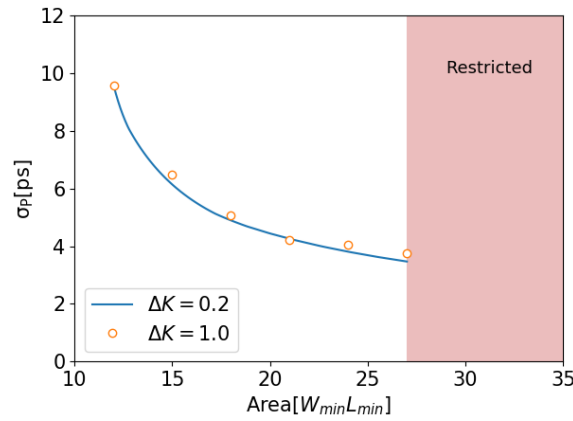


Figura 3.8: Comparación del σ_D entre la optimización base y discreta

Optimización con correlación espacial y aumento continuo en el tamaño de las compuertas lógicas

El objetivo de introducir la correlación espacial es únicamente mejorar la precisión del modelo analítico por lo que esta no modifica la metodología de optimización discutida anteriormente solo aumenta la complejidad de los cálculos numéricos ya que el último termino en las ecuaciones 2.107 y 3.7 es generalmente diferente de cero.

Para incluir la correlación espacial en el análisis es requerido primeramente una matriz cuadrada que especifique el nivel de correlación entre cada par de compuertas, los coeficientes de dicha matriz pueden ser calculados por medio del modelo exponencial una vez conocidas las distancias de separación entre las compuertas lógicas implementadas dentro del Layout [5], sin embargo para los

cálculos de esta sección se trabajara directamente con una matriz de correlación dada, misma que se muestra a continuación.

$$\rho = \begin{bmatrix} 1,00 & 0,82 & 0,75 & 0,62 \\ 0,82 & 1,00 & 0,88 & 0,79 \\ 0,75 & 0,88 & 1,00 & 0,95 \\ 0,62 & 0,79 & 0,95 & 1,00 \end{bmatrix} \quad (3.15)$$

Como se menciona anteriormente la correlación espacial no modifica la metodología de optimización a seguir sin embargo si modifica el espacio de optimización del camino lógico como es mostrado en la Figura 3.9.

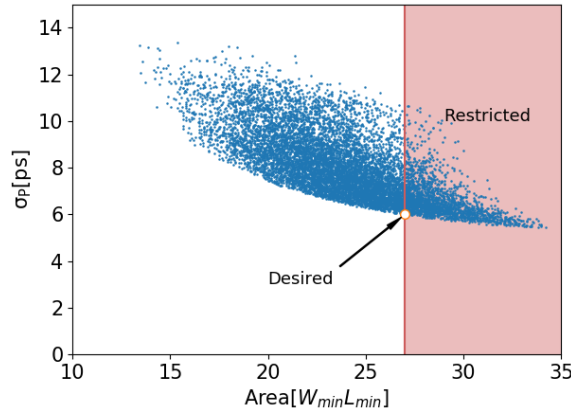


Figura 3.9: Espacio de optimización con correlación espacial

Claramente es observado que la correlación espacial incrementa el valor del σ_P , esto debido a los términos extras incluidos en la ecuación 2.107 que moldean dicha correlación. Para llevar a cabo la optimización se utilizara nuevamente la restricción de área especificada en los ejemplos anteriores ($A_C = 27$), por su parte el paso de iteración escogido será $\Delta K = 0,2$, usando estos parámetros los resultados obtenidos una vez llevada a cabo la optimización se muestran a continuación.

Iteración	G1		G2		G3		G4		σ_P [ps]	A_P [$W_{min}L_{min}$]
	K	M	K	M	K	M	K	M		
1	1.0	-21	1.0	-30	1.0	-85	1.0	-63	13.9	12.0
2	1.0	-19	1.0	-29	1.2	-55	1.0	-58	13.3	12.6
3	1.0	-17	1.0	-26	1.2	-48	1.2	-38	11.8	13.2
.
25	1.6	-3.8	1.8	-4.7	2.8	-4.0	2.6	-4.4	6.22	26.4
26	1.6	-4.1	2.0	-3.8	2.8	-4.1	2.6	-4.4	6.17	27.0

Tabla 3.5: Proceso de optimización iterativo con correlación espacial

Graficando algunos resultados.

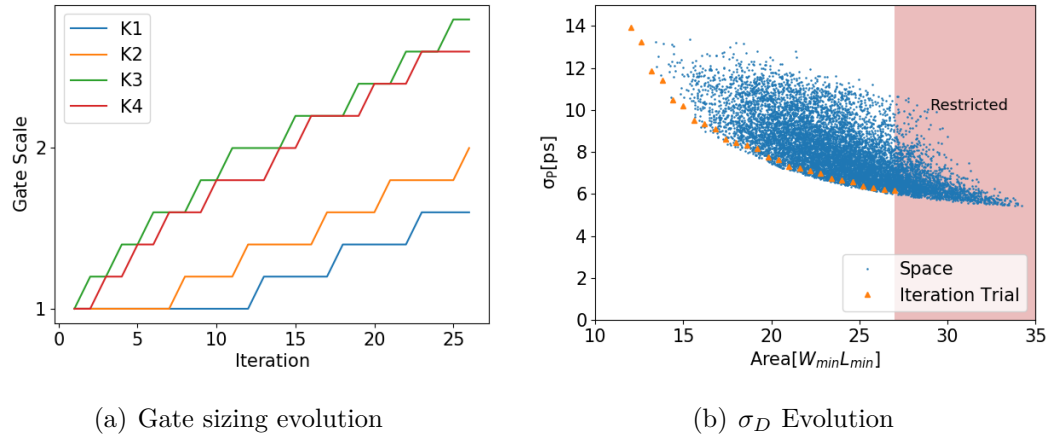


Figura 3.10: Resultados de optimización base

Comparando los resultados con la optimización base se forma la siguiente Tabla.

Claramente la correlación espacial modifica los tamaños óptimos conseguidos ya que el impacto que tienen cada compuerta en el σ_P es mayor para una escala dada, esto se ve reflejado en las métricas de cada compuerta como es visto en las iteraciones mostradas por la Tabla 3.10 en la cual las compuertas aumentaron considerablemente su métrica comparado con el caso base, por su parte el valor mínimo posible de conseguir dada la restricción también aumento en un 56 %.

Nuevamente se usa el simulador eléctrico SPICE para realizar la comparación entre los resultados anteriores y el verdadero punto óptimo, cabe destacar que por defecto SPICE no soporta la inclusión de correlación pero estas puede ser generadas a partir de un conjunto de variables independientes por medio del método Cholesky (Apendice A.7).

Parámetro	Caso base	Caso correlacionado
K1	1.6	1.6
K2	1.8	2.0
K3	2.4	2.8
K4	3.2	2.6

Tabla 3.6: Comparación entre la optimización base y con correlación espacial

Los resultados de la comparación de resultados entre el modelo analítico y SPICE auxiliado por la transformación de Cholesky se muestran en la Tabla 3.7.

Parámetro	Modelo Analítico	SPICE
K1	1.6	1.6
K2	2.0	1.7
K3	2.8	2.6
K4	2.6	3.0
σ_P [ps] Valor real en tecnología	4.25	4.20

Tabla 3.7: Comparación de valor optimo bajo restricción de área considerando correlación espacial

Calculando el error de esta segunda evaluación.

$$\text{Error relativo del } \sigma_D = 1,18\% \quad (3.16)$$

3.3.2. Optimización de un camino lógico compuesto por compuertas de diferente tipo

El camino lógico que se analizará en esta sección es mostrado en la Figura 3.11.

En el cual se considera.

$$C_1 = 1[\text{fF}] \quad (3.17)$$

$$C_2 = 3[\text{fF}] \quad (3.18)$$

$$C_3 = 5[\text{fF}] \quad (3.19)$$

$$C_4 = 9[\text{fF}] \quad (3.20)$$

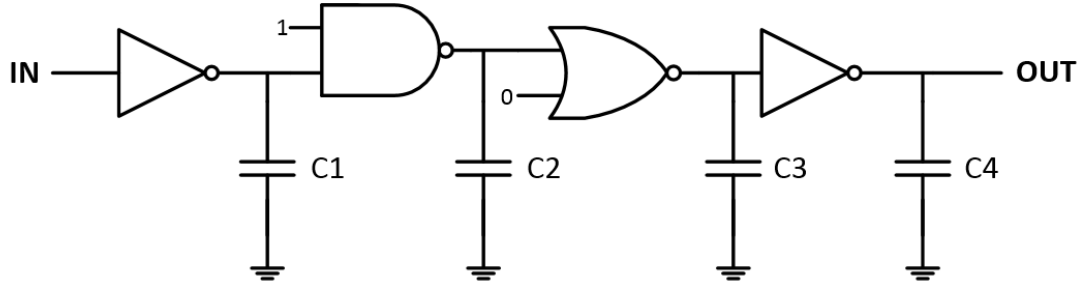


Figura 3.11: Circuito bajo prueba

En éste análisis se tomara en cuenta la correlación espacial entre compuertas, la cual se define por la siguiente matriz.

$$\rho = \begin{bmatrix} 1,00 & 0,81 & 0,84 & 0,69 \\ 0,81 & 1,00 & 0,96 & 0,84 \\ 0,84 & 0,96 & 1,00 & 0,81 \\ 0,69 & 0,84 & 0,81 & 1,00 \end{bmatrix} \quad (3.21)$$

Si el nodo de entrada conmuta de la forma $0 \rightarrow 1$, a partir de la Tabla 2.1 se puede encontrar que:

$$FD_1 = 2 \quad (3.22)$$

$$FD_2 = 2 \quad (3.23)$$

$$FD_3 = 4 \quad (3.24)$$

$$FD_4 = 1 \quad (3.25)$$

Aplicar la metodología de optimización estudiada anteriormente a un camino compuesto por compuertas de diferente tipo no es del todo correcto, ya que aún cuando en dicha metodología se está midiendo el impacto en el σ_D ésta no llega a considerar que crecer una compuerta puede tener un impacto en el área del camino diferente según el tipo de compuerta, ya que por ejemplo escalar un inversor en un factor ΔK tiene un costo de área diferente que incrementar una compuerta NAND3 en la misma cantidad ΔK esto se puede apreciar en la ecuación 3.9 al sustituir los factores diseño de cada compuerta (2.1).

Para tomar las consideraciones anteriores se plantea una métrica general que pueda tomar en cuenta combinaciones mixtas de compuertas, dicha métrica es

una combinación de la métrica de la metodología anterior junto al impacto de área discutido, la forma matemática de dicha métrica se muestra a continuación.

$$M_i = \frac{\partial \sigma_P^2}{\partial K_i} / \frac{\partial A_P}{\partial K_i} \quad (3.26)$$

Claramente el primer factor corresponde con la métrica de la metodología utilizada en los análisis anteriores y la cual es expresada en 3.6, por su parte el segundo término puede calcularse de manera directa utilizando 3.12, por lo cual:

$$\frac{\partial A_P}{\partial K_i} = \frac{\partial \sum_{i=1}^x A_i}{\partial K_i} = \frac{\partial A_i}{\partial K} \quad (3.27)$$

Utilizando 3.9 es directo encontrar que:

$$\frac{\partial A_P}{\partial K_i} = (\#nMOS_i \times FD_{n_i} + \#pMOS_i \times FD_{p_i})[W_{min}L_{min}] \quad (3.28)$$

Antes de iniciar la optimización se muestra el espacio de trabajo, el cual se genera para $A_C = 40$.

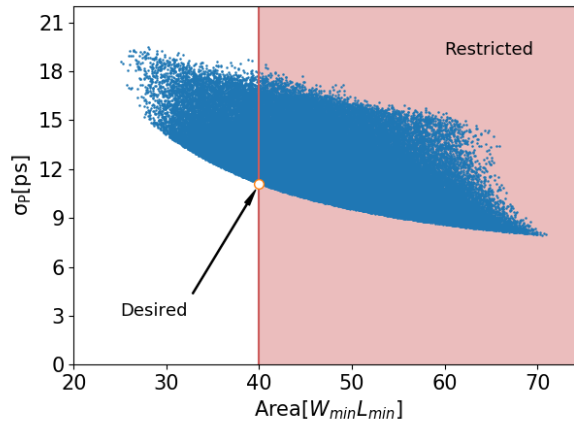


Figura 3.12: Espacio de optimización considerado

Llevando acabo la optimización con la métrica definida en 3.26 son obtenidos los siguientes resultados.

Iteración	G1		G2		G3		G4		σ_P [ps]	A_P [$W_{min}L_{min}$]
	K	M	K	M	K	M	K	M		
1	1.0	-2.1	1.0	-2.7	1.0	-3.7	1.0	-19	20.52	24.0
2	1.0	-1.9	1.0	-2.5	1.0	-3.5	1.2	-10	18.61	24.6
.
14	1.2	-0.8	1.4	-0.8	1.6	-0.9	2.4	-1.0	11.77	38.0
15	1.2	-0.8	1.4	-0.8	1.6	-0.9	2.6	-0.7	11.53	39.6

Tabla 3.8: Proceso de optimización del circuito bajo prueba

Graficando algunos resultados

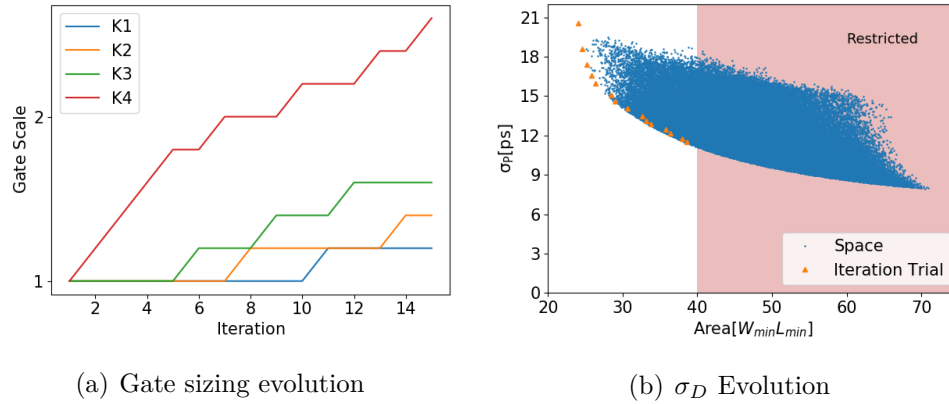


Figura 3.13: Resultados de optimización para el circuito bajo prueba

Comparando el resultado obtenido en una tecnología real contra el circuito óptimo que puede proporcionar dicha tecnología se encuentra que:

Computando el error entre el valor de σ_D predicho por el modelo analítico y su valor real en la tecnología.

$$\text{Error relativo del } \sigma_D = \frac{8,07 - 7,86}{8,07} = 2,60\% \quad (3.29)$$

Para finalizar se muestra el histograma generado si se implementa el diseño del modelo analítico contra el circuito optimo de la tecnología.

Parámetro	Modelo Analítico	Tecnología
K1	1.2	1.38
K2	1.4	1.50
K3	1.6	1.45
K4	2.6	3.18
σ_P [ps] Valor real en tecnología	8.07	7.86

Tabla 3.9: Comparación de valor óptimo para circuito de prueba

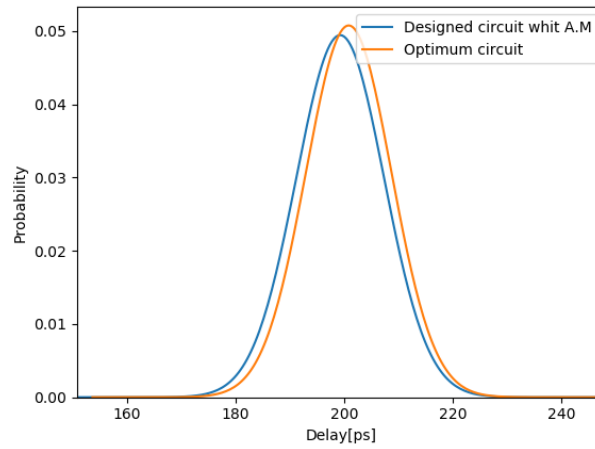


Figura 3.14: Comparación entre los dos circuitos especificados en la Tabla 3.9

3.4. Optimización sin aumentar la capacitancia de entrada del camino lógico

Un tipo importante de optimización es aquella que mantiene el valor de la capacitancia de entrada de la primera compuerta en un valor definido, esto es así por que dicho valor también es la capacitancia de entrada del camino, la cual sirve para conocer el efecto de acoplar directamente diferentes etapas lógicas [1].

Para este analisis se usara como referencia el camino lógico mostrado en la Figura 3.4.

En esta optimización el consumo de área no es una variable de interés primario ya que la forma del σ_P se vuelve convexa con respecto al área consumida, lo cual quiere decir que aunque se disponga de mas área a partir de un cierto punto no

se puede mejorar el desempeño del camino, este efecto es plasmado gráficamente en la Figura 3.15.

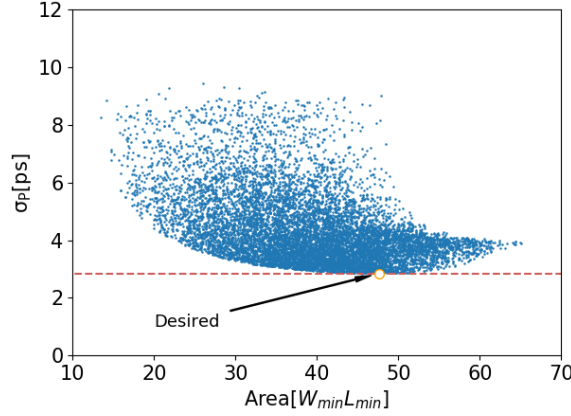
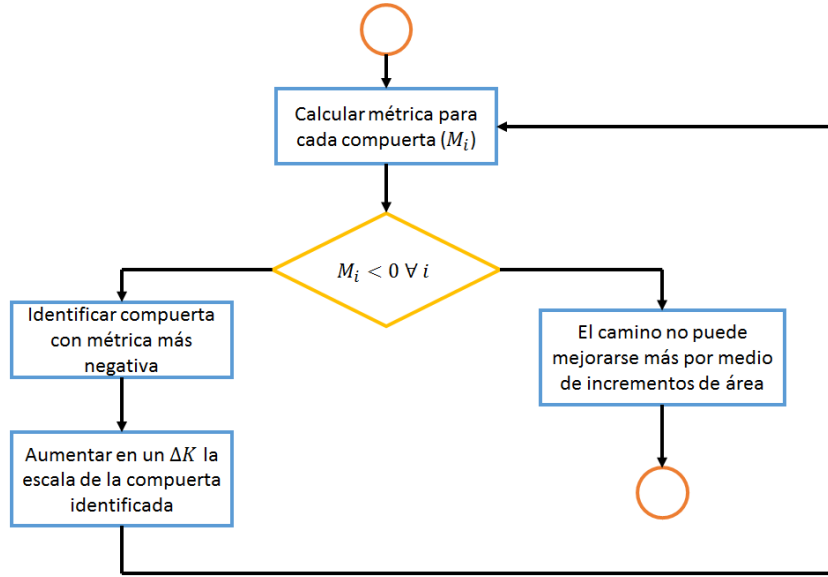


Figura 3.15: Espacio de optimización conservando el valor de C_{in} de la primera compuerta

Cabe mencionar que en el desarrollo de la Figura 3.15 no se considero ningún tipo de correlación espacial.

El por que de este efecto convexo es fácil de explicar conociendo que el aumentar la escala de una compuerta también aumenta su capacitancia de entrada (2.9), como la carga de una compuerta en un camino lógico es siempre dependiente de la capacitancia de entrada de la compuerta posterior en el camino implica que el aumentar el tamaño de cualquier compuerta diferente a la primera siempre cargara a su respectiva compuerta anterior aumentando de manera proporcional al valor del σ_D (2.95) de dicha compuerta, naturalmente este aumento se puede relajar si se dimensiona la compuerta afectada en la proporción adecuada, sin embargo esto genera una reacción en cadena ya que la carga se transmite a la compuerta inmediata anterior repitiendo el ciclo hasta llegar a la primera compuerta la cual esta fijada bajo este tipo de optimización haciendo imposible la eliminación de los efectos de carga transmitidos, cuando dichos efectos superen al factor de reducción dado por cualquier dimensionamiento de compuertas utilizar mas área solo tendrá efectos perjudiciales en el valor del σ_P encontrado en el camino lógico, utilizando el calculo diferencial es fácil demostrar que cuando el proceso de dimensionar una compuerta perjudica el valor del σ_P dicha compuerta tendrá una métrica positiva, por lo cual es posible plantar la metodología de optimización por medio del diagrama de flujo mostrado en la Figura 3.16.


 Figura 3.16: Diagrama de flujo para la optimización conservando el valor de C_{in} de la primera compuerta

Cabe destacar que el criterio de paro por la métrica puede ser engañoso debido a que un valor absoluto como cero no tiene una cabida en un algoritmo numerico como el desarrollado, esto debido a que el algoritmo se mueve en la curva por pasos finitos (Ver Figura 3.17) por lo que es muy improbable que el algoritmo pase por exactamente el punto cero en la evaluación de la métrica, en estos caso es mas conveniente establecer un rango de convergencia. Los resultados así como algunas iteraciones intermedias obtenidas al aplicar el proceso de optimización se muestran en la siguiente Tabla, en dichos cálculos no se considero la correlación espacial y se tomo como paso de iteración el factor $\Delta K = 0,2$.

Iteración	G1		G2		G3		G4		σ_P [ps]	A_P [$W_{min}L_{min}$]
	K	M	K	M	K	M	K	M		
1	1.0	NA	1.0	-3.2	1.0	-13	1.0	-22	9.57	12.0
2	1.0	NA	1.0	-3.2	1.0	-13	1.2	-13	8.33	12.6
.
40	1.0	NA	2.0	-0.7	3.8	-0.7	5.0	-0.8	3.06	35.4

 Tabla 3.10: Proceso de optimización conservando el valor de C_{in}

Graficando algunos resultados de la Tabla 3.10.

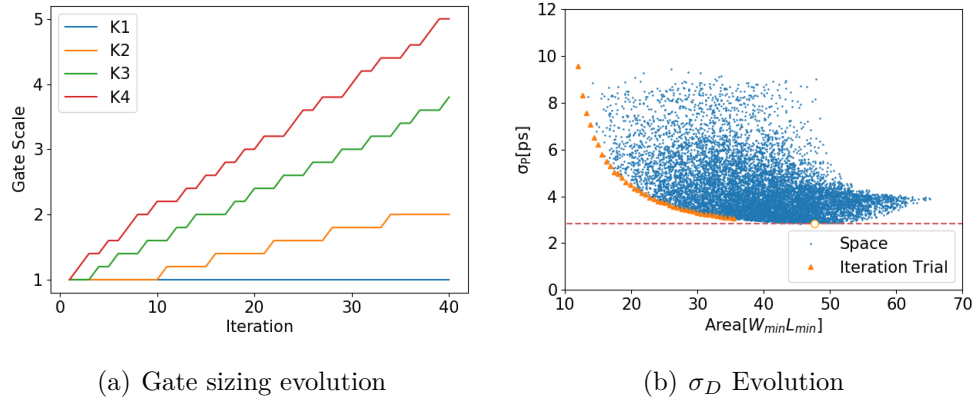


Figura 3.17: Resultados de optimización conservando el valor de C_{in}

Validando los resultados anteriores por medio de SPICE se obtiene la Tabla 3.11.

Parámetro	Modelo Analítico	SPICE
K1	1.0	1.0
K2	2.0	2.0
K3	3.8	7.0
K4	5.0	10
σ_P [ps]		
Valor real en tecnología	2.74	2.52

Tabla 3.11: Comparación de valor optimo conservando el valor de C_{in}

Calculando el error

$$\text{Error relativo del } \sigma_D = 8,03\% \quad (3.30)$$

Como se menciona anteriormente es de destacar de la Figura 3.17 que en este tipo de optimización el criterio de paro es difícil de definir ya que la pendiente de las curva óptima es pequeña en las fronteras del valor mínimo, tal que la convergencia final depende exactamente del valor o rango en el cual dicha pendiente se considere despreciable, para ejemplificar esto ultimo se repite la simulación pero indicando un numero mas cercano a cero como criterio de paro.

Si se repite la tabla de comparación en SPICE para este segundo modelo se encuentra la Tabla 3.12.

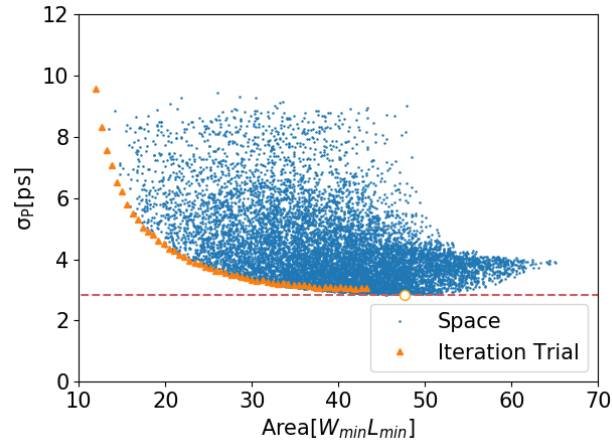


Figura 3.18: Repetición de la optimización anterior con un criterio de paro mas fino

Parámetro	Modelo Analítico	SPICE
K1	1.0	1.0
K2	2.9	2.4
K3	7.1	7.0
K4	10	10
σ_P [ps]	2.54	2.52
Valor real en tecnología		

Tabla 3.12: Repetición de la comparación de valor optimo conservando el valor de C_{in} con un criterio de para mas fino

Calculando el error

$$\text{Error relativo del } \sigma_D = 0,7\% \quad (3.31)$$

Claramente al afinar el criterio de paro se consiguen resultados óptimos, sin embargo debe tenerse cuidado con el valor de ΔK utilizado ya que al haber cambios tan pequeños en la pendiente si el paso es muy grande se puede brincar el punto en el cual se alcanza el criterio de para llevando al sistema iterativo a una no convergencia, esta situación por el contrario no sucede cuando se tiene un paro por restricción área ya que el numero de iteraciones siempre sera definido, en general se recomienda siempre tener un criterio de para basado en área para no entrar en ciclos infinitos.

Aun cuando el criterio de paro original era menos estricto los resultados obtenidos

son mas que aceptables ya que el cambio en el σ_P es poco y el área utilizada es mucho menor.

3.5. Conclusiones

En éste capítulo también se mostró una aplicación directa del modelo analítico desarrollado en el capítulo 2, la optimización enfocada al proceso de dimensionamiento de las compuertas de un camino lógico se estudio tanto de forma cuantitativa como cualitativa las cantidades de las derivadas respectivas de la desviación estándar del retardo con respecto al factor de escala K de una compuerta lógica, esto tanto a nivel compuerta como camino lógico estudiar las derivadas es importante por que permiten definir el cambio producido en una función bajo pequeños cambios en sus entradas, si se aplica lo anterior en el contexto de caminos logicos claramente las derivadas definidas permiten la comparación directa de la mejora en la desviación estándar del retardo por el proceso de dimensionar una compuerta logica por lo que la cantidad puede usarse como un indicador para cuantificar que compuertas son mas eficiente de dimensionar tal que debido a esto se denomino a la derivada del camino con respecto a la compuerta i como la métrica de la compuerta i . Posteriormente se uso esta definición de la métrica en diferentes ejemplos de optimización, la exactitud de la metodología fue validada ampliamente con respecto a SPICE.

Capítulo 4

Guías prácticas de diseño para optimización de un camino lógico

4.1. Comportamiento del retardo con respecto a los parámetros de diseño

En esta sección el valor del σ_D de las compuertas lógicas es analizado con respecto a los principales parámetros de diseño, los parámetros considerados son: 1) La escala de la compuerta lógica (K), 2) La capacitancia de carga (C_L) y 3) El apilamiento de transistores en serie dentro de la compuerta (Stack).

4.1.1. Factor de escala de la compuerta

Las ecuaciones 2.87 y 2.95 muestran que tanto el valor nominal del retardo como su desviación estándar dependen fuertemente del factor de escala (K), el cual es un parámetro directamente controlable al diseñar una topología lógica. También la Figura 4.1 muestra que la desviación estándar del retardo se reduce conforme se aumenta la escala de la compuerta lógica, claramente dos regiones de operación se pueden distinguir en la figura anterior. En la primera región (lado izquierdo) la desviación estándar se reduce de manera significativa con respecto a un aumento en el factor de escala, mientras que en la segunda región (lado derecho) la desviación estándar presenta solo una pequeña reducción con respecto al factor de escala.

Con base al párrafo anterior la reducción del retardo estadístico se puede combatir redimensionando adecuadamente las compuertas lógicas, esta estrategia es conocida como Gate-sizing [5].

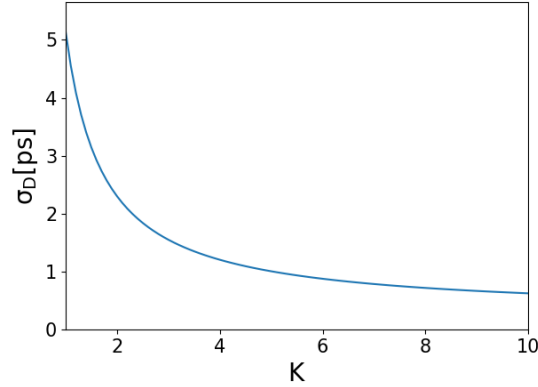


Figura 4.1: Comparación del σ_D con respecto a la escala de la compuerta, en un inversor con $C_L = 4$ [fF]

4.1.2. Capacitancia de carga

Las ecuaciones 2.87 y 2.95 muestran que tanto el retardo nominal como la desviación estándar son directamente proporcionales a la capacitancia de carga, esto se puede visualizar gráficamente en la Figura 4.2, la cual se aproxima a una función lineal.

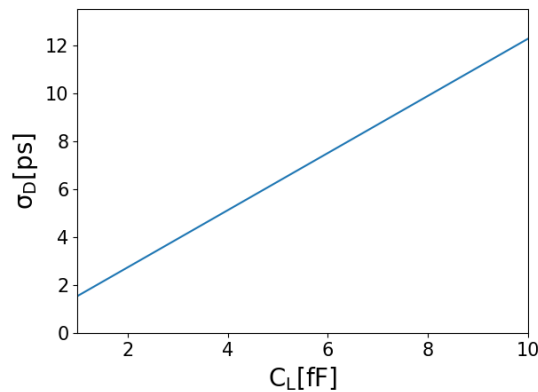


Figura 4.2: Comparación del σ_D con respecto a la carga de la compuerta, en un inversor con $K = 1$

Por lo tanto en un circuito digital se deberá tener especial cuidado en el diseño de las compuertas que dirijan en sus salidas nodos con un gran valor capacitivo.

4.1.3. El Stack

Conforme se aumenta el Stack de transistores para computar una operación lógica se aumenta el número de nodos intermedios y el FD de cada transistor en la compuerta. Por lo que la contribución en la variación dada por los nodos intermedios será mayor, de igual forma por la situación discutida también incrementará el consumo de área de la compuerta lógica.

Aparte de esto el Stack también incrementa de forma directamente proporcional a la contribución de las variaciones independientes en la desviación estándar y de una forma más compleja y significativa (ver ecuación 2.103) en las contribuciones de las variaciones correlacionadas.

Por lo tanto es recomendable no utilizar compuertas con un Stack excesivo para implementar funciones lógicas.

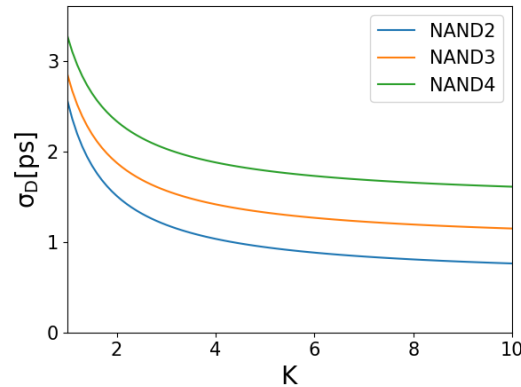


Figura 4.3: Comparación del σ_D con respecto al Stack con $C_L = 2$ [[fF]]

4.2. Guías de diseño

Generalmente un diseñador de IC está interesado en identificar las compuertas que al aumentar su escala generen la mayor reducción total al valor de la desviación estándar del retardo en un camino lógico.

La métrica mas básica que se puede discutir esta basada en la desviación estándar del retardo de una compuerta ($\partial\sigma_D/\partial K$) dicha ecuación puede ser computada por la ecuación 3.1 su gráfica se muestra en la Figura 4.4. Visualizando la gráfica se observa que los valores de la ecuación 3.1 son mayores cuando la relación C_L/K es grande, en otras palabras para el mismo incremento de escala y carga fija una compuerta con $K = 2$ genera una reducción del σ_D mayor que una compuerta dimensionada con tamaño $K = 4$. Utilizando la relación marcada por las variables independientes de la ecuación 3.1 se puede plantear como métrica la ecuación 4.6.

$$M_i = \frac{C_{L_i}}{K_i^{a_1}} \quad (4.1)$$

Donde a_1 es una constante empírica que depende del tipo de variaciones consideradas y toma el valor 2.5 en la tecnología de 65nm cuando se consideran variaciones simultaneas de W , L , T_{ox} y V_T , los detalles de como calcular dicho coeficiente se muestran en el Apéndice B.1.

Claramente cuando una compuerta tiene un puntaje alto en su métrica esta reduce en mayor medida el valor de su σ_D en comparación a una compuerta con un puntaje inferior. Del análisis anterior es claro que la métrica tiende a presentar valores más altos para compuertas más pequeñas. Por otra parte la métrica tiende a presentar valores inferiores para una compuerta más grande y en estos casos aumentar la escala de la compuerta no provee beneficios reales de reducción en el σ_D .

En un camino lógico cada compuerta puede tener diferentes factores de escala y cargas, lo cual implica diferentes puntajes y claramente en base a lo discutido aquellas compuertas con mayor puntaje son las que generan la mayor reducción en su contribución individual al σ_P .

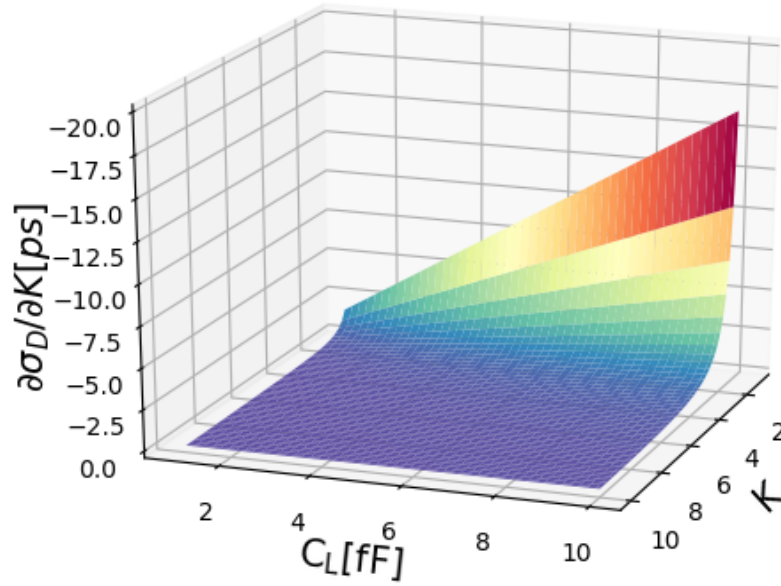


Figura 4.4: Valores de la derivada de la desviación estándar del retardo para una compuerta lógica

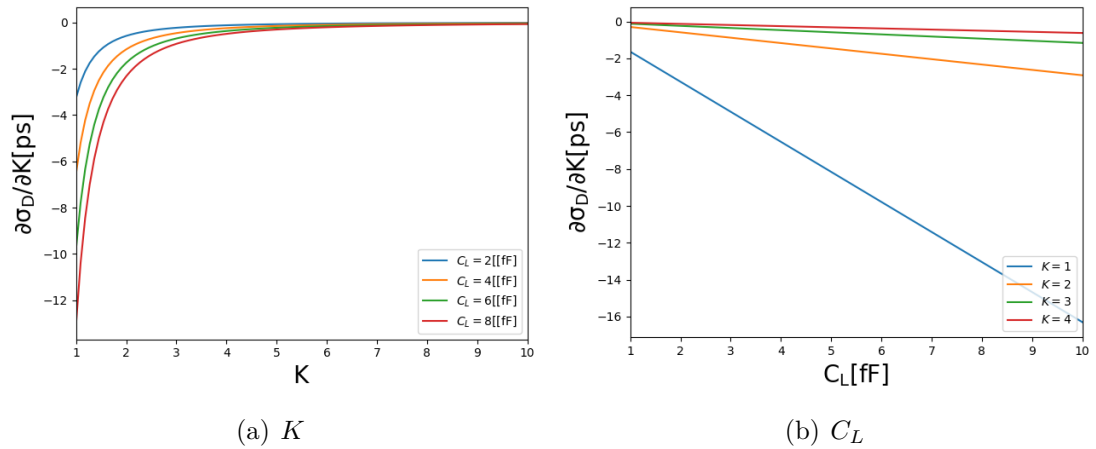


Figura 4.5: Vista en cortes de la Figura 4.4

Para comparar que la métrica dada en la ecuación 4.6 sigue la tendencia de la ecuación 3.1 se muestra la Figura 4.6 en la cual se compara el valor normalizado de

la derivada con respecto a la métrica discutida con el fin de visualizar la tendencia de ambas curvas.

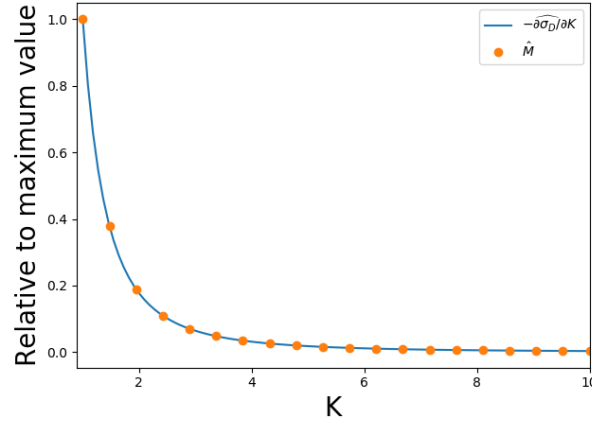


Figura 4.6: Comparación entre la tendencia de la métrica y la derivada completa, con $a = 2,5$

La nomenclatura \hat{X} hace referencia a una variable normalizada con respecto a su máximo. Cabe destacar que el valor normalizado es independiente de C_L ya que este afecta de forma lineal tanto a la derivada como a la métrica.

Al usar la ecuación 4.6 como métrica se están asumiendo las siguientes situaciones.

1. El camino lógico se compone del mismo tipo de compuertas.
2. Se ignora las interacciones de las compuertas en el camino
3. No se considera la correlación espacial.

4.2.1. Métrica para diferentes tipos de compuerta

Cuando se consideran diferentes tipos de compuertas no se debe ponderar solo la reducción en el σ_D individual si no también el incremento producido en el área de la compuerta, esto debido a que diferentes topologías tienen un consumo diferente de área para un mismo aumento de K [11], en general es directo observar que para una compuerta lógica su incremento en área por un aumento en su factor de escala se encuentra al derivar la expresión 3.9.

$$\frac{\partial A_D}{\partial K} = (\#nMOS_i \times FD_{n_i} + \#pMOS_i \times FD_{p_i})[W_{min}L_{min}] \quad (4.2)$$

Tal que siguiendo la relación marcada en 4.6 y modificando adecuadamente se puede generar una métrica que compare el diferente ΔA_D generado por el factor de escala en cada tipo de compuerta, en base a lo anterior se puede plantear una nueva métrica que pondere esta consideración.

$$M_i = \frac{C_{L_i}}{K_i^{a_1}} \times \frac{1}{\#nMOS_i \times FD_{n_i} + \#pMOS_i \times FD_{p_i}} \quad (4.3)$$

4.2.2. Métrica considerando efectos adyacentes

La cadena de inversores mostrada en la Figura 4.7 es utilizada para ejemplificar el impacto de aumentar el factor de escala de una compuerta dentro de un camino lógico, como se estudio anteriormente conforme se aumenta las dimensiones de una compuerta lógica se disminuye tanto el retardo nominal como la desviación estándar. Sin embargo la contribución de una compuerta en la desviación estándar total del camino no solo se genera por su propia contribución individual si no también por las interacciones con las otras compuertas dentro del camino.

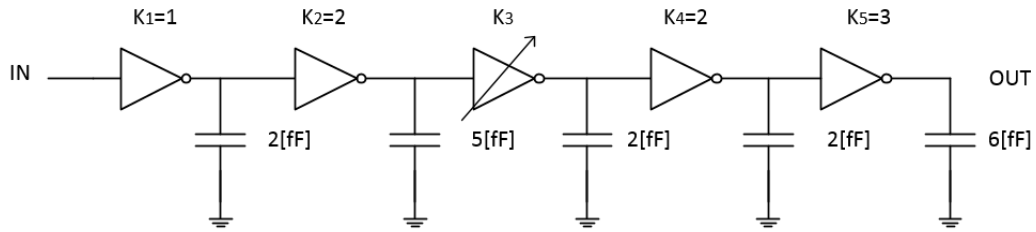


Figura 4.7: Cadena de inversores bajo redimensionamiento

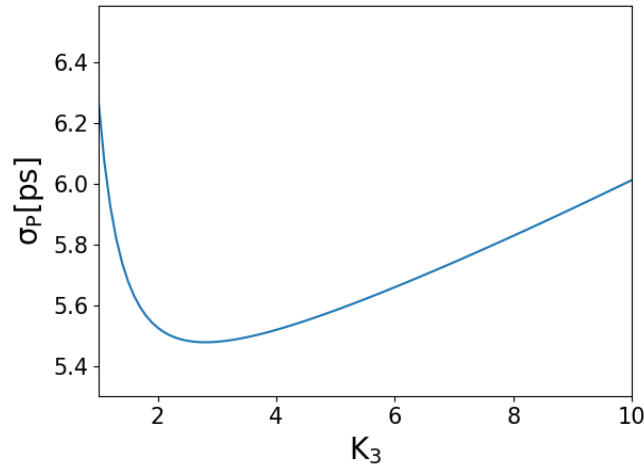


Figura 4.8: Comportamiento del camino lógico anterior bajo un dimensionamiento indiscriminado de K_3

Al aumentar la escala de una compuerta se aumenta su capacitancia de entrada (ver ecuación 2.9), en un camino lógico la capacitancia de entrada de una compuerta es generalmente la carga de la compuerta predecesora. Por lo que incrementar el K la compuerta i generalmente aumenta la desviación estándar de la compuerta $i - 1$ y debido a que el valor de σ_{D_i} tiende a saturarse en su mejora con respecto a un incremento indefinido de K_i en algún punto el incremento del $\sigma_{D_{i-1}}$ de la etapa anterior producido por el aumento de C_{in_i} llegara a ser dominante y el aumentar K_i traerá consigo resultados adversos al σ_P del camino lógico, este comportamiento se ejemplifica en la Figura 4.8.

De tal forma que para adecuar el valor del σ_P a un valor optimo no basta con examinar la contribución aislada de cada compuerta sino también sus interacciones, para incluir dichos efectos en la métrica se analiza la ecuación 3.8 de tal manera que se propone como métrica.

$$M_i = \frac{C_{L_i}}{K_i^{a_1}} - \frac{C_{ox} W_{min} L_{min} \times (FD_{n_i} + FD_{p_i})}{K_{i-1}^{a_1}} \quad (4.4)$$

4.2.3. Métrica con inclusión de covarianza

La covarianza generada por dos compuertas lógicas es un función de: 1) el coeficiente de correlación ρ , 2) el factor de escala K y 3) la carga de la compuerta C_L .

La correlación espacial entre dos compuertas decrece exponencialmente con la distancia como se ve en la ecuación 2.106, por lo cual para una compuerta lógica el calculo de la covarianza se puede truncar considerando solo a la periferia de la compuerta.

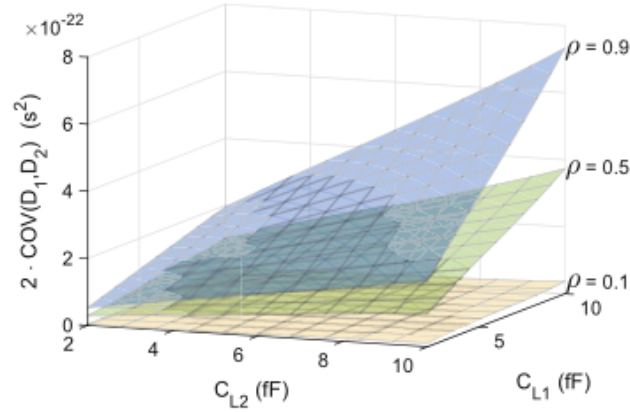


Figura 4.9: Covarianza y coeficiente de correlación para dos compuertas inversoras [5]

Examinando las similitudes entre las ecuaciones 2.95 y 2.107 se puede ver que para dos compuertas cercanas ($\rho \approx 1$) su covarianza se puede ver como la contribución de una compuerta con dimensiones $K_{eff} = \sqrt{K_1 K_2}$ y carga $C_{eff} = \sqrt{C_{L1} C_{L2}}$ que no pose variaciones independientes ($\delta_{V_T} = 0$). Tal que la expresión bajo estos cambios se muestra a continuación.

$$\text{COV}(D_i, D_j) = \rho \frac{\kappa^2 C_{eff}^2}{\beta_i \beta_j} \times \sqrt{\Upsilon_i \Upsilon_j} \times \left(\frac{\delta_W}{(K_{eff}^2 F D_{eff}^2)^2} + \frac{\delta_L}{K_{eff}^2 F D_{eff}^2} + \frac{\delta_{T_{ox}}}{K_{eff}^2 F D_{eff}^2} \right) \quad (4.5)$$

Por lo cual una primera aproximación de diseño para considerar las covarianza puede realizarse examinando la periferia de las compuertas mas criticas de un camino y calcular la covarianza por medio del modelo de la compuerta efectiva.

Si bien la contribución aislada de cada compuerta lógica es generalmente mas grande que la covarianza entre dicha compuerta y cualquier otra compuerta en el camino, el numero de términos totales necesarios para calcular la covarianza de un camino lógico crece de forma acelerada conforme al numero de compuertas por lo cual la covarianza no puede ser negada cuando existe una alta profundidad lógica en el camino.

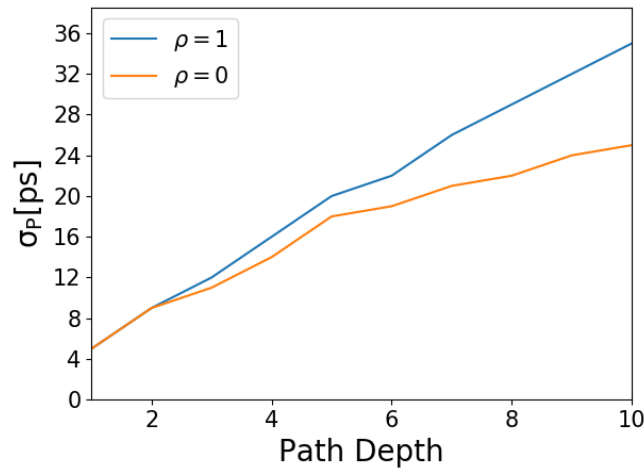


Figura 4.10: Impacto de la profundidad en el σ_P bajo condición de covarianza

Ajustando la métrica para incluir los efectos discutidos de la covarianza se plantea la siguiente ecuación.

$$M_i = \frac{C_{L_i}}{K_i^{a_1}} + \sum_{m=1}^n \rho_m \frac{C_{eff}}{K_{eff}^{a_2}} \quad (4.6)$$

Debido a que en el termino de covarianza solo existen variaciones correlacionadas el coeficiente a_2 deberá ser caracterizado para considerar solo a dichos términos, lo cual se logra siguiendo la metodología del Apéndice B.1 pero estableciendo $\delta_{V_T} = 0$.

Es importante destacar que ya que ρ se reduce de forma exponencial el calculo anterior puede ser truncado a la periferia del circuito.

4.2.4. Métrica Completa

Si se toman en consideración todos los efectos anteriores se puede plantear la siguiente métrica genérica para examinar cualquier camino lógico.

$$M_i = \left(\frac{C_{L_i}}{K_i^{a_1}} + \sum_{m=1}^n \left[\rho_m \frac{C_{eff}}{K_{eff}^{a_2}} \right] - \frac{(FD_{n_i} + FD_{p_i})(C_0)}{K_{i-1}^{a_1}} \right) \times \frac{1}{\Delta A} \quad (4.7)$$

Donde

$$C_0 = C_{ox} W_{min} L_{min} \quad (4.8)$$

$$\Delta A = \#nMOS_i \times FD_{n_i} + \#pMOS_i \times FD_{p_i} \quad (4.9)$$

En la ecuación 4.7 claramente el primer factor multiplicativo está formado por tres términos, el primero de ellos hace referencia a que tanto será la disminución en la contribución individual que tiene la compuerta en el σ_P , el segundo se refiere a la disminución de los factores de covarianza relacionados con la compuerta examinada y por su parte el último y tercer término se refiere al efecto de carga producido en la etapa anterior es decir que tanto aumenta la contribución en el σ_P la etapa anterior.

El segundo factor multiplicativo hace referencia a la ponderación de los diferentes incrementos de área causado por las diferentes topologías lógicas.

Capítulo 5

Herramienta CAD para la optimización de caminos lógicos considerando variaciones de proceso

5.1. Introducción

Las herramientas de diseño asistido por computadora (CAD) son la base del diseño VLSI, en el panorama actual existen herramientas de apoyo en cada una de las partes del flujo de diseño de los IC, en lo que respecta a el análisis de retardo existen herramientas que implementan diferentes metodologías para dicho análisis siendo el análisis estático de tiempos (STA) el mas utilizado [2], esta metodología se basa en caracterizar el comportamiento de las compuertas lógicas dentro de un conjunto de librerías con información del retardo de las compuertas bajo diferentes condiciones eléctricas y ambientales. Posteriormente una vez generada la librería esta es usada para el computo del retardo de los caminos lógicos de interés.

Si bien las las herramientas STA son el estándar de la industria [2], estas no toman en cuenta el impacto de las variaciones de proceso sino que solo computan el retardo en el peor caso dando como resultado diseños pesimistas [9]. Reducir este sobre-diseño es una de las bases en las cuales se propone el análisis estadístico de tiempo (SSTA).

Dada esta problemática, se planteo el desarrollo de una herramienta CAD la cual fue implementada por medio del lenguaje de programación Python para la tecnología de 65nm.

El capítulo comienza discutiendo la herramienta desarrollada y dando una introducción a la misma posteriormente se discute la teoría del diseño de experimentos y como utilizar la misma para desarrollar modelos empíricos confiables, posteriormente se aplica dicho conocimiento para caracterizar y validar las compuertas digitales frente a variaciones de proceso. Con los modelos de compuerta validados se estudia como utilizar los mismos para optimizar caminos lógicos tanto por medio de optimización exacta (Método de LaGrange) como por métodos Heurísticos los puntos anteriores fueron la base para el desarrollo de la herramienta.

5.2. Descripción general de la herramienta CAD

La herramienta nació por la necesidad encontrada al existir una falta de software especializado para el análisis de los efectos de las variaciones de proceso en los circuitos integrados, las cuales debido al continuo escalamiento de la tecnología de integración tiene una participación cada vez mucho mayor y sus efectos ya no pueden ser ignorados.

Si bien el clásico simulador eléctrico SPICE permite estudiar estos efectos con precisión, el uso del mismo dentro del análisis de variaciones de proceso en los circuitos VLSI es limitado debido a que este simulador necesita de un sistema Montecarlo para llevar acabo este tipo de análisis, el cual es inviable computacionalmente para circuitos relativamente grandes.

La herramienta utiliza el modelo lineal del retardo basado en sensibilidades precaracterizadas para lograr computar y/o optimizar los parámetros de diseño (dimensiones de los transistores) de un camino lógico con el fin de que un usuario pueda alcanzar diferentes objetivos.

5.2.1. Características

La herramienta CAD es un programa completo para el análisis de caminos lógicos desarrollado en Python 3.3, es un programa integrado y relativamente complejo

ya que cuenta con mas de 3000 líneas de código de alto nivel.

Entre sus características principales de la herramienta se destacan:

- Identificación inteligente de todos los caminos críticos del circuito entrado con el fin de que el usuario aislé el de interés, la identificación. también ordena (por defecto) los caminos en base al retardo de primer orden.
- Permite escoger que tipo de variaciones de proceso serán consideradas. en el catalogo se encuentran V_T , W , L , T_{ox} o cualquier combinación de estas.
- Permite habilitar o deshabilitar correlaciones espaciales en el camino lógico analizado.
- Permite habilitar o deshabilitar los efectos de las variaciones de proceso en la propagación de la señal del camino lógico. analizado.
- Soporte de las compuertas lógicas NAND, NOR y circuitos inversores.
- Soporte de la adición de condensadores para simular efectos de capacitancia en las interconexiones de compuertas.
- Permite la presentación de la capacitancia total en cada nodo del circuito calculándola a partir del numero y tipo de compuertas conectadas al nodo mas los condensadores añadidos.
- Generación de reportes y gráficos en tiempo real.
- Optimización caminos lógicos por el método de multiplicadores de LaGrange con base a diferentes funciones objetivos y restricciones, entre las F.O que es posible escoger se encuentran todas las variables de retardo y el área.
- Aplicación del método heurístico para obtener resultados cuasi-óptimos con un costo computacional bajo.
- Funciona en base a una interfaz gráfica de usuario (GUI), con el fin de ser más amigable al usuario.
- Aun cuando la herramienta solo analiza un camino lógico tiene como soporte la entrada de un circuito combinacional descrito en términos. de compuertas lógicas, con el fin de que el usuario pueda tomar en cuenta los efectos de Fan-Out que cargan los nodos.

La herramienta cuenta con un sinfín de características que facilitan su uso como lo son:

- Interfaz principal bilingüe y totalmente interactiva.
- Generador automático del esquemático para el camino lógico bajo análisis.
- Soporte de esquemáticos y Netlist provenientes de lenguajes externos. como SPICE y Verilog esto a través de dos traductores integrados.
- Capacidad de reportar los análisis realizados por medio de archivos de texto.
- Generación de gráficos pertinente para la visualización del análisis en tiempo real.

Lo anterior solo son actividades de soporte, como se menciona el principal objetivo de la herramienta es el análisis y optimización de caminos lógicos, para esto herramienta cuenta con un conjunto de polinomios pre caracterizados que determinan el comportamiento de las compuertas ante diversos estímulos y cargas, la herramienta utiliza entonces el modelo lineal del retardo para calcular la media y desviación estándar del retardo en caminos lógicos.

Para realizar lo anterior la herramienta

- Calcula las capacitancias en cada nodo del camino lógico en base al numero y tipo de compuertas lógicas conectadas al nodo, mas la posibilidad de agregar capacitancias externas a cada nodo para simular efectos de interconexiones.
- Calculo de la correlación espacial entre las compuertas en base a la posición de Layout (especificada en el Netlist) y el modelo exponencial [5].
- La herramienta automáticamente sensibiliza las compuertas para permitir propagación de señal

5.2.2. Diagrama de flujo general

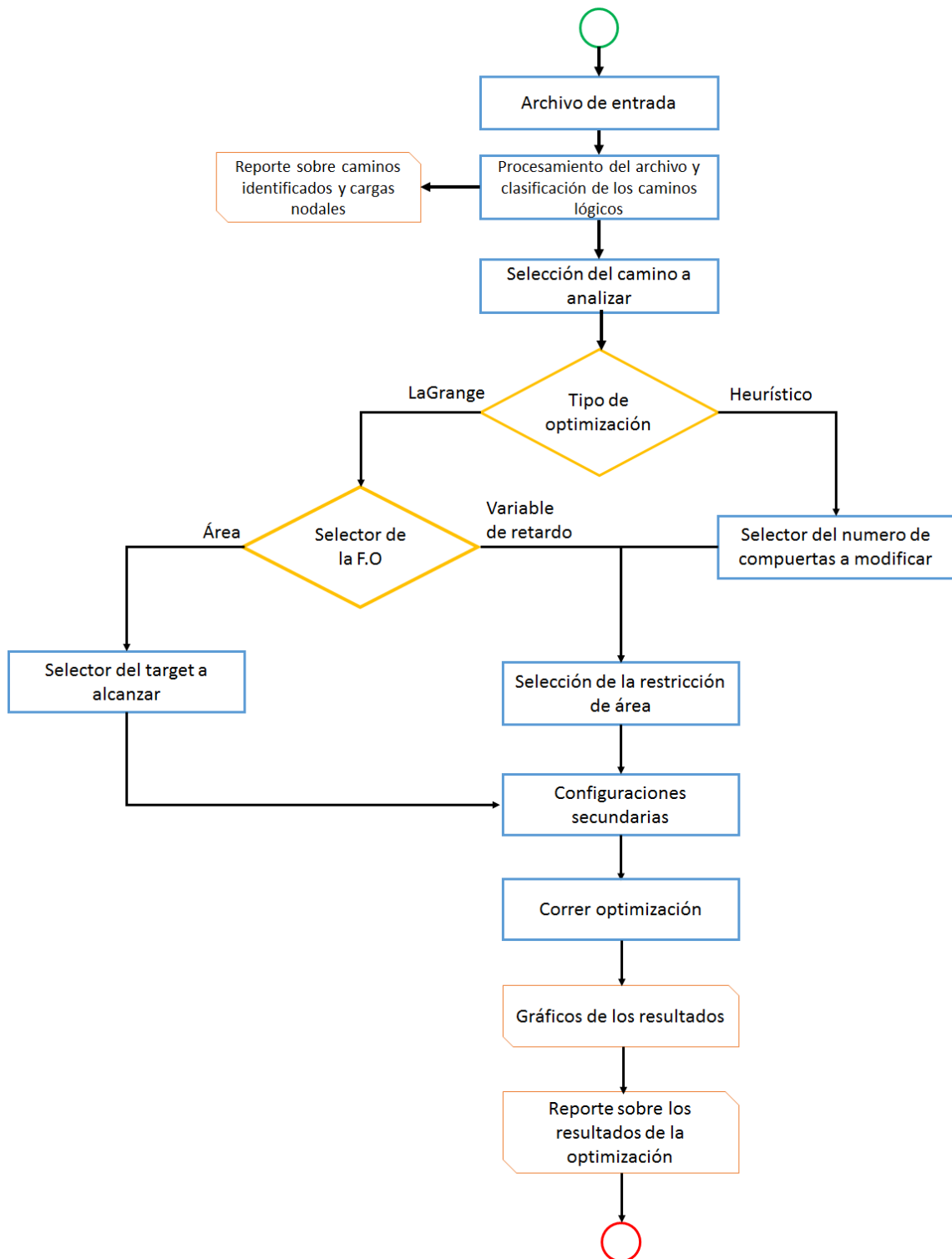


Figura 5.1: Diagrama de flujo compacto de la herramienta CAD

5.2.3. Archivos de entrada y salida de la herramienta

Archivo de entrada

- Contiene la descripción en términos de compuertas lógicas, NAND2 y NOR2 y circuitos inversores.
- Soporte de archivos de entrada en lenguajes externos como Verilog y SPICE.
- Puede incluir capacitancias en los nodos para simular efectos de carga por las interconexiones.
- Se debe indicar si se desea activar o desactivar la correlación espacial en el análisis.
- Debe incluirse en el las variaciones de proceso a considerar.
- Contiene la descripción completa de un circuito, con el fin de contar efectos de carga en los nodos del camino lógico a analizar.

Reporte sobre caminos identificados y cargas nodales

- Contiene datos básicos del circuito analizado: nombre, numero de nodos, número de compuertas ect.
- Contiene las capacitancias efectivas en cada nodo, muestra una lista de los nodos y su carga ordenados de mayor a menor, por medio de las opciones del netlist se puede definir el largo de la lista.
- También incluye información sobre los diferentes caminos lógicos (numero de etapas, retardo en primer orden...) del circuito, los caminos se presentan en forma de lista y son ordenados por su retardo nominal.

Gráficos de los resultados

Los gráficos generados por la herramienta son

1. Histogramas de las distribución del retardo antes y después de la optimización.
2. Comparación entre el área del camino antes y después de la optimización.
3. Comparación de dimensiones de las compuertas antes y después de la optimización.

A continuación se muestran dichos gráficos para un camino de prueba

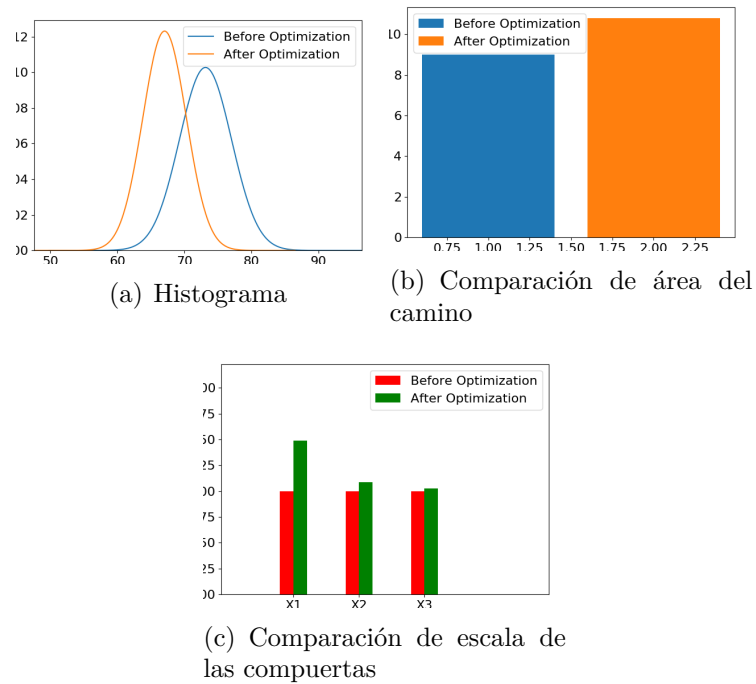


Figura 5.2: Gráficos generados por la herramienta CAD

Reporte sobre los resultados de la optimización

En este reporte se muestra

1. Datos básicos del camino lógico analizado, profundidad, nodo de entrada, nodo de salida y compuertas dentro del camino.
2. Comparativa del σ_P , μ_P y área del camino antes y después de la optimización.
3. Para cada compuerta se muestra:
 - Escala (K),
 - Retardo nominal (μ_D)
 - Desviación estandar del retardo de la compuerta (σ_D)
 - Área utilizada por la compuerta
 - Capacitancia de entrada
 - Auto capacitancia,
 - Capacitancia de carga
 - Contribución de cada variación al valor del σ_D .

Esto es mostrado para antes y después de la optimización

4. La raíz cuadrada de la covarianza entre todo par de compuertas adyacentes debido a la pendiente de entrada, esto es mostrado para antes y después de la optimización
5. Capacitancias total en los nodos del camino antes y después de la optimización
6. Matriz de correlación espacial en caso de incluirse en el análisis.

Un manual completo de la herramienta se puede encontrar en el Apéndice [C.1.2](#).

5.3. Diseño de experimentos

Para el desarrollo de la herramienta CAD es necesario generar modelos confiables para el calculo del retardo y su desviación estándar. Para esto se usa la técnica conocida como diseño de experimentos (DOE).

DOE es una metodología experimental que ayuda a investigar los efectos de un conjunto de variables de entrada sobre una variable de salida. Estos experimentos consisten en una serie de corridas en las que se realizan cambios intencionales en las variables de entrada. En cada corrida se recolectan los datos necesarios que posteriormente pueden ser analizados de manera matemática para obtener un modelo que relacione las salidas del sistema en función de sus entradas.

5.3.1. Diseño Factorial

Existen muchas variantes del método DOE siendo el método factorial la versión mas robusta cuando se examina mas de una sola variable de entrada [12], en el diseño factorial se analizan todas las posibles combinaciones entre las entradas mediante el conjunto de experimentos realizados con el fin de estudiar el efecto conjunto generado sobre la respuesta.

Llevados acabo los experimentos se puede generar una tabla que relacione cada posible combinación de variables de entrada con las salidas obtenidas para dicha corrida, con todo el espacio de estos datos se puede entonces ajustar un polinomio de grado m para generar la función requerida .

A las variables de entrada estudiadas dentro del diseño factorial se les denomina propiamente como factores (f), para su análisis cada factor es dividido en un conjunto finito de valores dentro de su rango de acción a cada punto dentro de dicho conjunto se le denomina como nivel (l), cuando el diseño factorial tiene el mismo numero de niveles para cada variable se le denomina a este como diseño factorial completo y requiere un numero total de corridas f^l , de otra forma si el numero de niveles utilizado es asimétrico entre los factores se habla de un diseño factorial mixto.

Tomando como referencia el estudio de un fenómeno donde se tienen bajo control dos factores llamados A y B que pueden tomar dos niveles distintos. Si se quiere entonces encontrar la relación de estos factores con la observación del fenómeno llamado Y se puede llevar acabo un diseño factorial completo combinando los

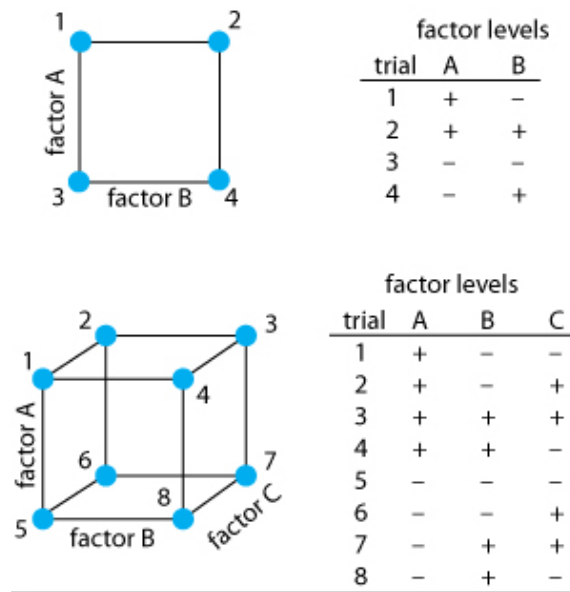


Figura 5.3: Diseño factorial de dos y tres factores

diferentes valores de los factores y registrando el resultado de cada observación como lo muestra la Tabla 5.1

Experimento	A	B	Y
1	-1	-1	20
2	-1	1	30
3	1	-1	40
4	1	1	52

Tabla 5.1: Ejemplo de diseño factorial

Utilizando la Tabla 5.1 se puede realizar un gráfico de las interacciones, esto se muestra en la Figura 5.4.

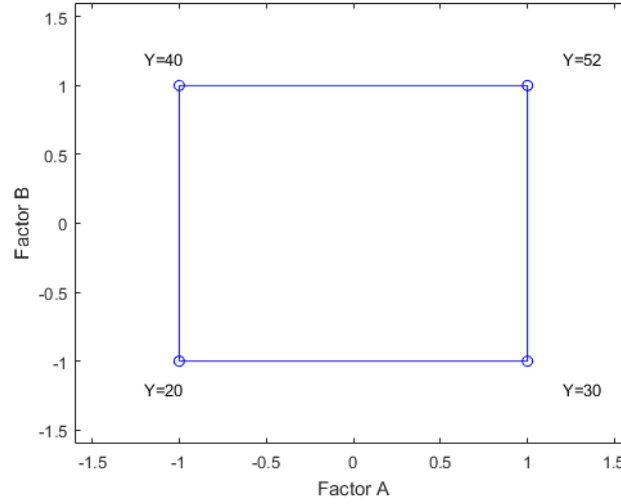


Figura 5.4: Diseño factorial con dos factores

Modelo DOE

Utilizando los datos del experimento anterior se puede plantear un modelo de regresión.

$$Y = C_{00} + C_{10}A + C_{01}B + C_{11}AB \quad (5.1)$$

Donde los parámetros C representan la ponderación aislada de cada factor y sus interacciones, para calcular a estos coeficientes se utiliza la técnica de los mínimos cuadrados la cual seleccionara los parámetros de ponderación minimizando el cuadrado de la diferencia entre el modelo propuesto y las observaciones [13], para buscar este punto mínimo se utilizan técnicas de optimización las cuales serán discutidas a detalle mas adelante, por ahora se muestra únicamente detalles sobre los cálculos realizados para llevar a cabo el ajuste.

Definiendo el residuo entre el modelo y el punto de evaluación i, j .

$$\epsilon_{i,j}(\mathbf{C}) = y_{i,j} - (C_{00} + C_{10}A_{i,j} + C_{01}B_{i,j} + C_{11}A_{i,j}B_{i,j}) \quad (5.2)$$

Tal que

$$\epsilon_{0,0}(\mathbf{C}) = 20 - (C_{00} - 1C_{10} - 1C_{01} + 1C_{11}) \quad (5.3)$$

$$\epsilon_{0,1}(\mathbf{C}) = 30 - (C_{00} - 1C_{10} + 1C_{01} - 1C_{11}) \quad (5.4)$$

$$\epsilon_{1,0}(\mathbf{C}) = 40 - (C_{00} + 1C_{10} - 1C_{01} - 1C_{11}) \quad (5.5)$$

$$\epsilon_{1,1}(\mathbf{C}) = 52 - (C_{00} + 1C_{10} + 1C_{01} + 1C_{11}) \quad (5.6)$$

Se define entonces el termino sumando cuadrático del error (S) como:

$$S(\mathbf{C}) = \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (\epsilon_{i,j})^2 \quad (5.7)$$

El mínimo de la función S se puede encontrar por medio de la raíz su respectivo gradiente (∇S), los valores de los coeficientes C en dicho punto corresponden con el ajuste optimo, bajo esta consideración se muestra el modelo obtenido para el ejemplo bajo consideración.

$$f(A, B) = 35,5 + 10,5A + 5,5B + 0,5AB \quad (5.8)$$

La ecuación 5.8 se gráfica en la Figura 5.5.

Si bien el ajuste de los coeficientes por medio de la técnica de mínimos cuadrados permite obtener el ajuste mas optimo del modelo, esto no garantiza que el modelo sea adecuado para moldear el fenómeno estudiado. Para conocer la confiabilidad del modelo deben utilizarse métodos estadísticos como el análisis de variación (ANOVA) [12].

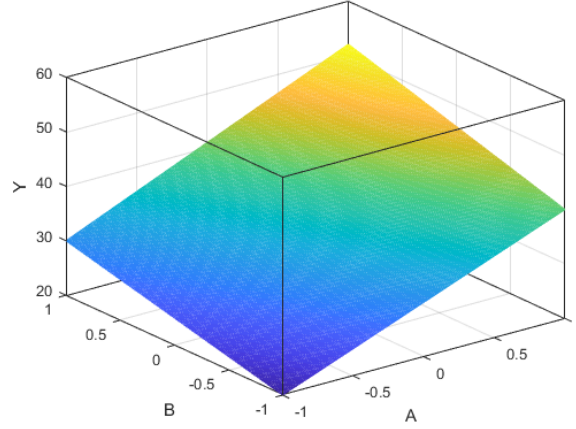


Figura 5.5: Superficie de respuesta

ANOVA es una metodología para comparar diferencias entre dos o mas grupos bajo un conjunto de consideraciones o supuestos [12], dichos supuestos se mencionan a continuación.

1. Los datos de cada uno de los grupos bajo análisis son independientes.
2. Los datos bajo análisis siguen una distribución normal en sus residuos.
3. Los datos bajo análisis presentan una varianza equitativamente distribuida en su rango.

Los principios anteriores se denominan respectivamente como: 1) principio de linealidad, 2) principio de normalidad y 3) principio de homegenidad

Si se toman dichos principios como axiomas la varianza del experimento se puede expresar en términos de la varianza entre los factores y su error [12]:

$$\sum (y_i - \bar{y})^2 = \sum (f_i - \bar{y})^2 + \sum (y_i - f_i)^2 \quad (5.9)$$

En el ANOVA se utilizan la siguientes nomenclaturas para definir las cantidades anteriores.

$$SS_T = \sum (y_i - \bar{y})^2 \quad (5.10)$$

$$SS_R = \sum (f_i - \bar{y})^2 \quad (5.11)$$

$$SS_E = \sum (y_i - f_i)^2 \quad (5.12)$$

Comunmente el ANOVA es utilizado en el contexto de regresión para realizar una prueba de significancia de los parámetros ajustados, esta prueba se realiza al plantear una hipótesis nula:

$$H_0 : C_{00} = C_{01} = C_{10} = C_{11} = 0 \quad (5.13)$$

El rechazo de dicha hipótesis implica que al menos uno de los coeficientes de ajuste contribuye significativamente para reducir el error entre los datos y el modelo desarrollado.

Si se denomina al numero de parámetros examinados p y al numero de coeficientes de ajuste como k , se puede probar que el cociente de varianzas F_0 se puede expresar como:

$$F_0 = \frac{SS_R/k}{SS_E/(n-p)} \quad (5.14)$$

Si el valor de F_0 es mayor que el coeficiente Fisher $F_{0,5,k,n-p}$ dado por la Tabla F se puede entonces rechazar a la hipotesis nula.

Un cuantificador de desempeño global del ajuste del modelo se determina como sub-producto de un análisis ANOVA completo y se denomina como coeficiente de determinación R^2 , el cual es un estadístico que mide el nivel de ajuste de los datos, este numero se encuentra acotado entre 0 y 1, el valor máximo representa un ajuste perfecto.

$$R^2 = 1 - \frac{SS_E}{SS_T} \quad (5.15)$$

Cabe destacar que el coeficiente R^2 aumenta con el numero de coeficientes de ajuste k el cual es una medida proporcional al grado del polinomio utilizado, sin embargo conforme mayor se hace este grado los puntos interpolados por el análisis empiezan a sufrir el llamado fenómeno de Runge que hace a estos puntos oscilar, por lo que conforme aumenta el grado se mejora el ajuste en los puntos donde se evaluó el experimento pero se reduce el valor del modelo para interpolar datos dentro del rango de acción.

5.3.2. Caracterización de la librería digital

El proceso de caracterizar las compuertas digitales para la herramienta CAD esta basado en el DOE. La metodología consiste en simular un análisis transitorio para cada tipo de compuerta variando los parámetros de entrada con el fin de medir el efecto de estos en el retardo para posteriormente realizar el ajuste polinomial, en general las diferentes variables que afectan el retardo de una compuerta son la pendiente de entrada de la compuerta (SR_I), la capacitancia de carga C_L , escala de la compuerta K , tipo de conmutación, tipo de compuerta y entrada por la cual se genera la conmutación [5].

El simulador seleccionado para realizar la caracterización sera la herramienta SPICE, las variables a medir de interés serán:

Variable	Simbolo	Variable	Simbolo
Retardo Nominal	μ_D	Pendiente nominal de salida de la compuerta	SR_O
Sensibilidad del retardo con respecto al V_T	$S_{V_T}^D$	Sensibilidad del retardo con respecto al W	S_W^D
Sensibilidad del retardo con respecto al L	S_L^D	Sensibilidad del retardo con respecto al T_{ox} .	$S_{T_{ox}}^D$
Sensibilidad del retardo con respecto a SR_I	$S_{SR_I}^D$	Sensibilidad de SR_O con respecto a SR_I	$S_{SR_I}^{SR_O}$
Sensibilidad del SR_O con respecto a V_T	$S_{V_T}^{SR_O}$	Sensibilidad del SR_O con respecto a W	$S_W^{SR_O}$
Sensibilidad del SR_O con respecto a L	$S_L^{SR_O}$	Sensibilidad del SR_O con respecto a T_{ox}	$S_{T_{ox}}^{SR_O}$

Tabla 5.2: Parámetros a caracterizar

Todas las variables de interés serán caracterizadas para cada una de las entradas de las compuertas NOT, NAND2 y NOR2 con respecto a K , C_L y SR_I , para esto se utiliza una lista de simulación variando los parámetros anteriores en el siguiente rango.

Capítulo 5. Herramienta CAD para la optimización de caminos lógicos considerando variaciones de proceso

Factor	Simbolo	Valor minimo	Valor maximo	Paso	Total
Escala de la compuerta	K	1	10	1	10
Carga externa de la compuerta	C_L	0.5	15	0.5	30
Pendiente de Entrada	SR_I	5	300	5	60

Tabla 5.3: Variables a examinar

En general para cada compuerta analizada se realizaran 18 000 simulaciones por cada entrada de la misma.

Caracterización de parámetros nominales (D, SR_O)

Las dos primeras variables de interés a caracterizar son el retardo nominal D y pendiente de salida SR_O , debido a que estas cantidades son totalmente nominales su medición es directa atreves de las simulaciones, la metodología a seguir se muestra en la Figura 5.6.

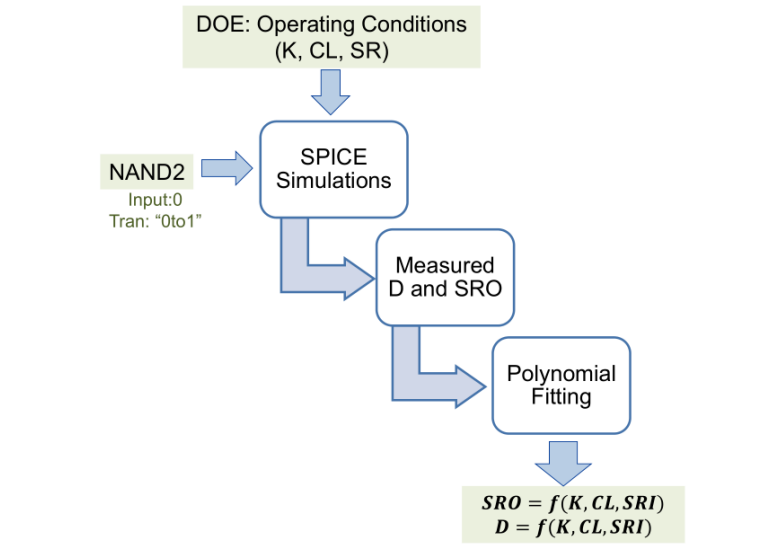


Figura 5.6: Metodología de caracterización de cantidades nominales

Este proceso deberá repetirse para todas las entradas y conmutaciones de las compuertas de interés.

	<i>K</i>	<i>CL</i>	<i>SRI</i>	<i>D</i>	<i>SRO</i>	<i>NAND2,n0,"0to1"</i>
	<i>K</i>	<i>CL</i>	<i>SRI</i>	<i>D</i>	<i>SRO</i>	<i>NAND2,n0,"1to0"</i>
	<i>K</i>	<i>CL</i>	<i>SRI</i>	<i>D</i>	<i>SRO</i>	<i>NAND2,n1,"0to1"</i>
	<i>K</i>	<i>CL</i>	<i>SRI</i>	<i>D</i>	<i>SRO</i>	<i>NAND2,n1,"1to0"</i>
	1	0.2fF	20ps	10ps	12ps	
	2	0.4fF	20ps	15ps	16ps	
	4	0.6fF	30ps	20ps	19ps	
	6	0.8fF	50ps	25ps	23ps	
	⋮	⋮	⋮	⋮		

Figura 5.7: Resultados de la metodología de caracterización nominal, los valores numéricos son únicamente con fin de ilustración

Caracterización y análisis de las sensibilidades

El objetivo final de caracterizar las sensibilidades es simplemente obtener el valor del σ_D como función de los parámetros de entrada a través del modelo estadístico del retardo, se puede cuestionar la duda de por que es necesario los niveles de sensibilidad y no simplemente abordar directamente el valor de σ_D , la respuesta a esto se debe a que no es muy viable por simulación obtener directamente dicho valor ya que esto requiere simulaciones del tipo Monte Carlo que son excesivamente costosas computacionalmente, esto juntado al alto numero de simulaciones requeridas por el DOE del tipo factorial vuelve el proceso completamente ineficiente [5].

Para computar los modelos de sensibilidad se puede hacer uso de la definición de de sensibilidad la cual simplemente indica que la misma es la derivada de la variable de interés con respecto al parámetro examinado, sin embargo comúnmente la derivada no sera una constante atreves de todo el rango y tomar en cuenta esto vuelve nuevamente inviable el análisis en términos computacionales, una alternativa practica es definir la sensibilidad en solo dos puntos muestras que linealicen la función en un intervalo de interés como lo muestra la Figura 5.8.

Bajo este análisis la sensibilidad se puede calcular simplemente como

$$S_L^D = \frac{D' - D_n}{\Delta L} \quad (5.16)$$

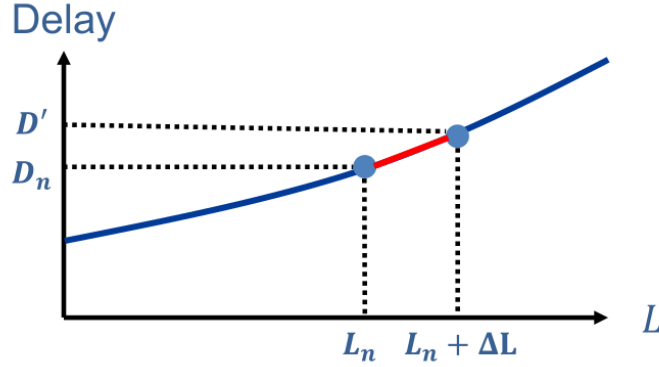


Figura 5.8: Metodología de caracterización de la sensibilidad

Donde D_n es el retardo de la compuerta evaluada por su parte D' es el retardo de una compuerta bajo las mismas condiciones eléctricas y ambientales pero con un cambio en la variable de proceso examinada este cambio llamado L' se escoge a una distancia de $2\sigma_L$ tal que genera un $\Delta L = 2\sigma_L$.

Claramente el proceso se puede utilizar directamente para obtener todo tipo de sensibilidades. En compuertas con redes de transistores en serie es posible solo usar una variable aleatoria para moldear el comportamiento estadístico de las variaciones en W , L y T_{ox} ya que debido a la alta proximidad de los transistores en el Layout la correlación entre dichas variables sera total, por el contrario variaciones en el V_T son del tipo no correlacionadas y por lo tanto requieren de sensibilidades separadas para su modelado, en redes en paralelo el modelo caracterizado asume que solo un transistor conmuta por lo cual el modelado solo requiere cuantificar el comportamiento de un solo transistor.

5.3.3. Computo del σ_D

Utilizando las sensibilidades se puede obtener el σ_D de una compuerta lógica siguiendo el modelo estadístico del retardo:

$$\sigma_D^2 = (S_{V_T}^D \sigma_{V_T})^2 + (S_W^D \sigma_W)^2 + (S_L^D \sigma_L)^2 + (S_{T_{ox}}^D \sigma_{T_{ox}})^2 + (S_{SR_I}^D \sigma_{SR_I})^2 \quad (5.17)$$

El valor de σ_{SR_I} para una compuerta es una función de las sensibilidades correspondientes de la etapa anterior.

$$\sigma_{SR_i}^2 = (S_{V_{T_{i-1}}}^{SR} \sigma_{V_{T_{i-1}}})^2 + (S_{W_{i-1}}^{SR} \sigma_{W_{i-1}})^2 + (S_{L_{i-1}}^{SR} \sigma_{L_{i-1}})^2 + (S_{T_{ox_{i-1}}}^{SR} \sigma_{T_{ox_{i-1}}})^2 + (S_{SR_{I_{i-1}}}^{SR} \sigma_{SR_{I_{i-1}}})^2 \quad (5.18)$$

Estas ultimas situaciones causan un nuevo tipo de correlación llamada correlación por propagación de señal, la cual relaciona las variaciones de proceso de la compuerta $i - 1$ con la señal aplicada a la compuerta i , mas adelante se describirá de manera funcional este tipo de correlación.

El retardo nominal de un camino lógico se puede computar con base a la ecuación 2.98, por su parte la desviación estándar se analiza con base a la ecuación 2.99, las dos ecuaciones discutidas anteriormente se muestran en esta sección para mayor comprensión.

$$\mu_P = \sum_{i=1}^n \mu_{D_i} \quad (5.19)$$

$$\sigma_P^2 = \sum_{i=1}^n \sigma_{D_i}^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{COV}(D_i, D_j) \quad (5.20)$$

El termino de correlación esta compuesto de dos terminos como lo muestra la ecuación 5.21.

$$\text{COV}(D_i, D_j) = \lambda + \Psi \quad (5.21)$$

Donde el primer termino se debe a la correlación espacial descrita en el Capitulo 2, este termino puede expresarse en términos de las sensibilidades de cada etapa:

$$\lambda = \rho_{i,j} \left((S_{W_i}^D S_{W_j}^D) \sigma_W^2 + (S_{L_i}^D S_{L_j}^D) \sigma_L^2 + (S_{T_{ox_i}}^D S_{T_{ox_j}}^D) \sigma_{T_{ox_j}}^2 \right) \quad (5.22)$$

Es importante recordar que las variaciones en el V_T son independientes por lo que no presentan correlación espacial, por su parte ρ vendrá dado por una función entre la distancia física entre la compuerta i y j como se estudio en el Capitulo 2.

Por su parte Ψ representa la correlación debida a la propagación de señal, la cual se debe a que la generación de variaciones en el SR_I proviene en realidad de las variaciones de proceso encontradas en la etapa anterior, matemáticamente esto se puede expresar como:

$$\begin{aligned} \Psi = & S_{V_{T_{i-1}}}^{SR} \sigma_{V_{T_{i-1}}} S_{SR_{I_i}}^D \sigma_{SR_{I_i}} + S_{W_{i-1}}^{SR} \sigma_{W_{i-1}} S_{SR_{I_i}}^D \sigma_{SR_{I_i}} + \\ & S_{L_{i-1}}^{SR} \sigma_{L_{i-1}} S_{SR_{I_i}}^D \sigma_{SR_{I_i}} + S_{T_{ox_{i-1}}}^{SR} \sigma_{T_{ox_{i-1}}} S_{SR_{I_i}}^D \sigma_{SR_{I_i}} + S_{SR_{I_{i-1}}}^{SR} \sigma_{SR_{I_{i-1}}} S_{SR_{I_i}}^D \sigma_{SR_{I_i}} \end{aligned} \quad (5.23)$$

5.4. Validación de la caracterización

Caracterizada la librería digital se procede a realizar una comparación con el simulador eléctrico SPICE para comprobar el correcto funcionamiento de los polinomios obtenidos por el proceso de diseño de experimentos, en general se llevaran a cabo dos tipos de validaciones: 1) Validación a nivel compuerta lógica y 2) Validación a nivel camino lógico.

La primera validación hace referencia únicamente al desempeño de los polinomios caracterizados, por su parte la segunda validación se refiere a la actuación del modelo estadístico del retardo (ecuaciones 2.98 y 2.99) implementado por medio de los polinomios, esta ultima validación se analizará con y sin la inclusión de la correlación espacial en el camino lógico.

5.4.1. Validación a nivel compuerta lógica

La validación de las compuertas lógicas se lleva a cabo para las compuertas NOT, NAND2 y NOR2 pre-caracterizadas por medio del proceso descrito en la sección 5.3.2, para la comparación con SPICE se analizara el caso de compuertas de dimensiones mínimas bajo diferentes condiciones de carga, las cantidades a comparar serán el retardo nominal y la desviación estándar, se examinara para cada compuerta el desempeño de la red PullDown y PullUp. Los resultados de las validaciones se muestran en las Figuras 5.9 a 5.11

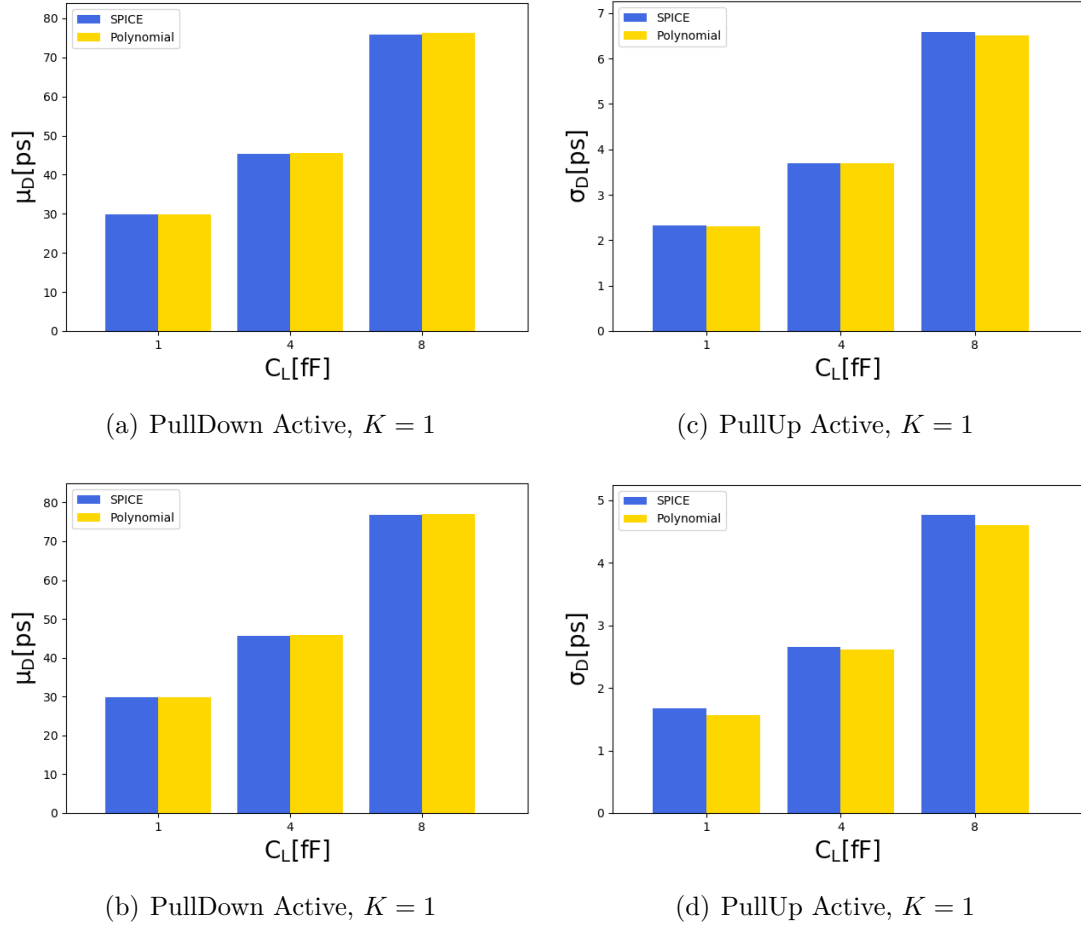


Figura 5.9: Comparación entre los polinomios caracterizados y SPICE en el Inversor

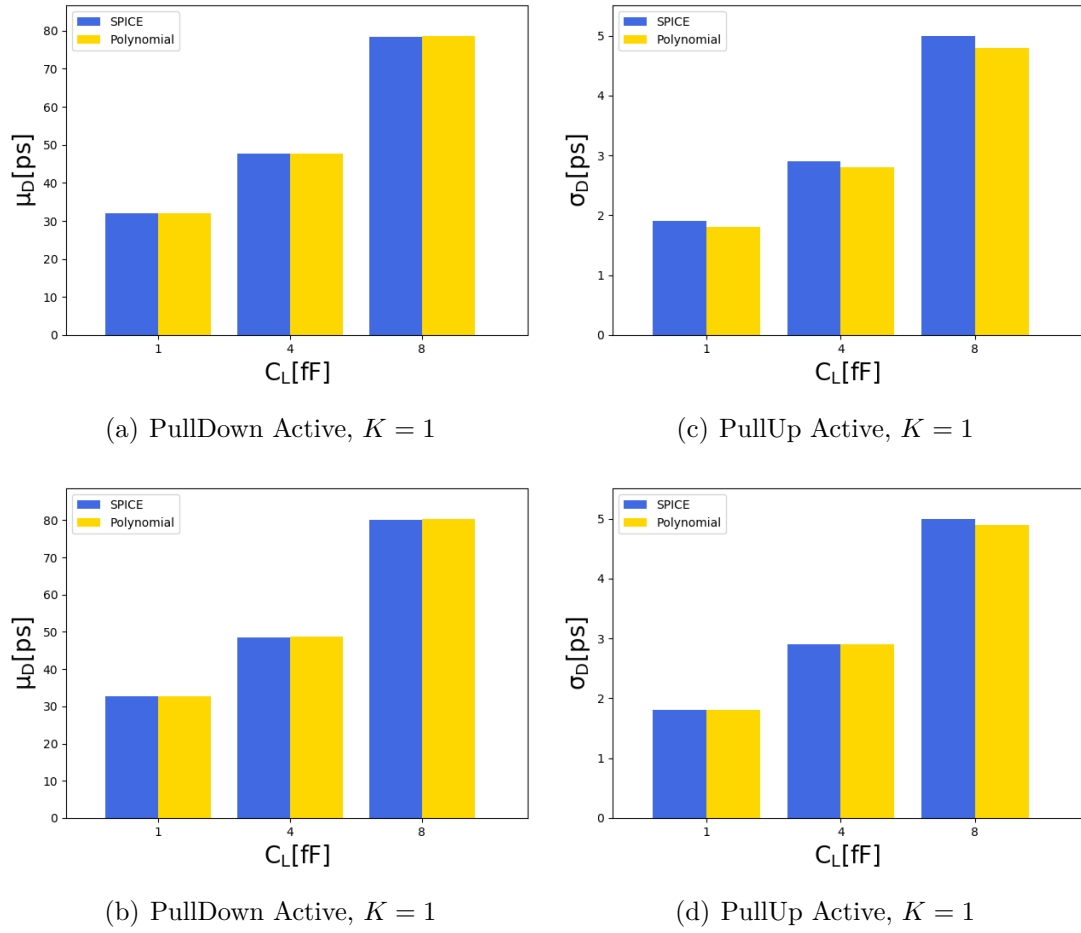


Figura 5.10: Comparación entre los polinomios caracterizados y SPICE en la NAND2

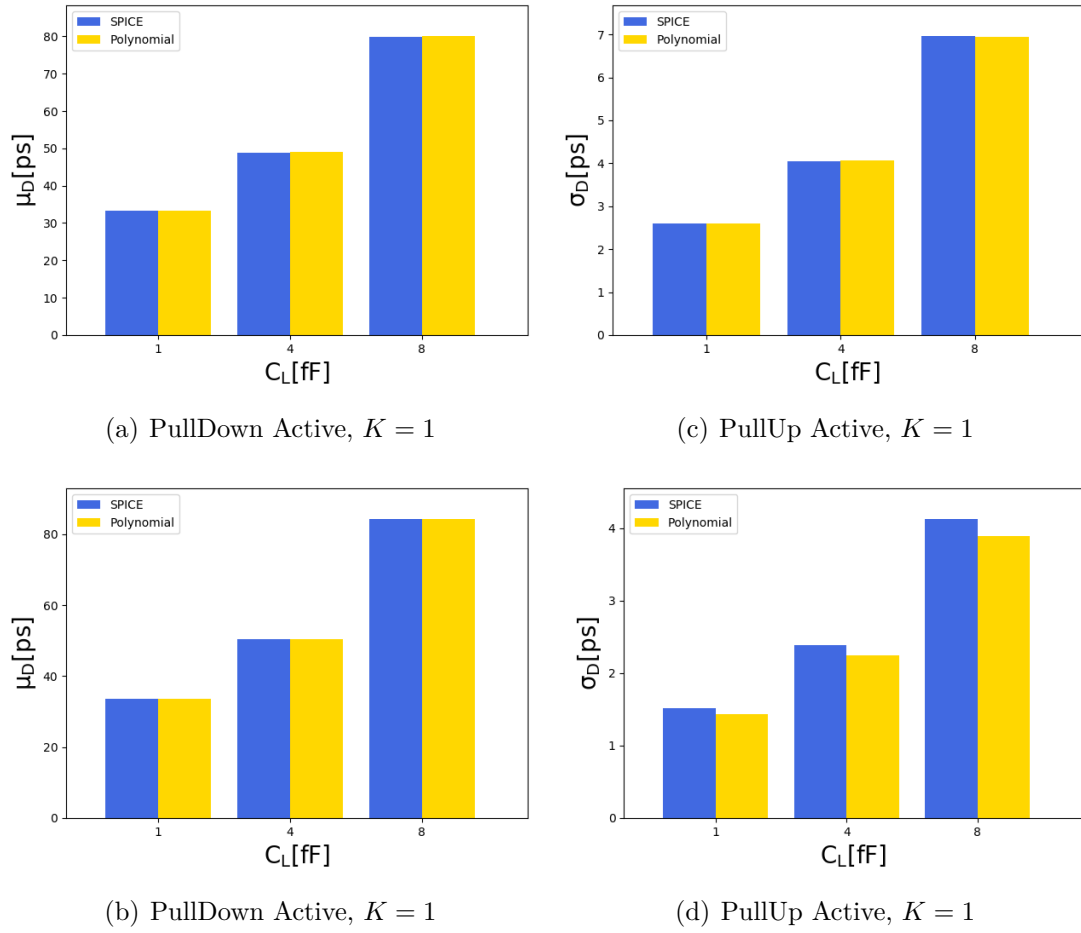


Figura 5.11: Comparación entre los polinomios caracterizados y SPICE en la NOR2

5.4.2. Validación en camino lógico básico

Para esta validación se considera el camino lógico mostrado en la Figura 5.12.

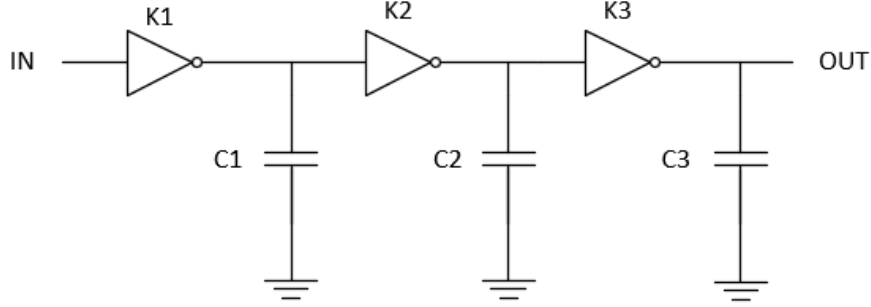


Figura 5.12: Camino básico para validación

Tal que

$$K_1 = K_2 = K_3 = 1 \quad (5.24)$$

$$C_1 = C_2 = C_3 = 5[\text{fF}] \quad (5.25)$$

Para incluir la correlación espacial en el camino se toma la matriz de correlaciones de la ecuación 5.26, dicha matriz es propuesta en base al comportamiento típico encontrado en los caminos lógicos (Ver Figura 2.15).

$$\rho = \begin{bmatrix} 1,00 & 0,83 & 0,62 \\ 0,83 & 1,00 & 0,74 \\ 0,62 & 0,74 & 1,00 \end{bmatrix} \quad (5.26)$$

Comparando el retardo nominal μ_P y la desviación estándar σ_P del retardo para el camino lógico se forma la Tabla 5.4.

Engine	μ_P	σ_P	Units
SPICE	213.80	11.20	[ps]
Polynomial Based Model	204.79	12.44	[ps]
Error	4.85	7.38	[%]

Tabla 5.4: Comparación en el camino básico

Es de destacar que SPICE no permite la inclusión variables aleatorias correlacionadas de manera directa, sin embargo se puede utilizar la transformación de Cholesky descrita en el Apéndice A.7 para superar esta inconveniente.

5.4.3. Validación en un camino lógico compuesto

Para esta validación se considera el camino lógico mostrado en la Figura 5.13.

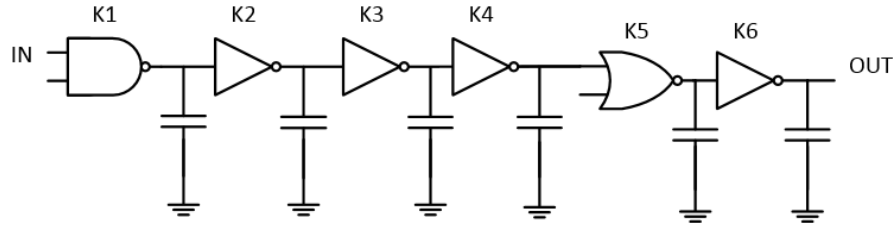


Figura 5.13: Camino compuesto para validación

Tal que

$$K_1 = K_2 = K_3 = K_4 = K_5 = K_6 = 1 \quad (5.27)$$

$$C_1 = C_2 = C_3 = C_4 = C_5 = C_6 = 5[\text{fF}] \quad (5.28)$$

Se incluye la correlación espacial por medio de la matriz 5.29 la cual se propone nuevamente en base al comportamiento típico observado en los caminos lógicos.

$$\rho = \begin{bmatrix} 1,00 & 0,83 & 0,69 & 0,57 & 0,47 & 0,39 \\ 0,83 & 1,00 & 0,83 & 0,69 & 0,57 & 0,47 \\ 0,69 & 0,83 & 1,00 & 0,83 & 0,69 & 0,57 \\ 0,57 & 0,69 & 0,83 & 1,00 & 0,83 & 0,69 \\ 0,47 & 0,57 & 0,69 & 0,83 & 1,00 & 0,83 \\ 0,39 & 0,47 & 0,57 & 0,69 & 0,83 & 1,00 \end{bmatrix} \quad (5.29)$$

Tal que al realizar la simulación eléctrica apoyada por la transformación de Cholesky (Apéndice A.7) y comparar con el modelo Polinómico se forma la Tabla 5.5.

Engine	μ_P	σ_P	Units
SPICE	470.50	21.77	[ps]
Polynomial Based Model	448.64	20.27	[ps]
Error	4.85	7.38	[%]

Tabla 5.5: Comparación del camino lógico compuesto

Se concluye la validación.

5.5. Multiplicadores de LaGrange

Los multiplicadores de LaGrange son un método que permite encontrar los puntos notables (máximos y mínimos) de funciones de múltiples variables sujetas a restricciones, el método de LaGrange permite plantear un sistema de $n + k$ ecuaciones donde n es el número de variables de la función a optimizar y k es el número de restricciones a las cuales se encuentra sujeta y las cuales agregan cada una un nuevo tipo de variable independiente al sistema, estas variables auxiliares se conocen propiamente como multiplicadores de LaGrange. Una vez planteado el sistema a resolver este puede ser solucionado con un algoritmo para encontrar las raíces o ceros de un sistema de ecuaciones como por ejemplo Newton-Raphson.

El tipo más simple de optimización es la optimización sin restricciones o no acotada, la cual es una aplicación del cálculo diferencial ya que cuando no existen restricciones el problema de optimización se reduce directamente a hallar los puntos máximos o mínimos de la función a optimizar (f), lo cual se logra al solucionar la ecuación 5.30.

$$\nabla f = \vec{0} \quad (5.30)$$

En la ecuación anterior ∇f representa el gradiente de la función f el cual es una operación de carácter vectorial que opera sobre una función escalar para producir

un campo vectorial cuyos elementos tienen como magnitud la máxima razón de cambio posible para un incremento infinitesimal de la función en el punto de evaluación, todos los elementos del campo siempre apunta en dirección de los valores máximos de la función [14].

Para una función multivariable $f(x_1, x_2 \dots x_n)$ se define su gradiente como:

$$\nabla f(x_1, x_2 \dots x_n) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \quad (5.31)$$

Gráficamente el gradiente es representado como un campo vectorial que siempre apunta en la dirección de mayor ascenso en la función (por su parte el negativo del gradiente apunta hacia la dirección de mayor descenso), una propiedad del gradiente es que en un mapa de curvas de nivel este siempre es ortogonal a las curvas (ya que por definición en el contorno de una curva de nivel no existe cambio).

Un ejemplo gráfico de lo discutido anteriormente se muestra a continuación en la Figura 5.14.

En un problema de optimización sin restricciones el valor optimo se encuentra buscando los puntos donde el gradiente de la función toma el valor cero, ya que estos puntos representan máximos y mínimos (locales o globales) que pueden ser evaluados con respecto a un criterio como las condiciones de Karush–Kuhn–Tucker para saber a que categoría pertenece.

Cuando se agregan restricciones al problema de optimización buscar los puntos notables comúnmente no proporcionara soluciones dentro de los márgenes de restricción por lo que para considerar estas se deben incluir dentro de la función a optimizar los multiplicadores de LaGrange (λ_i) los cuales son variables auxiliares que permiten resolver el problema de optimización y una vez resuelto pueden ser desechadas, Si se tiene una función f como función objetivo (F.O) para optimizar acompañada por un conjunto de funciones g_i que especifican restricciones que debe cumplir la F.O, se denomina entonces como el Lagrangiano del sistema a la cantidad.

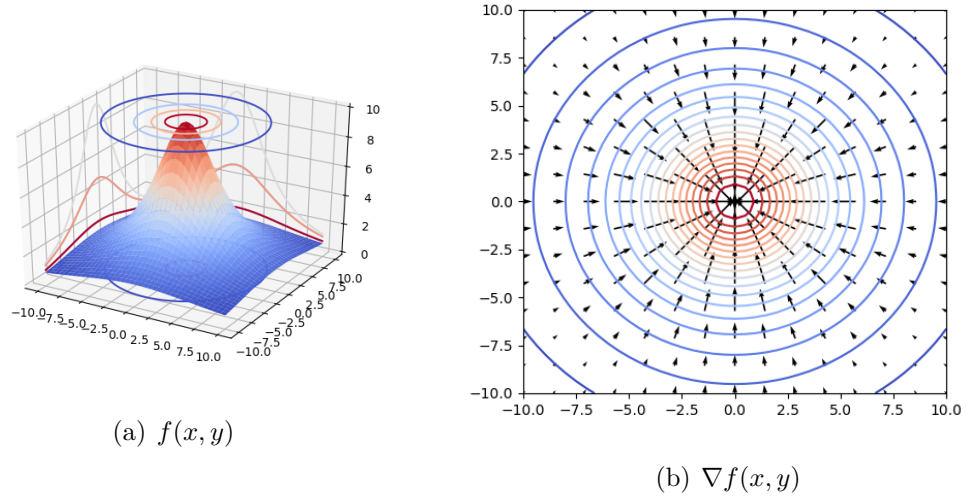


Figura 5.14: El gradiente es un campo vectorial cuyos elementos siempre apuntan en torno al máximo de la función

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^k \lambda_i g_i(\vec{x}) \quad (5.32)$$

Tal que los puntos notables de f que satisfacen el conjunto de restricciones se pueden encontrar por medio de la solución simultanea de las raíces para cada componente del $\nabla \mathcal{L}$, para ejemplificar esto se proporcionan ejemplos de aplicación de los multiplicadores se encuentra en el Apéndice C.1.

En la mayoría de los lenguajes de programación modernos existen paquetes o librerías estandarizadas para abordar los problemas de optimización por ejemplo en Python la librería *scipy.optimize* ofrece de forma integrada la subrutina *minimize* la cual permite realizar optimización con o sin restricciones por medio de diversos métodos numéricos, dentro de dicha subrutina se debe especificar la función objetivo, funciones de restricciones, punto de inicio y método de solución. Un diagrama de bloque de este proceso se muestra en la Figura 5.15.

Para el desarrollo de la herramienta SSTA se usara la subrutina *minimize* discutida anteriormente con el método de solución *SLSQP*, el cual se basa en plantear el Lagrangiano del problema para posteriormente encontrar su raíz usando el método Newton-Raphson y evaluando cada punto por medio de las condiciones de Karush–Kuhn–Tucker para identificar cada categoría de los puntos notables.

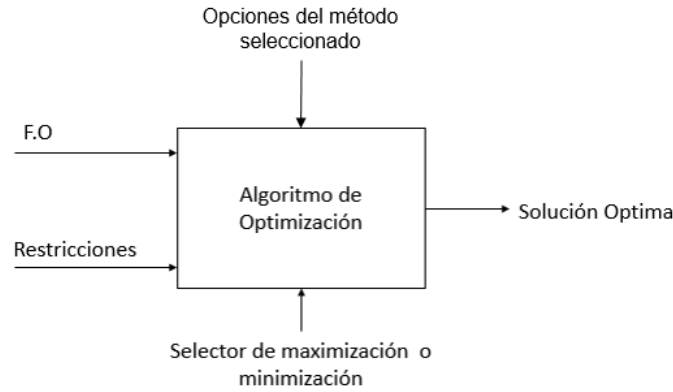


Figura 5.15: Diagrama de bloques de proceso de optimización

5.5.1. Optimización de caminos lógicos bajo una variable de retardo

Utilizando los polinomios caracterizados por DOE se puede obtener el retardo nominal y desviación estándar de cualquier camino lógico formado por las compuertas caracterizadas, gracias a la metodología de los multiplicadores de LaGrange se puede usar la función que describe el valor del σ_P o μ_P o incluso una combinación de ambas como una F.O a minimizar.

Como se mencionó anteriormente los tamaños de las compuertas en el camino son las variables principales que tiene el diseñador bajo su control, en general tanto el σ_P como μ_P son inversamente proporcionales a dichos tamaños por lo que compuertas más grandes son generalmente más rápidas y robustas a variaciones de proceso, sin embargo el área es uno de los recursos más preciados en el diseño de circuitos integrados por lo que incrementar la escala de las compuertas indefinidamente no es una opción factible, más bien se busca una repartición de recursos, por lo que utilizando los modelos matemáticos desarrollados anteriormente junto con técnicas de optimización se pueden implementar metodologías de diseño robustas. Claramente podemos notar entonces que para una cadena de N compuertas sometidas a capacitancias nodales fijas se tienen N variables ($K_1, K_2 \dots K_n$) donde cada una de estas variables será conocida como variables de decisión o de diseño ya que como se mencionó la escala de las compuertas es el parámetro que está bajo la influencia directa del diseñador de circuitos integrados.

Debido a que tanto el área del camino como el valor de las variables de retardo son funciones del tamaño de cada una de las compuertas en el camino

$(K_1, K_2 \dots K_n)$, el problema de optimización es directo de plantear, si bien en este análisis F.O sera una variable de retardo el seleccionar la misma no es trivial y diversas propuestas se han realizado por diferentes grupos de investigación, esto es así por que conocemos que las variaciones de proceso convierten el retardo en una variable aleatoria con distribución normal la cual queda definida por dos parámetros σ_P y μ_P , por lo que la función objetivo deberá ser una función que utilice una o ambas variables.

Es importante recordar que el valor de μ_P representa el valor esperado del retardo en el circuito por lo que implica que a valores más pequeños los circuitos fabricados serán en promedio cada más rápidos, también como se mencionó anteriormente este valor representa el retardo nominal de un camino lógico por lo que μ_P no es dependiente de la variaciones de proceso. Por otro lado el σ_P representa la desviación estándar de los circuitos fabricados y es generada por las variaciones de proceso, es una medida de la dispersión sobre la media es decir un valor grande representa que la velocidad de los circuitos fabricados será repartida en un rango amplio mientras que un valor pequeño indicara que la velocidad de los circuitos estarán concentrados muy cerca de la media.

Observando los párrafos anteriores notamos que la selección de la F.O no es trivial y puede variar de acuerdo a la necesidad de diseño. Por ejemplo si un diseñador quisiera realizar el diseño más rápido permitido por la tecnología él podría pensar en minimizar μ_P sin embargo la distribución resultante podría ser muy dispersa debido a las variaciones de proceso. Por otro lado tal vez sería necesario para un diseñador que su circuito sea lo más robusto a las variaciones de proceso, por lo que σ_P podría ser una mejor opción de optimización ya que generaría una curva de distribución más fina pero mas lenta de forma nominal, combinaciones de ambas variables permiten ponderar estas cualidades.

Los casos discutidos con anterioridad (con algunas modificaciones) tienen aplicaciones en diferentes nichos de la industria de la electrónica digital, los diseñadores de circuitos FullCustom tienen por lo general como tarea sacar el mayor provecho a las tecnologías desarrolladas en cuanto a velocidad de procesamiento se refiere por lo que minimizar μ_P puede ser uno de los factores principalmente considerados en la etapa de diseño, por otro lado la industria de los ASIC'S tiene como prioridad generar una solución confiable a un cliente específico de forma rápida y robusta por lo que el diseño de un ASIC'S es repartido en módulos a diversas casas de diseño de IC con el fin de disminuir el tiempo al mercado, estos módulos son especificados en términos de área y velocidad de respuesta, para garantizar el correcto acoplo de los módulos los diseñadores ASIC'S tiene como principal objetivo hacer un diseño robusto a las variaciones de proceso, por lo cual el interés en la minimización de la desviación

estándar crece considerablemente en este nicho.

Funciones objetivo basadas en una variable de retardo

Las F.O propuestas para el análisis estadístico dadas por la literatura se muestran a continuación.

$$F.O \rightarrow \sigma_P$$

Esta función objetivo es dedicada a volver la curva gaussiana del retardo del camino lo más estrecha posible por lo que permite genera los caminos más robustos a las variaciones de proceso al ser minimizada, sin embargo bajo esta optimización no se tiene control directo sobre el valor de μ_P , aun así tanto μ_P como σ_P son inversamente proporcionales al área consumida como puede verse de un análisis analítico del retardo, por lo que es esperado como efecto secundario caminos más rápidos de forma nominal.

$$F.O \rightarrow \frac{\sigma_P}{\mu_P}$$

Esta F.O es conocida como el delay variability o coeficiente de variación y representa a σ_P como porcentaje de la media del retardo de propagación, es una medida estandarizada porcentual del grado de variación por lo que tiene un comportamiento similar a la F.O anterior, sin embargo al tener a μ_P en el denominador esta función converge de forma más rápida ya que μ_P se hace mas pequeño al crecer el área del camino, mientras que σ_P llega mas rapido a un punto de saturación por lo que F.O tiende a crecer después del punto donde σ_P empieza a saturarse evitando que la optimización crezca indeterminadamente ahorrando de esta forma área en el diseño, sin embargo como contra efecto es que al estar μ_P en el denominador el camino lógico siempre será más lento de forma nominal que si este se llevara a cabo con la optimización de σ_P

$$F.O \rightarrow \mu_P + 3\sigma_P$$

Esta función objetivo tiene como argumentos minimizar tanto σ_P como μ_P , el factor tres permite mantener un balance entre ambos parámetros, esta F.O es una medida directa del retardo máximo del camino, por lo que es común denominar a esta F.O como Worst Case.

Comparación de las F.O

Diseñar un circuito óptimo es un concepto naturalmente subjetivo pero a lo largo de este documento se entenderá como optimo el diseño que permita tener un mejor Yield, donde el Yield se define simplemente como

$$\text{Yield} = \frac{\text{Circuitos fabricados que cumplen la especificación}}{\text{Total de circuitos fabricados}} \quad (5.33)$$

El Yield también puede ser definido de forma equivalente como la probabilidad de que un circuito cumpla una especificación, llamando a una especificación de retardo como T_D y sabiendo que el retardo de un camino lógico tiene una distribución de probabilidad normal, se puede definir entonces:

$$\text{Yield} = P(D \leq T_D) = \int_{-\infty}^{T_D} \phi(D) dD \quad (5.34)$$

Donde $\phi(D)$ es la función de probabilidad del camino lógico, tal que

$$\phi(D) \sim N(\mu_P, \sigma_P) = \frac{1}{2\pi\sigma_P} \exp\left(\frac{-(D - D_P)^2}{2\sigma_P^2}\right) \quad (5.35)$$

Es importante recordar que toda distribución normal se puede normalizar y obtener la probabilidad bajo la curva en base a la tabla Z, alternatively se pueden utilizar software para calcular los valores de probabilidad.

Para comparar el resultado de dimensionar un circuito en base a las minimización propuesta de las diversas F.O se partirá de un caso básico de estudio para una cadena de 5 inversores, como el mostrado en la Figura 5.16.

Se empezara la comparación generando una optimización bajo una restricción de área que solo permita un incremento de 10 %, se analizaran resultados y posteriormente se incrementara la restricción para permitir un incremento del 100 % y finalmente limitando el camino solamente por los valores máximos de escalas en las compuertas que son alcanzables.

Todos los análisis partirán de dimensiones mínimas y bajo la condición que cada compuerta está separada una distancia de 100nm con respecto a sus compuertas adyacente, generando correlación con base al modelo exponencial [5].

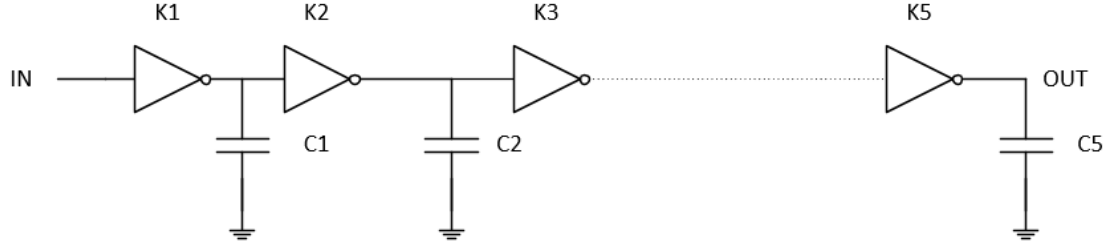


Figura 5.16: Ejemplo de camino lógico, donde $C_1 = 2\text{fF}$, $C_2 = 2\text{fF}$, $C_3 = 4\text{fF}$, $C_4 = 2\text{fF}$, $C_5 = 5\text{fF}$

Resultados de optimización con una restricción de incremento en el área de un 10 % sobre el área original

F.O	μ_P [ps]	σ_P [ps]
σ_P	247.87	12.09
σ_P/μ_P	248.10	12.10
$\mu_P + 3\sigma_P$	244.29	12.41

Tabla 5.6: Comparación de resultados

La comparación gráfica de las diferentes distribuciones de probabilidad obtenidas por cada optimización se muestra en la Figura 5.17. La integral de la figura anterior nos devuelve el Yield, el cual es mostrado en la Figura 5.18..

Nótese que tanto la F.O σ_P y σ_P/μ_P devuelven prácticamente los mismos valores, esta tendencia se mantiene en diferentes caminos siempre que la restricción de área permita solo incrementos pequeños en el camino final.

De los resultados anteriores notamos las siguientes observaciones.

- Las funciones objetivo σ_P y $\frac{\sigma_P}{\mu_P}$ devuelven prácticamente el mismo resultado esto es cierto siempre que el incremento de área permitido por la restricción es pequeño.
- La F.O $\mu_P + 3\sigma_P$ devuelve el mejor Yield, pero no el mejor valor de σ_P .

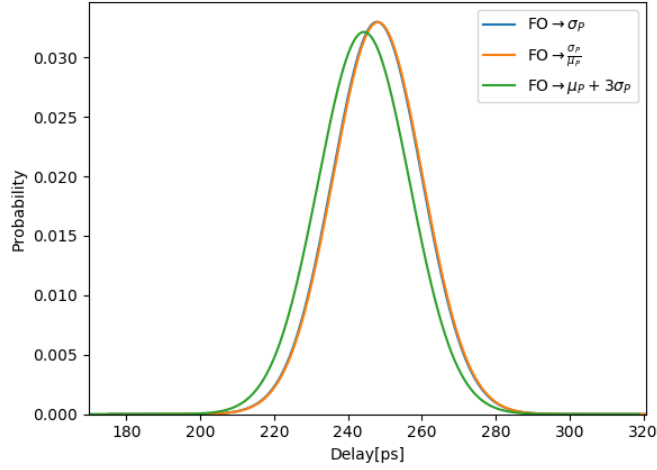


Figura 5.17: Comparación de resultados

Para explicar estos fenómenos recordemos que el algoritmo de LaGrange utiliza principalmente el gradiente de la F.O para moverse por el espacio de optimización, notando el primer resultado podemos ver que las funciones objetivos tienen derivadas que apuntan en la misma dirección, por lo que el ritmo de cambio en el cociente de $\frac{\sigma_P}{\mu_P}$ debe ser dominado por σ_P , comparado con el inverso del parámetro μ_P .

Sabiendo que el circuito original antes de la optimización presenta.

$$\mu_P = 263,08[\text{ps}] \text{ Y } \sigma_P = 14,37[\text{ps}]$$

La diferencia entre los valores de μ_P y σ_P antes y después de la optimización se muestran en la Tabla 5.7.

F.O	$\Delta\mu_P$ [ps]	$\Delta\sigma_P$ [ps]
σ_P	-15.21	-2.28
σ_P/μ_P	-14.96	-2.70
$\mu_P + 3\sigma_P$	-18.79	-1.96

Tabla 5.7: Cambios entre las variables de retardo para las optimizaciones bajo restricción de área $A_C = 1,1A_{P0}$

Viendo estos resultados es fácil destacar que μ_P tiene una derivada más negativa que σ_P por lo que este parámetro decrece más rápido con respecto a las variables

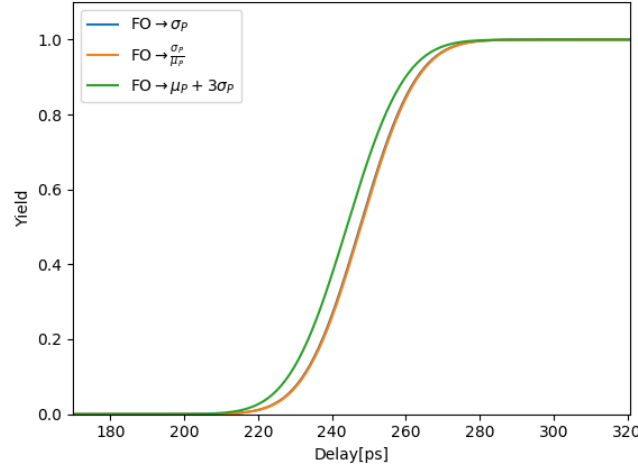


Figura 5.18: Yield devuelto por las diferentes funciones objetivos bajo un incremento de área pequeño.

de decisión, sin embargo en la F.O σ_P/μ_P no se trabaja directamente con μ_P si no con el recíproco de μ_P , tal que.

$$\Delta \frac{1}{D_{P,F.O \rightarrow \frac{\sigma_P}{\mu_P}}} = -0,053[1/\text{ps}] \quad (5.36)$$

Como vemos el ritmo de cambio de este factor es muy pequeño por lo que en la F.O $\frac{\sigma_P}{\mu_P}$ el término $\frac{1}{\mu_P}$ parece casi constante, lo cual explica el comportamiento similar entre las dos primeras F.O, en general examinado la fórmula de la derivada sabemos que si una función y cambia rápidamente en una dirección su recíproco $1/y$ (no confundir con función inversa) cambiara lentamente en la otra dirección. Cabe destacar que tanto σ_P como μ_P son funciones que tienden a saturarse a partir un punto, por lo que en algún punto en la función $\frac{\sigma_P}{\mu_P}$ el término $1/\mu_P$ será dominante tal que la F.O crecerá con respecto a las variables de decisión, terminando la minimización, esto puede traer como consecuencia que el algoritmo pare antes de asignar completamente los recursos disponibles a las compuertas debido al punto de inflexión donde la función objetivo empieza a crecer.

La discordancia entre el comportamiento de la F.O $\mu_P + 3\sigma_P$ es debido simplemente a que:

$$-\frac{\partial \mu_P}{\partial K} > -\frac{\partial \sigma_P}{\partial K} \quad (5.37)$$

Por lo el algoritmo de LaGrange empieza a moverse principalmente por el camino de la derivada del μ_P , pero ponderando el valor del σ_P en una cantidad no despreciable.

Resultados de optimización con restricción que permite un incremento del 100 % en el área original

F.O	μ_P [ps]	σ_P [ps]
σ_P	177.75	6.82
σ_P/μ_P	184.02	6.97
$\mu_P + 3\sigma_P$	175.21	6.96

Tabla 5.8: Comparación de resultados

La comparación gráfica de las diferentes distribuciones de probabilidad se muestra a continuación.

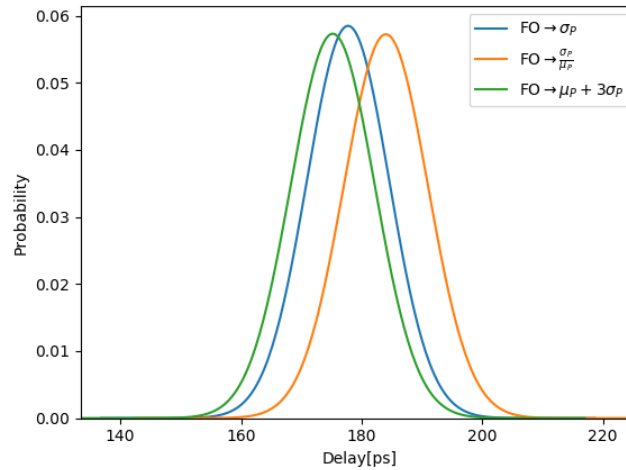


Figura 5.19: Histograma para optimización con incremento de área considerable

Calculando el Yield atreves de la Figura 5.19.

De los resultados anteriores vemos que:

- Las funciones objetivo σ_D y $\frac{\sigma_P}{\mu_P}$ ya no devuelven resultados idénticos como se tenía bajo incrementos pequeños de área, esto es debido a que en un

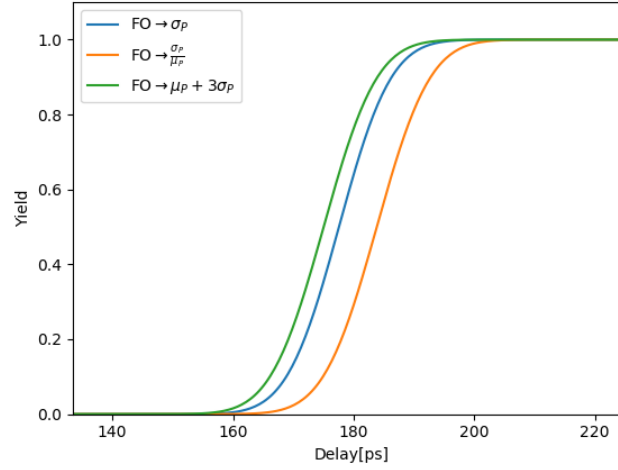


Figura 5.20: Yield para optimización con incremento de área considerable

punto de la optimización el factor $1/\mu_P$ empieza a ser dominante por lo que el algoritmo empieza a seguir el camino que no minimice el valor de μ_P como fue discutido en la sección anterior, esto se refleja que de las tres optimizaciones la F.O $\frac{\sigma_P}{\mu_P}$ devuelve el mayor valor de μ_P comparado con el resto de las F.O, esto es una tendencia encontrada para este tipo de análisis.

- La función objetivo $\mu_P + 3\sigma_P$ nuevamente devuelve el circuito con el mejor Yield.

Resultados de optimización sin restricción de área

F.O	μ_P [ps]	σ_P [ps]
σ_P	81.56	2.84
σ_P/μ_P	112.34	3.99
$\mu_P + 3\sigma_P$	81.56	2.84

Tabla 5.9: Comparación de resultados

El criterio de paro la F.O σ_P y $\mu_P + 3\sigma_P$ fue el hecho que el valor máximo de K para una compuerta es 10 ya que este fue el limite de la caracterización, por su parte la F.O $\frac{\sigma_P}{\mu_P}$ termino la optimización usando un menor porcentaje de área que este máximo permitido.

La comparación gráfica de las diferentes distribuciones de probabilidad se muestra a continuación.

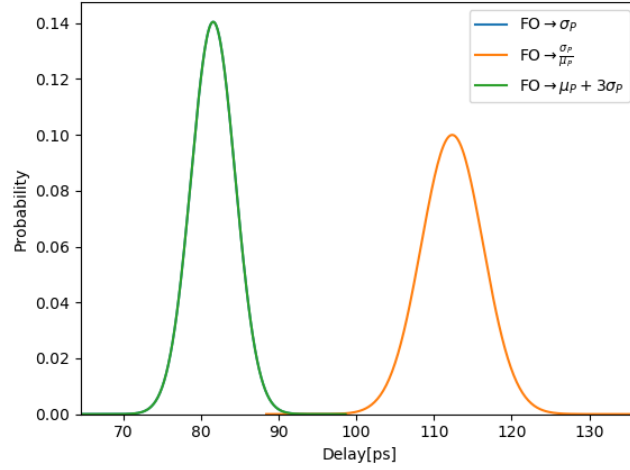


Figura 5.21: Histograma para optimización sin restricciones

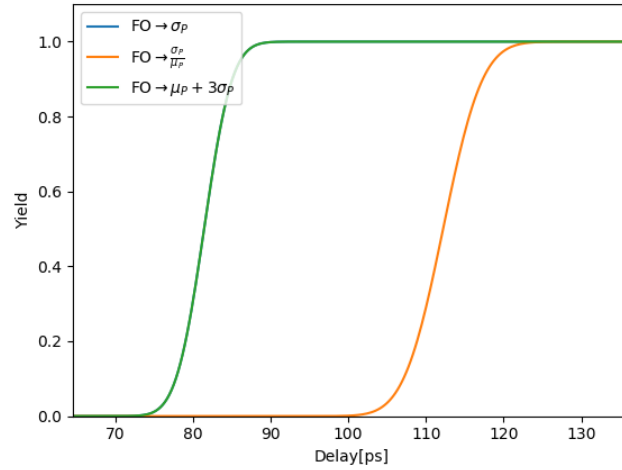


Figura 5.22: Yield para optimización sin restricción

En este análisis cabe destacar que la función $\frac{\sigma_P}{\mu_P}$ utilizó un área mucho menor que el resto de las F.O.

5.5.2. Optimización de caminos lógicos minimizando el área

Este tipo de optimización es mas directa de abordar que su contra parte basada en una variable del retardo ya que la F.O esta completamente definida como el área del camino lógico.

Debido a que una minimización de área sin restricción devuelve una respuesta trivial (el camino mas pequeño que puede fabricarse) esta optimización siempre debe ser acompañada de una restricción temporal la cual puede ser descrita con base a las funciones objetivo estudiadas en la sección anterior, se nota sin embargo que dar una restricción basada únicamente en el σ_P no es una opción practica ya que no se sabe nada del retardo nominal solo se pide que este tenga una distribución con una dispersión sobre la media basada en un intervalo definido, esto claramente es confuso y poco practico para un objetivo de diseño, por otro lado el delay variability genera un problema similar solo que ahora midiendo el intervalo en unidades porcentuales, la única variable de retardo que funciona como una restricción practica es $\mu_P + 3\sigma_P$ ya que esta restricción se puede interpretar como el retardo máximo del camino permitido, es decir el problema de optimización completo tratara de buscar un diseño que posea un mínimo uso de área para alcanzar un retardo máximo permitido por el camino lógico.

Debido a lo discutido en el párrafo anterior la optimización para minimización de área tomara como F.O el área del camino y dicha optimización siempre estará acompañada de una restricción del tipo $\mu_P + 3\sigma_P = T_D$, donde el parámetro T_D se denomina Target de diseño y corresponde con el peor retardo que tendrá el camino cuando es sujeto a variaciones de proceso.

Aunque sencilla la optimización anterior tiene un gran potencial de aplicación sobretodo en el diseño de ASIC'S donde el sistema a diseñar se divide en módulos que desean diseñarse para cumplir una especificación de retardo.

Aparte de esto circuitos con un menor uso de área generalmente consumen menos potencia eléctrica de la alimentación al conmutar.

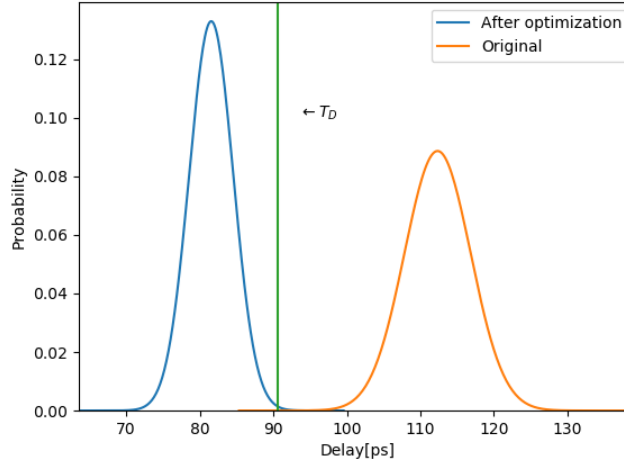


Figura 5.23: Ejemplo de optimización minimizando el área

5.6. Método Heurístico

Debido a que la metodología completa de los multiplicadores de LaGrange requiere un alto costo [11] computacional, han surgido diversas alternativas para obtener soluciones quasi-optimas fáciles de computar, dichos métodos son conocidos como heurísticos y se basan en el cálculo de una métrica para las compuertas lógicas de un camino, esta métrica comúnmente mide el impacto en el σ_P de redimensionar la compuerta evaluada.

El funcionamiento básico de los métodos heurísticos en el diseño es definir una métrica que permita seleccionar las compuertas dentro de un camino lógico cuyo re-dimensionamiento tenga un mayor impacto en la reducción de la variación del retardo total del camino lógico.

Para este proceso muchas métricas se han sido propuestas, en el trabajo [15] se utiliza la derivada de la desviación estándar del retardo de la compuerta con respecto al ancho de canal, esta métrica fue utilizada en la sección 3.3 del capítulo 2 y se denominara como la métrica completa, sin embargo dicha métrica es costosa de computar ya que toma en cuenta toda la correlación entre las compuertas del camino lógico lo cual requiere un total de N^N cálculos por cada iteración de evaluación en el camino lo cual es muy costoso si se utilizan modelos complejos como los polinomios desarrollados en este capítulo, para contrarrestar esto otras métricas se han propuesto en la literatura, una de éstas métricas fue propuesta en el trabajo [11], en dicho trabajo se utiliza el concepto de segmento de un

camino lógico lo cual se refiere a todo posible conjunto ternario de compuertas lógicas adyacentes que pueden formarse, un camino lógico de N etapas tendrá N segmentos y por ende solo necesitará N evaluaciones por iteración.

El concepto de segmento se basa en que la mayor contribución de la correlación espacial y de señal se da en las dos compuertas adyacentes respecto a la compuerta evaluada, el desarrollo matemático de la métrica completa podrá encontrarse en el trabajo [11], la ecuación final se muestra a continuación:

$$M_i = (\sigma_{D_{i-1},V} + \sigma_{D_i,SRI}) \frac{\partial \sigma_{D_{i-1},V}}{\partial K_i} + (\sigma_{D_i,V} + \sigma_{D_{i+1},SRI}) \frac{\partial \sigma_{D_i,V}}{\partial K_i} + \sigma_{D_{i+1},V} \frac{\partial \sigma_{D_{i+1},V}}{\partial K_i} + \sigma_{D_{i-1},V} \frac{\partial \sigma_{D_i,SRI}}{\partial K_i} + \sigma_{D_i,V} \frac{\partial \sigma_{D_{i+1},SRI}}{\partial K_i} \quad (5.38)$$

Ejemplos de aplicación de la métrica se puede encontrar en el trabajo [11]. Una de las principales características del método heurístico es que va a dimensionar una compuerta por iteración, lo cual puede aprovecharse para definir un tipo de optimización donde solo se modifiquen un numero pequeño de compuertas, esto es recomendado por que en un camino lógico ya sintetizado el modificar solo algunas de las compuertas en contra medida de dimensionar todo el camino permitirá a la optimización ser mas atractiva ya que el re trabajo de ajustar el resto del circuito una vez optimizado el camino es menor.

El diagrama de flujo general del método Heurístico se muestra en la Figura 5.24.

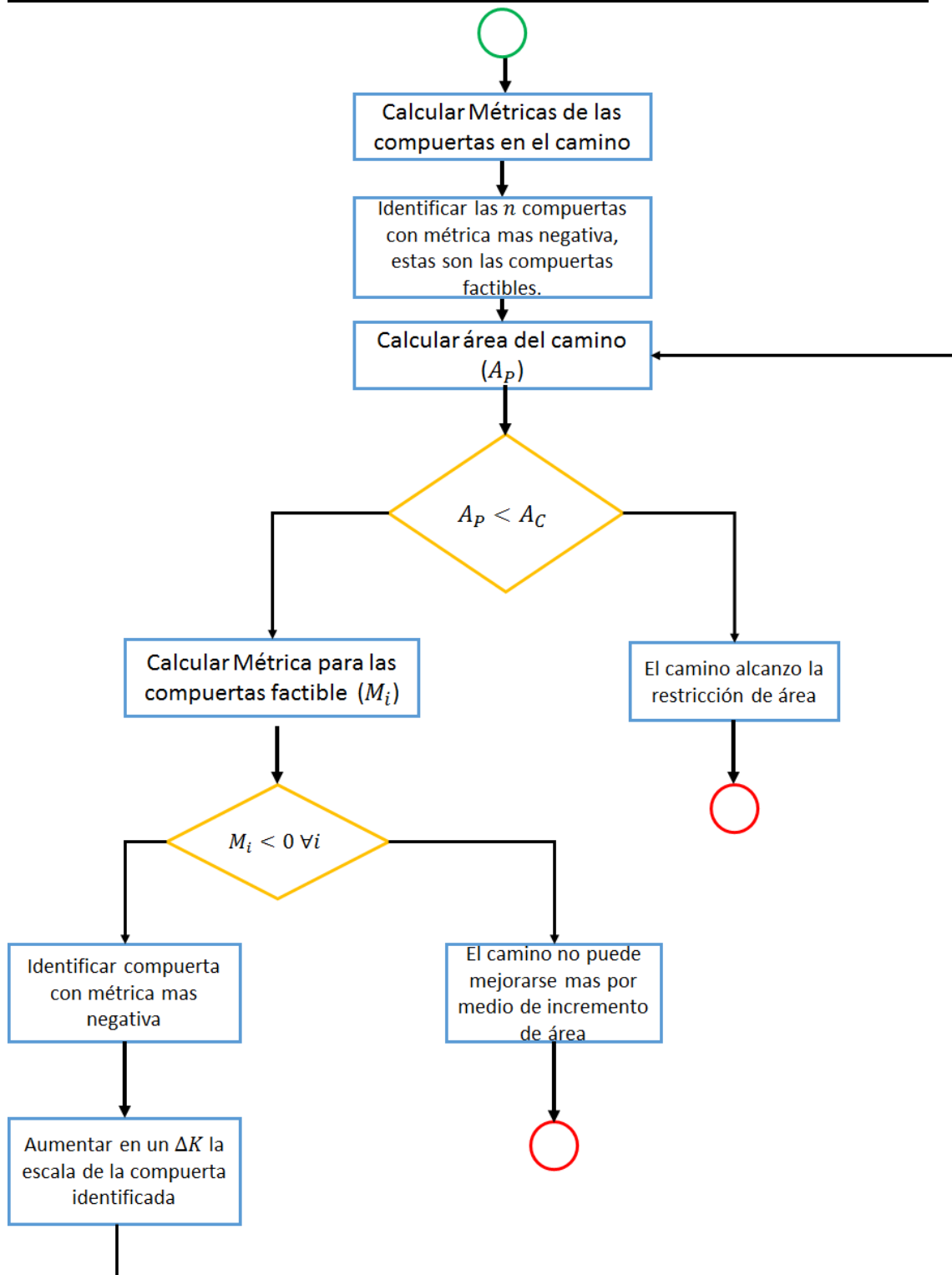


Figura 5.24: Diagrama de flujo para optimización heurística

La comparación del desempeño de la métrica con el método LaGrange se muestra en la Tabla 5.10.

Característica	LaGrange	Heurístico
Tiempo de Computo	Alto costo computacional requiere el calculo del gradiente de la F.O a optimizar.	Costo computacional bajo, basado en la evaluación de una formula pre-establecida para cada compuerta en el camino.
Correlación espacial	Incluye todos los efectos de la correlación espacial para la optimización.	Solamente moldea los efectos de correlación entre las compuertas pertenecientes al segmento.
Variaciones consideradas	Todas	Unicamente variaciones en el V_T y en el SR_I .
Tipo de resultados	Resultados optimos	Resultados quasi-optimos
Principal aplicación	Circuitos pequeños y sumamente criticos	Versatil en la mayoría de las aplicaciones

Tabla 5.10: Comparación de resultados

5.7. Identificación de caminos lógicos

Si bien la herramienta SSTA estará enfocada en optimizar una de las diferentes funciones objetivos descritas anteriormente para un camino lógico, un sistema digital real esta formado por múltiples caminos lógicos los cuales deben ser pre-identificados para poder llevar acabo el proceso de optimización.

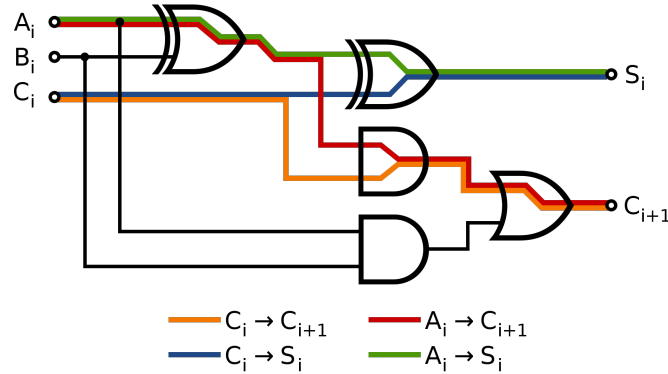


Figura 5.25: Ejemplo de un circuito digital y algunos de sus caminos lógicos

Los caminos lógicos son fáciles de identificar por medio de algoritmos si se plantea al circuito por su grafo equivalente, debido a la falta de retroalimentación y a que cada nodo de un circuito no puede ser conducido por mas de una compuerta (de otra forma se genera una situación de contención lógica) se puede demostrar que un camino lógico tiene un grafo equivalente en forma de árbol donde cada nodo del circuito es un nodo del grafo y cada compuerta lógica forma una arista que une sus nodos de entrada con la salida de la compuerta, la información temporal del retardo puede indicarse como la ponderación de la arista que se interpreta como el costo (temporal) de cruzar del nodo de entrada a la salida, por ultimo el grafo de un camino lógico es siempre dirigido ya que son las entradas las que conducen la salida y nunca la situación contraria.

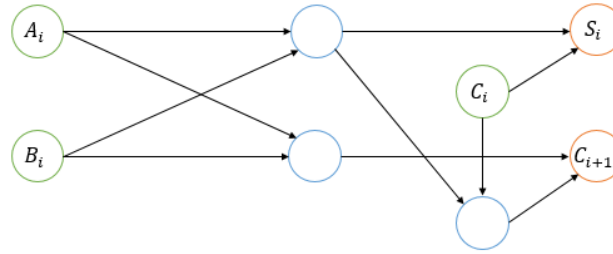


Figura 5.26: Grafo dual del circuito digital anterior

Para trabajar con grafos en Python se puede usar la librería terciaria NetworkX, la cual pertenece a la comunidad de software libre y esta especializada en su uso para la investigación científica de redes gráficas complejas.

Las subrutinas mas importantes que sera usadas de esta librería dentro de la herramienta SSTA son:

- *allsimplepaths*($G, N1, N2$): Función especializada para encontrar todos los caminos del grafo G que relacionen a los nodos $N1$ a $N2$, la subrutina esta basada en el clásico algoritmo Depth-first-search para identificar los caminos. Este algoritmo esta caracterizado por tener un tiempo computacional $O(N!(V + E))$ para encontrar todos los caminos de $N1$ a $N2$ por lo cual solo es adecuado si el numero de nodos N del grafo es bajo, como la función debe ser evaluada para cada par de relación entre entradas y salidas el identificar todos los caminos en diseños VLSI complejos es inviable.
- *dijkstra-pathlength*($G, N1, N2$): Función utilizada para detectar el camino mas corto (de menor ponderación) en el grafo G entre el nodo $N1$ y $N2$, si se quiere identificar el camino mas largo se puede trabajar con el inverso

del peso de las aristas y utilizar el mismo algoritmo, ya que esta especializado en encontrar un solo camino (el mas corto) el tiempo computacional de este algoritmo es mucho menor al de detectar todos los caminos lógicos y corresponde con un tiempo computacional $O(E + V \log(V))$, al ser dicho costo relativamente bajo este algoritmo puede usarse aun en diseños complejos para identificar un conjunto de caminos críticos.

En general cuando en la herramienta SSTA el numero de nodos del grafo sea bajo ($N < 16$) se utilizara la subrutina *allsimplepaths* para mostrar y permitir seleccionar cualquiera de los posibles caminos en el circuito, mientras que para un numero de nodos altos se usara la subrutina *dijkstra pathlength* por cuestiones de mantener el tiempo de identificación de caminos en valores razonables.

Debido a que como se menciona en el grafo dual cada compuerta corresponde con una arista la ponderación de dichas aristas es el retardo de las compuertas lógicas dicho retardo sera calculado para esta sección de clasificación y identificación de caminos con el modelo analítico para reducir el costo computacional, una vez identificado el camino a optimizar se utilizara el modelo de polinomios desarrollado en esta sección.

5.8. Conclusiones

En este capítulo se propuso y desarrolló una herramienta computacional SSTA, la idea de esta herramienta nació al buscar una alternativa ligera pero más precisa que el modelo analítico para la predicción del comportamiento estadístico del retardo en los caminos lógicos esto debido a que los clásicos métodos MonteCarlo son demasiado costosos en términos computacionales para la mayoría de las aplicaciones requeridas por la industria VLSI, para lograr su cometido la herramienta utiliza el modelo estadístico del retardo visto en capítulos anteriores pero utilizando parámetros precisamente pre-caracterizados por medio de un número exhaustivo de simulaciones (18 000 por compuerta lógica), estas pre-caracterizaciones permiten tener márgenes de error tan bajos como un $2 \sim 3\%$ como fue visualizado durante la validación.

Una vez lograda la alta habilidad de predicción por el modelo polinomial se implementó en la herramienta algoritmos de optimización tanto exactos como heurísticos bajo un conjunto de funciones objetivo propuestas en el estado del arte de la investigación para el área de EDA. Un análisis sobre cada función

objetivo así como la comparación entre todas estas fue dado de forma detallada durante el desarrollo del capítulo.

Aparte de lo anterior diversas técnicas de programación orientada objetos (POO) y librerías estándar de software fueron utilizadas en la herramienta no solo para implementar sus caracterizaras funcionales sino también para generar una GUI amigable e intuitiva al usuario.

Un manual completo sobre el uso de la herramienta se puede encontrar en el Apéndice [C.1.2](#).

Capítulo 6

Conclusiones

A lo largo del trabajo se estudió el efecto y modelado de las variaciones de proceso en los sistemas digitales, en especial considerando variaciones correlacionadas espacialmente en el ancho y largo de canal así como en el espesor del oxido de los transistores, también se consideró como una variación independiente el valor del voltaje umbral.

Los modelos analíticos que moldean el impacto de estas variaciones en el retardo fueron desarrollados a partir de ver a las compuertas CMOS por medio de una red RC distribuida, dichos modelos están especificados en términos del parámetro de escala de la compuerta (K) y permiten computar no solo el valor nominal del retardo nominal (μ_D) sino también su desviación estándar (σ_D). Posteriormente se utilizó dicho modelo para encontrar el comportamiento del retardo a nivel camino lógico.

De manera alternativa se presentó una metodología para obtener funciones polinómicas que molden el impacto de las variaciones en el retardo por medio de un simulador eléctrico, estos modelos son extraordinariamente precisos pero mucho más complejos que el modelado analítico anteriormente discutido, los modelos polinomicos fueron incluidos dentro de una herramienta de software desarrollada para computar el valor del retardo estadístico (valor nominal y desviación estándar) de los caminos lógicos.

Si bien los modelos analíticos son menos precisos que las funciones polinómicas estos permiten observar el impacto directo de las variables de diseño en una expresión compacta, esto fue aprovechado para estudiar a las funciones a través de sus derivadas las cuales definen métricas que pueden usarse junto a métodos

de optimización heurísticos para obtener metodologías de optimización de bajo costo computacional la cual fue presentada, los modelos analíticos también pueden ser usados por los diseñadores de IC para obtener un insight del efecto de las variaciones de proceso en el retardo de los sistemas.

Por su parte la herramienta de software utiliza los polinomios junto a sub-rutinas de optimización como el método de LaGrange para generar el dimensionamiento óptimo de las compuertas en un camino lógico, dicho dimensionamiento puede realizarse para minimizar el área consumida respetando un target de retardo o alternativamente minimizar una variable de retardo sujeta a un consumo de área máximo, los resultados de la herramienta pueden ser exportados o visualizados en tiempo real de uso todo a través de una interfaz gráfica de usuario.

Durante el trabajo se mostró la validación y precisión tanto del modelo analítico como del polinómico.

En el trabajo también se proporcionó para la herramienta de software anteriormente discutida una guía sobre su uso, prestaciones y características importantes. Un manual completo se muestra en el Apéndice [C.1.2](#), mientras que el funcionamiento interno de los algoritmos fue discutido durante el cuerpo del trabajo de tesis.

Se espera que el trabajo aquí desarrollado sea una base para el grupo de trabajo que permita expandir todo discutido a aplicaciones directas con la industria.

Apéndice A

A.1. Demostración de la Ecuación 2.17

Partiendo de:

$$C_{self} = \frac{1}{2}(\mathfrak{o}_n C_{G_n} + \mathfrak{o}_p C_{G_p}) + (\mathfrak{o}_n C_{DB0_n} + \mathfrak{o}_p C_{DB0_p}) \quad (A.1)$$

En base a la ecuación 2.6 se puede escribir el primer termino de la ecuación A.1 como:

$$\frac{1}{2}(\mathfrak{o}_n C_{G_n} + \mathfrak{o}_p C_{G_p}) = \frac{1}{2}(\mathfrak{o}_n C_{ox}(W_n L_n) + \mathfrak{o}_p C_{ox}(W_p L_p)) \quad (A.2)$$

Utilizando las condiciones de dimensionamiento discutidas para las compuertas digitales.

$$\frac{1}{2}(\mathfrak{o}_n C_{G_n} + \mathfrak{o}_p C_{G_p}) = \frac{1}{2}(\mathfrak{o}_n C_{ox} F D_n K W_{min} L_{min} + \mathfrak{o}_p C_{ox} F D_p K W_{min} L_{min}) \quad (A.3)$$

Por lo tanto

$$\frac{1}{2}(\mathfrak{o}_n C_{G_n} + \mathfrak{o}_p C_{G_p}) = \frac{1}{2}(C_{ox} K W_{min} L_{min})(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) \quad (A.4)$$

Mientras que para el segundo termino sustituyendo la ecuación 2.15.

$$(\mathfrak{o}_n C_{DB0_n} + \mathfrak{o}_p C_{DB0_p}) = \mathfrak{o}_n (C_{j0} A_{D_n} + C_{jsw} P_{D_n}) + \mathfrak{o}_p (C_{j0} A_{D_p} + C_{jsw} P_{D_p}) \quad (A.5)$$

Acomodando términos.

$$(\mathfrak{o}_n C_{DB0_n} + \mathfrak{o}_p C_{DB0_p}) = C_{j0}(\mathfrak{o}_n A_{D_n} + \mathfrak{o}_p A_{D_p}) + C_{jsw}(\mathfrak{o}_n P_{D_n} + \mathfrak{o}_p P_{D_p}) \quad (A.6)$$

Se sustituirá en la ecuación anterior las expresiones de área y perímetro mostradas en 2.13 y 2.14, pero antes se empieza por examinar la componente de área usando las condiciones encontradas en el dimensionamiento de compuertas digitales.

$$A_D = 2,5LW = 2,5L_{min}(W_{min}FDK) \quad (A.7)$$

Por lo que usando dicha condiciones mas una factorización de términos comunes se obtiene que:

$$C_{j0}(\mathfrak{o}_n A_{D_n} + \mathfrak{o}_p A_{D_p}) = C_{j0}(2,5L_{min}KW_{min})(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) \quad (\text{A.8})$$

Por su parte examinando el perímetro con las condiciones de dimensionamiento de compuertas digitales.

$$P_D = 2W + 5L = 2KW_{min}FD + 5L_{min} \quad (\text{A.9})$$

Tal que sustituyendo y factorizando:

$$C_{jsw}(\mathfrak{o}_n P_{D_n} + \mathfrak{o}_p P_{D_p}) = C_{jsw}((2KW_{min})(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + 5L_{min}(n_n + n_p)) \quad (\text{A.10})$$

Aplicando los resultados anteriores en la expresión completa de la auto capacitancia tenemos:

$$\begin{aligned} C_{self} = & 0,5(C_{ox}KW_{min}L_{min})(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + \\ & C_{j0}(2,5L_{min}KW_{min})(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + \\ & C_{jsw}((2KW_{min})(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + 5L_{min}(\mathfrak{o}_n + \mathfrak{o}_p)) \end{aligned} \quad (\text{A.11})$$

Agrupando las constantes dependientes de la tecnología:

$$a_1 = 0,5(C_{ox}W_{min}L_{min}) \quad (\text{A.12})$$

$$a_2 = C_{j0}(2,5L_{min}W_{min}) \quad (\text{A.13})$$

$$a_3 = C_{jsw}(2W_{min}) \quad (\text{A.14})$$

$$B = C_{jsw}5L_{min} \quad (\text{A.15})$$

Se puede expresar

$$C_{self} = a_1(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + a_2(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + a_3(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + B(n_n + n_p) \quad (\text{A.16})$$

Claramente

a_1 : Corresponde con la capacitancia generada debido al acople capacitivo del canal con la compuerta del lado del drenaje, en un transistor de dimensiones mínimas.

a_2 : Corresponde con la capacitancia parásita generada por la componente de área del drenaje para un transistor de dimensiones mínimas.

a_3 : Corresponde con la capacitancia parásita generada por la componente del ancho del drenaje para un transistor de dimensiones mínimas.

B : Corresponde con la capacitancia parásita generada por la componente del largo del drenaje para un transistor de dimensiones mínimas.

Agrupando terminos la expresión resultante obtiene la forma de la ecuación 2.17.

$$C_{self} = A(\mathfrak{o}_n F D_n + \mathfrak{o}_p F D_p) + B(n_n + n_p) \quad (\text{A.17})$$

Donde

$$A = a_1 + a_2 + a_3 \quad (\text{A.18})$$

Q.E.D.

A.2. Demostración de la Ecuación 2.101

Sabiendo que

$$\text{COV}(D_i, D_j) = \sum_{k=1}^{\chi} \rho_{i,j} \times (S_{P_k,i}^D \sigma_{P_k,i}) \times (S_{P_k,j}^D \sigma_{P_k,j}) \quad (\text{A.19})$$

Al expandir en términos de las variaciones de proceso consideradas y negando las contribuciones de los nodos intermedios.

$$\begin{aligned} \text{COV}(D_i, D_j) = & \rho_{V_{T_i,j}} \times \left(\sqrt{\Lambda_i} S_{V_{T_i}}^D \sigma_{V_{T_i}} \times \sqrt{\Lambda_j} S_{V_{T_j}}^D \sigma_{V_{T_j}} \right) + \rho_{W_{i,j}} \times \left(\sqrt{\Upsilon_i} S_{W_i}^D \sigma_{W_i} \times \sqrt{\Upsilon_j} S_{W_j}^D \sigma_{W_j} \right) + \\ & \rho_{L_{i,j}} \times \left(\sqrt{\Upsilon_i} S_{L_i}^D \sigma_{L_i} \times \sqrt{\Upsilon_j} S_{L_j}^D \sigma_{L_j} \right) + \rho_{T_{ox,i,j}} \times \left(\sqrt{\Upsilon_i} S_{T_{ox_i}}^D \sigma_{T_{ox_i}} \times \sqrt{\Upsilon_j} S_{T_{ox_j}}^D \sigma_{T_{ox_j}} \right) \end{aligned} \quad (\text{A.20})$$

Donde

$$\Lambda_x = \sum_{x=1}^{\chi} \xi_x^2 \quad (\text{A.21})$$

$$\Upsilon_x = \sum_{x=1}^{\chi} \xi_x^2 + 2 \sum_{x=1}^{\chi} \sum_{x_2=x+1}^{\chi} \xi_x \xi_{x_2} \quad (\text{A.22})$$

Identificando que

$$S_{P_i}^D \sigma_{P_i} = \sigma_{D_{P_i}} \quad (\text{A.23})$$

Se reconoce que la expresión solicitada ya fue calculada para cada variación de proceso en las ecuaciones 2.66 a 2.69, notando que entre una compuerta i y una compuerta j los únicos cambios que pueden a ver en dichas ecuaciones son:

- La escala de la compuerta (K)
- La carga de la compuerta (C_L)
- La topología (FD)

Por lo cual sustituyendo se tiene que:

$$\begin{aligned}
 \text{COV}(D_i, D_j) = & \\
 & \rho_{V_{T_{i,j}}} \times \left(\left[\frac{\kappa \sqrt{\Lambda_i}}{\beta_{0i}} \times \frac{C_{L_i}}{K_i} \times \frac{\delta_{V_T}}{\sqrt{K_i} \times F D_i^{1,5}} \right] \left[\frac{\kappa \sqrt{\Lambda_j}}{\beta_{0j}} \times \frac{C_{L_j}}{K_j} \times \frac{\delta_{V_T}}{\sqrt{K_j} \times F D_j^{1,5}} \right] \right) + \\
 & \rho_{W_{i,j}} \times \left(\left[\sqrt{\Upsilon_i} \frac{\kappa}{\beta_{0i}} \times \frac{C_{L_i}}{K_i} \times \frac{\delta_W}{F D_i^2 K_i} \right] \left[\sqrt{\Upsilon_j} \frac{\kappa}{\beta_{0j}} \times \frac{C_{L_j}}{K_j} \times \frac{\delta_W}{F D_j^2 K_j} \right] \right) + \\
 & \rho_{L_{i,j}} \times \left(\left[\sqrt{\Upsilon_i} \frac{\kappa}{\beta_{0i}} \times \frac{C_{L_i}}{K_i} \times \frac{\delta_L}{F D_i} \right] \left[\sqrt{\Upsilon_j} \frac{\kappa}{\beta_{0j}} \times \frac{C_{L_j}}{K_j} \times \frac{\delta_L}{F D_j} \right] \right) + \\
 & \rho_{T_{ox_{i,j}}} \times \left(\left[\sqrt{\Upsilon_i} \frac{\kappa}{\beta_{0i}} \times \frac{C_{L_i}}{K_i} \times \frac{\delta_{T_{ox}}}{F D_i} \right] \left[\sqrt{\Upsilon_j} \frac{\kappa}{\beta_{0j}} \times \frac{C_{L_j}}{K_j} \times \frac{\delta_{T_{ox}}}{F D_j} \right] \right) \quad (\text{A.24})
 \end{aligned}$$

Factorizando

$$\begin{aligned}
 \text{COV}(D_i, D_j) = & \left[\left(\frac{\kappa^2}{\beta_{0i} \beta_{0j}} \right) \left(\frac{C_{L_i}}{K_i} \frac{C_{L_j}}{K_j} \right) \right] \times \\
 & \left[\rho_{V_{T_{i,j}}} \left(\frac{\sqrt{\Lambda_i \Lambda_j} \delta_{V_T^2}}{\sqrt{(K_i)(K_j)(F D_i F D_j)^{1,5}}} \right) + \right. \\
 & \left. \frac{\sqrt{\Upsilon_i \Upsilon_j}}{F D_i F D_j} \left(\rho_{W_{i,j}} \left(\frac{\delta_W^2}{K_i K_j F D_i F D_j} \right) + \rho_{L_{i,j}} \delta_L^2 + \rho_{T_{ox_{i,j}}} \delta_{T_{ox}}^2 \right) \right] \quad (\text{A.25})
 \end{aligned}$$

Q.E.D.

A.3. Inferencia de la Ecuación 3.1

Recordando las expresiones 2.66 a 2.69

$$\sigma_{D_{V_T}} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K} \times \frac{\delta_{V_T}}{\sqrt{K}FD^{1,5}} \quad (\text{A.26})$$

$$\sigma_{D_W} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K^2} \times \frac{\delta_W}{FD^2} \quad (\text{A.27})$$

$$\sigma_{D_L} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K} \times \frac{\delta_L}{FD} \quad (\text{A.28})$$

$$\sigma_{D_{T_{ox}}} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K} \times \frac{\delta_{T_{ox}}}{FD} \quad (\text{A.29})$$

Al derivar dichas expresiones se obtiene que

$$\frac{\partial \sigma_{D_{V_T}}}{\partial K} = \frac{\kappa}{\beta_0} \frac{C_L}{FD^{1,5}} \frac{\partial \left(\frac{1}{K} \frac{1}{\sqrt{K}} \right)}{\partial K} = -1,5 \frac{\kappa}{\beta_0} \frac{C_L}{FD^{1,5}} \left(\frac{1}{K^{2,5}} \right) \quad (\text{A.30})$$

$$\frac{\partial \sigma_{D_W}}{\partial K} = -\frac{\kappa}{\beta_0} \frac{C_L \delta_W}{FD^2} \frac{\partial \left(\frac{1}{K^2} \right)}{\partial K} = -2 \frac{\kappa}{\beta_0} \frac{C_L \delta_W}{FD^2} \left(\frac{1}{K^3} \right) \quad (\text{A.31})$$

$$\frac{\partial \sigma_{D_L}}{\partial K} = \frac{\kappa}{\beta_0} \frac{\delta_L C_L}{FD} \frac{\partial \left(\frac{1}{K} \right)}{\partial K} = -\frac{\kappa}{\beta_0} \frac{\delta_L C_L}{FD} \left(\frac{1}{K^2} \right) \quad (\text{A.32})$$

$$\frac{\partial \sigma_{D_{T_{ox}}}}{\partial K} = \frac{\kappa}{\beta_0} \frac{\delta_{T_{ox}} C_L}{FD} \frac{\partial \left(\frac{1}{K} \right)}{\partial K} = -\frac{\kappa}{\beta_0} \frac{\delta_{T_{ox}} C_L}{FD} \left(\frac{1}{K^2} \right) \quad (\text{A.33})$$

Sabiendo que

$$\sigma_D = \sqrt{\Lambda_i \sigma_{D_{V_T}}^2 + \Upsilon_i \left(\sigma_{D_W}^2 + \sigma_{D_L}^2 + \sigma_{D_{T_{ox}}}^2 \right)} \quad (\text{A.34})$$

Se puede expresar la derivada de dicha ecuación por la regla de la cadena.

$$\frac{\partial \sigma_D}{\partial K} = -\frac{2\Lambda_i \sigma_{D_{V_T}} \frac{\partial \sigma_{D_{V_T}}}{\partial K} + \Upsilon_i \left(2\sigma_{D_W} \frac{\partial \sigma_{D_W}}{\partial K} + 2\sigma_{D_L} \frac{\partial \sigma_{D_L}}{\partial K} + 2\sigma_{D_{T_{ox}}} \frac{\partial \sigma_{D_{T_{ox}}}}{\partial K} \right)}{2\sqrt{\Lambda_i \left(\sigma_{D_{V_T}} \right)^2 + \Upsilon_i \left(\left(\sigma_{D_W} \right)^2 + \left(\sigma_{D_L} \right)^2 + \left(\sigma_{D_{T_{ox}}} \right)^2 \right)}} \quad (\text{A.35})$$

Simplificando

$$\frac{\partial \sigma_D}{\partial K} = -\frac{\Lambda_i \sigma_{D_{V_T}} \frac{\partial \sigma_{D_{V_T}}}{\partial K} + \Upsilon_i \left(\sigma_{D_W} \frac{\partial \sigma_{D_W}}{\partial K} + \sigma_{D_L} \frac{\partial \sigma_{D_L}}{\partial K} + \sigma_{D_{T_{ox}}} \frac{\partial \sigma_{D_{T_{ox}}}}{\partial K} \right)}{\sqrt{\Lambda_i \left(\sigma_{D_{V_T}} \right)^2 + \Upsilon_i \left(\left(\sigma_{D_W} \right)^2 + \left(\sigma_{D_L} \right)^2 + \left(\sigma_{D_{T_{ox}}} \right)^2 \right)}} \quad (\text{A.36})$$

Q.E.D.

A.4. Demostración de la Ecuación 3.8

Recordando que:

$$\sigma_{D_{i-1}} = \left(\frac{\kappa}{\beta_{0_{i-1}}} \frac{C_{L_{i-1}}}{K_{i-1}} \right) \sqrt{\Lambda_{i-1} \left(\frac{\delta_{V_T}}{\sqrt{K_{i-1}} \times FD_{i-1}^{1,5}} \right)^2 + \Upsilon_{i-1} \left(\left(\frac{\delta_W}{K_{i-1} \times FD_{i-1}^2} \right)^2 + \left(\frac{\delta_L}{FD_{i-1}} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD_{i-1}} \right)^2 \right)} \quad (\text{A.37})$$

Tal que:

$$C_{L_{i-1}} = C_{self} + C_{FO} + C_M \quad (\text{A.38})$$

En la ecuación anterior C_M representa cualquier capacitancia fuera de las producidas por las compuertas lógicas del camino (como la proveniente de una interconexión), por su parte es claro que en un camino lógico.

$$C_{FO} = C_{in_i} + C_{FO_G} \quad (\text{A.39})$$

Donde C_{FO_G} representa a cualquier compuerta que esta conectada al nodo de salida pero se encuentra fuera del camino bajo análisis, como:

$$C_{in_i} = (W_{min} L_{min} C_{ox})(K_i \times (FD_{in} + FD_{ip})) \quad (\text{A.40})$$

Se puede expresar la ecuación A.37 como:

$$\sigma_{D_{i-1}} = (C_{self} + C_{FO} + C_M) \gamma = \gamma C_{self} + \gamma (W_{min} L_{min} C_{ox} (FD_{in} + FD_{ip}) K_i) + \gamma C_{FO_G} + \gamma C_M \quad (\text{A.41})$$

Donde:

$$\gamma = \left(\frac{\kappa}{\beta_{0_{i-1}}} \frac{1}{K_{i-1}} \right) \sqrt{\Lambda_{i-1} \left(\frac{\delta_{V_T}}{\sqrt{K_{i-1}} \times FD_{i-1}^{1,5}} \right)^2 + \Upsilon_{i-1} \left(\left(\frac{\delta_W}{K_{i-1} \times FD_{i-1}^2} \right)^2 + \left(\frac{\delta_L}{FD_{i-1}} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD_{i-1}} \right)^2 \right)} \quad (\text{A.42})$$

Derivando

$$\frac{\partial \sigma_{D_{i-1}}}{\partial K_i} = \gamma (W_{min} L_{min} C_{ox} (FD_{in} + FD_{ip})) \quad (\text{A.43})$$

Finalmente.

$$\frac{\partial \sigma_{D_{i-1}}}{\partial K_i} = (FD_{i_n} + FD_{i_p}) \left(\frac{\kappa W_{min} L_{min} C_{ox}}{\beta_{0_{i-1}} K_{i-1}} \right)$$

$$\sqrt{\Lambda_{i-1} \left(\frac{\delta_{V_T}}{\sqrt{K_{i-1}} \times FD_{i-1}^{1,5}} \right)^2 + \Upsilon_{i-1} \left(\left(\frac{\delta_W}{K_{i-1} \times FD_{i-1}^2} \right)^2 + \left(\frac{\delta_L}{FD_{i-1}} \right)^2 + \left(\frac{\delta_{T_{ox}}}{FD_{i-1}} \right)^2 \right)}$$

(A.44)

Q.E.D.

A.5. Cálculos sobre las sensibilidades

A.5.1. Demostración de las sensibilidades

Con base a la ecuación 2.24.

$$D = \ln(2)C_L \times (R_{n,p}) \quad (\text{A.45})$$

Donde $R_{n,p}$ representa la resistencia de la red de conmutación activa, por su parte C_L representa la capacitancia de carga total de la compuerta ($C_{FO} + C_{self}$) suponiendo que:

$$C_{FO} \gg C_{self} \rightarrow C_L \approx C_{FO} \quad (\text{A.46})$$

Esta condición de trabajo es muy común en los diseños digitales y permite eliminar la dependencia de los parámetros del transistor en la carga, tal que cualquier derivada con respecto a un parámetro de los transistores es absorbida por el termino $R_{n,p}$, el cual puede encontrarse en la ecuación 2.1, misma que es expresada aquí para una mejor visualización de la demostración.

$$R_{n,p} = \frac{V_{DD}}{\mu C_{ox} \frac{W_{n,p}}{L_{min}} (V_{DD} - |V_T|)^\alpha} \quad (\text{A.47})$$

Para simplificar cálculos posteriores se escribirá la expresión encontrada sustituyendo la relación $C_{ox} = \frac{\epsilon_{ox}}{T_{ox}}$, tal que con estos cambios y al reacomodar términos:

$$R_{n,p} = \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min} T_{ox}}{\mu \epsilon_{ox} W_{n,p}} \quad (\text{A.48})$$

Claramente la derivada parcial se puede calcular en este punto con respecto a los parámetros de interés, tal que directamente se observa:

$$S_L^D = \frac{\partial D}{\partial L} = \frac{V_{DD} \ln(2) C_L}{(V_{DD} - |V_T|)^\alpha} \times \frac{T_{ox}}{\mu \epsilon_{ox} W_{n,p}} \quad (\text{A.49})$$

$$S_{T_{ox}}^D = \frac{\partial D}{\partial T_{ox}} = \frac{V_{DD} \ln(2) C_L}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min}}{\mu \epsilon_{ox} W_{n,p}} \quad (\text{A.50})$$

$$S_W^D = \frac{\partial D}{\partial W} = -\frac{V_{DD} \ln(2) C_L}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min} T_{ox}}{\mu \epsilon_{ox} W_{n,p}^2} \quad (\text{A.51})$$

Para expresar en términos de la escala de la compuerta lógica se recuerda que:

$$W_{n,p} = K \times FD \times W_{min} \quad (\text{A.52})$$

Realizando la sustitución del parámetro anterior en las ecuaciones de las sensibilidades encontradas:

$$S_L^D = \frac{\partial D}{\partial L} = \frac{\ln(2) C_L}{K \times FD} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{T_{ox}}{\mu \epsilon_{ox} W_{min}} \quad (\text{A.53})$$

$$S_{T_{ox}}^D = \frac{\partial D}{\partial T_{ox}} = \frac{\ln(2) C_L}{K \times FD} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min}}{\mu \epsilon_{ox} W_{min}} \quad (\text{A.54})$$

$$S_W^D = \frac{\partial D}{\partial W} = -\frac{\ln(2) C_L}{K^2 \times FD^2} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min} T_{ox}}{\mu \epsilon_{ox} W_{min}} \times \frac{1}{W_{min}} \quad (\text{A.55})$$

Q.E.D.

A.5.2. Modelos simplificados

Definiendo

$$\beta_0 = \frac{\mu\epsilon_{ox}W_{min}}{T_{ox}L_{min}} \quad (A.56)$$

$$\kappa = \ln(2) \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \quad (A.57)$$

Expresando el valor de $\sigma_{D_{V_T}}$ con base a las ecuaciones 2.36 y 2.41.

$$\sigma_{D_{V_T}} = S_{V_T}^D \times \left(\sigma_{V_{T0}} \times \sqrt{\frac{1}{K \times FD}} \right) \quad (A.58)$$

De forma similar utilizando las ecuaciones 2.42 a 2.44 se obtienen las sensibilidades de los otros parámetros de interés, las cuales al combinarse junto a los valores de las desviaciones estándar de cada parámetro expresadas en el conjunto de ecuaciones que abarca desde 2.45 a 2.47 permiten formar el valor del σ_D para cada uno del resto de parámetros de interés.

$$\sigma_{D_W} = S_W^D \times \sigma_W \quad (A.59)$$

$$\sigma_{D_L} = S_L^D \times \sigma_L \quad (A.60)$$

$$\sigma_{D_{T_{ox}}} = S_{T_{ox}}^D \times \sigma_{T_{ox}} \quad (A.61)$$

Se empieza por analizar el parámetro V_T .

$$\sigma_{D_{V_T}} = \frac{C_L}{K \times FD} \times \frac{\ln(2)V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min}T_{ox}}{\mu\epsilon_{ox}W_{min}} \times \left(\frac{\alpha}{V_{DD} - |V_T|} \right) \times \sigma_{V_{T0}} \sqrt{\frac{1}{K \times FD}} \quad (A.62)$$

Reconociendo

$$\sigma_{D_{V_T}} = \frac{C_L}{(K \times FD)^{1,5}} \times \frac{\kappa}{\beta_0} \times \left(\frac{\alpha}{V_{DD} - |V_T|} \right) \sigma_{V_{T0}} \quad (A.63)$$

Denominando

$$\delta_{V_T} = \left(\frac{\alpha}{V_{DD} - |V_T|} \right) \sigma_{V_{T0}} \quad (A.64)$$

Se puede escribir finalmente la forma compacta del $\sigma_{D_{V_T}}$ como:

$$\sigma_{D_{V_T}} = \frac{\kappa}{\beta_0} \times \frac{C_L}{(K \times FD)^{1,5}} \times \delta_{V_T} \quad (A.65)$$

Para el parámetro W se tiene:

$$\sigma_{D_W} = \frac{\ln(2)C_L}{K^2 \times FD^2} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min}T_{ox}}{\mu\epsilon_{ox}W_{min}} \times \frac{1}{W_{min}} \times \sigma_{D_W} \quad (A.66)$$

Si bien la sensibilidad del W es acompañada de un signo negativo se eliminó dicho signo a propósito por los resultados discutidos de la ecuación 2.38, en el cual se expresa que la relación del modelo lineal del retardo originalmente relacionaba factores al cuadrado, los cuales absorben cualquier signo negativo.

Reconociendo

$$\sigma_{D_W} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K^2 \times FD^2} \times \frac{1}{W_{min}} \times \sigma_{D_W} \quad (A.67)$$

Denominando

$$\delta_W = \left(\frac{1}{W_{min}} \times \sigma_W \right) \quad (A.68)$$

Se puede escribir finalmente la forma compacta del σ_{D_W} como:

$$\sigma_{D_W} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K^2 \times FD^2} \times \delta_{D_W} \quad (A.69)$$

Para el parámetro L se expresa

$$\sigma_{D_L} = \frac{\ln(2)C_L}{K \times FD} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{T_{ox}}{\mu\epsilon_{ox}W_{min}} \times \sigma_L \quad (A.70)$$

Reconociendo que

$$\beta_0 L_{min} = \frac{\mu\epsilon_{ox}W_{min}}{T_{ox}} \quad (A.71)$$

Se puede escribir

$$\sigma_{D_L} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K \times FD} \times \frac{1}{L_{min}} \times \sigma_L \quad (A.72)$$

Denominando

$$\delta_L = \left(\frac{1}{L_{min}} \times \sigma_L \right) \quad (A.73)$$

Se puede expresar finalmente la versión compacta del σ_{D_L} .

$$\sigma_{D_L} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K \times FD} \times \delta_L \quad (A.74)$$

Finalmente para el parámetro T_{ox} se expresa

$$\sigma_{D_{T_{ox}}} = \frac{\ln(2)C_L}{K \times FD} \times \frac{V_{DD}}{(V_{DD} - |V_T|)^\alpha} \times \frac{L_{min}}{\mu\epsilon_{ox}W_{min}} \times \sigma_{T_{ox}} \quad (A.75)$$

Reconociendo que

$$\beta_0 T_{ox} = \frac{\mu \epsilon_{ox} W_{min}}{L_{min}} \quad (A.76)$$

Se puede escribir

$$\sigma_{D_{T_{ox}}} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K \times FD} \times \frac{1}{T_{ox}} \times \sigma_{T_{ox}} \quad (A.77)$$

Denominando

$$\delta_{T_{ox}} = \left(\frac{1}{T_{ox}} \times \sigma_{T_{ox}} \right) \quad (A.78)$$

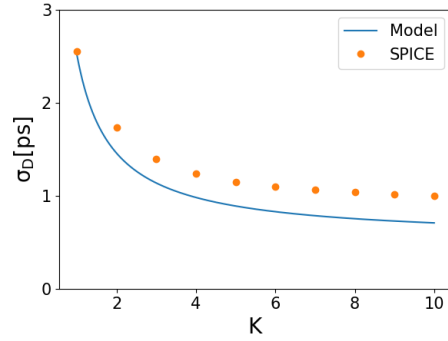
Se puede expresar finalmente la versión compacta del $\sigma_{D_{T_{ox}}}$.

$$\sigma_{D_{T_{ox}}} = \frac{\kappa}{\beta_0} \times \frac{C_L}{K \times FD} \times \delta_{T_{ox}} \quad (A.79)$$

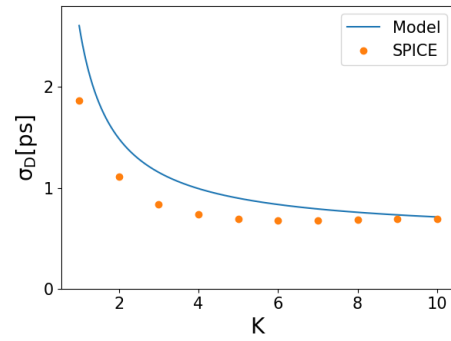
Q.E.D.

A.6. Gráficas auxiliares de validación para el modelo analítico a nivel compuerta

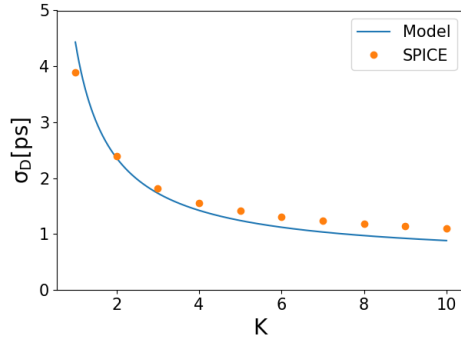
Figura A.1: Comparación entre el modelo analítico del σ_D y SPICE, en la compuerta NAND2



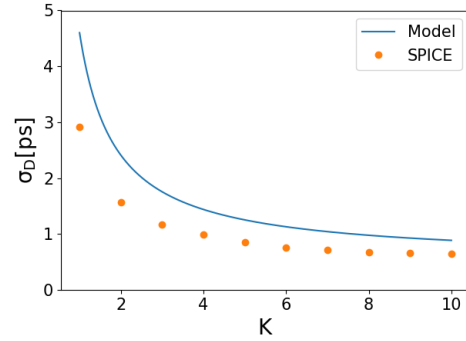
(a) Red PullDown, $C_L = 2[\text{fF}]$



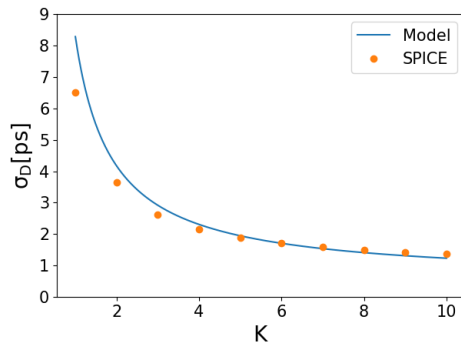
(d) Red PullUp, $C_L = 2[\text{fF}]$



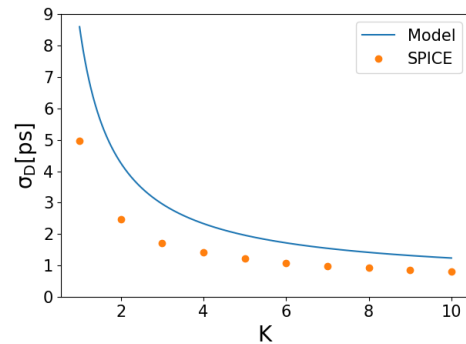
(b) Red PullDown, $C_L = 4[\text{fF}]$



(e) Red PullUp, $C_L = 4[\text{fF}]$

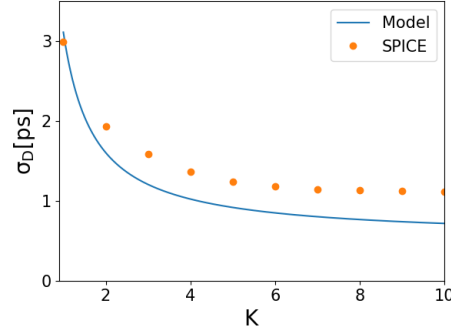


(c) Red PullDown, $C_L = 8[\text{fF}]$

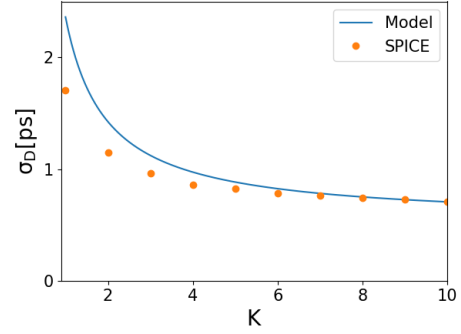


(f) Red PullUp, $C_L = 8[\text{fF}]$

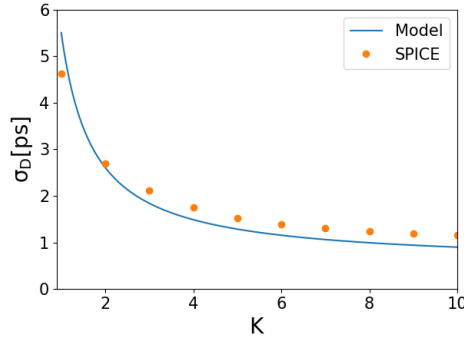
Figura A.2: Comparación entre el modelo analítico del σ_D y SPICE, en la compuerta NOR2



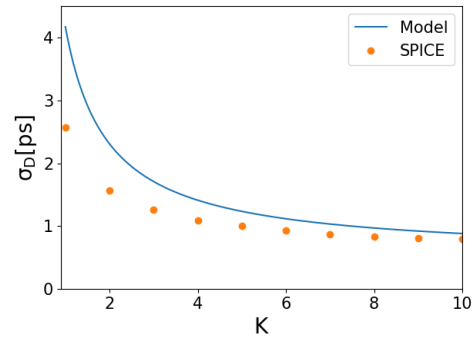
(a) Red PullDown, $C_L = 2[\text{fF}]$



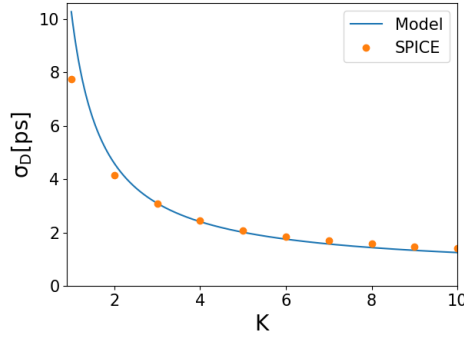
(d) Red PullUp, $C_L = 2[\text{fF}]$



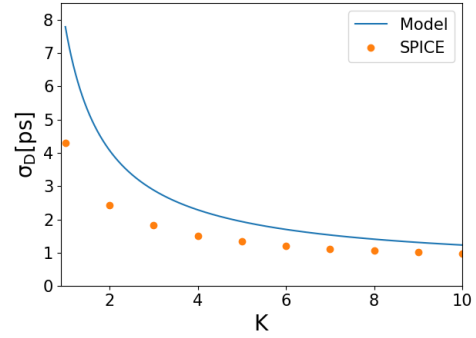
(b) Red PullDown, $C_L = 4[\text{fF}]$



(e) Red PullUp, $C_L = 4[\text{fF}]$



(c) Red PullDown, $C_L = 8[\text{fF}]$



(f) Red PullUp, $C_L = 8[\text{fF}]$

A.7. Transformación de Cholesky

La transformación de Cholesky es un proceso para generar un conjunto de n variables correlacionadas ($Z_{C_1}, Z_{C_2} \dots Z_{C_n}$) a partir de un conjunto de n variables independientes ($Z_1, Z_2 \dots Z_n$) y una matriz de correlación deseada.

El proceso para generar las variables correlacionadas se muestra gráficamente en la Figura A.3.

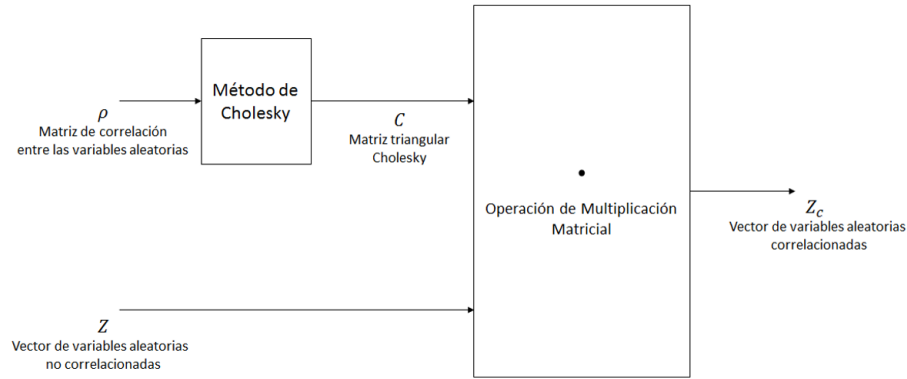


Figura A.3: Proceso completo de la transformación de Cholesky

No se debe confundir la transformación de Cholesky con el método de Cholesky, el primero es el proceso discutido para generar variables aleatorias correlacionadas mientras que el segundo es un algoritmo que opera sobre matrices cuadradas simétricas (como una matriz de covarianza) para generar una matriz triangular inferior [5], este algoritmo se basa en que cualquier matriz cuadrada con pivotes no nulos se puede escribir como:

$$\mathbf{A} = \mathbf{C}\mathbf{C}^T \quad (\text{A.80})$$

Se denomina al termino \mathbf{C} como de descomposición de Cholesky de la matriz \mathbf{A} , los elementos de dicha matriz se pueden calcular por medio de las siguientes ecuaciones.

$$c_{ii} = a_{ii} - \sum_{k=1}^{i-1} c_{ik}^2 \quad (\text{A.81})$$

$$c_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} c_{ik}c_{jk}}{c_{jj}} \quad \text{Para } i \neq j \quad (\text{A.82})$$

Una vez obtenida la descomposición de Cholesky se debe operar esta con el vector de variables aleatorias independientes Z para obtener el conjunto de variables correlacionadas.

$$\begin{bmatrix} Z_{C_1} \\ Z_{C_2} \\ \vdots \\ Z_{C_n} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ 0 & c_{22} & \dots & c_{2n} \\ \vdots & \ddots & \dots & \\ 0 & 0 & \dots & c_{nn} \end{bmatrix} \times \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} \quad (\text{A.83})$$

Claramente la transformación produce

$$\begin{aligned} Z_{C_1} &= c_{11}Z_1 + c_{12}Z_2 \dots + c_{1n}Z_n \\ Z_{C_2} &= c_{22}Z_2 \dots + c_{2n}Z_n \\ &\vdots \\ Z_{C_n} &= c_{nn}Z_n \end{aligned} \quad (\text{A.84})$$

El sistema mostrado en [A.84](#) puede ser introducido en una netlist de SPICE para generar las variables correlacionadas para una simulación MonteCarlo.

Apéndice B

B.1. Obtención de los coeficientes de la métrica propuesta

B.1.1. Coeficiente a_1

Tomando como referencia la ecuación:

$$\frac{\partial \sigma_D}{\partial K} = - \frac{\Lambda \sigma_{D_{V_T}} \frac{\partial \sigma_{D_{V_T}}}{\partial K} + \Upsilon \left(\sigma_{D_W} \frac{\partial \sigma_{D_W}}{\partial K} + \sigma_{D_L} \frac{\partial \sigma_{D_L}}{\partial K} + \sigma_{D_{Tox}} \frac{\partial \sigma_{D_{Tox}}}{\partial K} \right)}{\sqrt{\Lambda \left(\sigma_{D_{V_T}} \right)^2 + \Upsilon \left(\left(\sigma_{D_W} \right)^2 + \left(\sigma_{D_L} \right)^2 + \left(\sigma_{D_{Tox}} \right)^2 \right)}} \quad (B.1)$$

Donde

$$\frac{\partial \sigma_{D_{V_T}}}{\partial K} = -1,5 \frac{\kappa}{\beta_0} \frac{C_L \delta_{V_T}}{F D^{1,5}} \left(\frac{1}{K^{2,5}} \right) \quad (B.2)$$

$$\frac{\partial \sigma_{D_W}}{\partial K} = -2 \frac{\kappa}{\beta_0} \frac{C_L \delta_W}{F D^2} \left(\frac{1}{K^3} \right) \quad (B.3)$$

$$\frac{\partial \sigma_{D_L}}{\partial K} = - \frac{\kappa}{\beta_0} \frac{\delta_L C_L}{F D} \left(\frac{1}{K^2} \right) \quad (B.4)$$

$$\frac{\partial \sigma_{D_{Tox}}}{\partial K} = - \frac{\kappa}{\beta_0} \frac{\delta_{Tox} C_L}{F D} \left(\frac{1}{K^2} \right) \quad (B.5)$$

Y el modelo de métrica

$$M_i = \frac{C_{L_i}}{K_i^{a_1}} \quad (\text{B.6})$$

Si se toma como referencia una topología de un inversor conmutando por su red PullDown, la ecuación de la derivada es solo función de C_L y K , mientras que la métrica es función de C_L , K y a_1 , si solo se interesa en ajustar la tendencia de las curvas se pueden normalizar ambas con respecto a su valor máximo y posteriormente usar el parámetro a_1 para ajustar ambas curvas, dicho proceso se muestra en la Figura B.5.

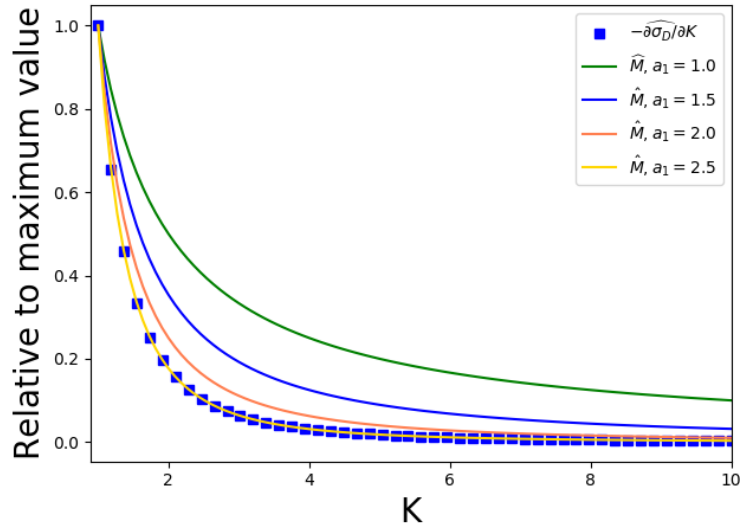


Figura B.4: Diferentes ajuste de curva por medio del parámetro a_1

La nomenclatura \widehat{X} hace referencia a una variable normalizada con respecto a su máximo. Cabe destacar que el valor normalizado es independiente de C_L ya que este afecta de forma lineal tanto a la derivada como a la métrica.

Para encontrar el valor de a_1 que genera la diferencia mínima entre las curvas se realiza la estrategia de los mínimos cuadrados. Definiendo el termino sumando cuadrático del error.

$$S(a_1) = \sum_{K=1}^{K_{max}} \left(-\frac{\partial \widehat{\sigma}_D}{\partial K} - \widehat{M}(a_1) \right)^2 \quad (\text{B.7})$$

En este trabajo se considera $K_{max} = 10$.

Para encontrar el mínimo valor del error se necesita resolver la ecuación

$$\frac{d}{da_1} S(a_1) = 0 \quad (\text{B.8})$$

Lo cual se puede hacer sustituyendo las ecuaciones B.1, B.6 en B.7 lo cual da información suficiente para computar B.8 y posteriormente utilizar un solucionador de raíces como el algoritmo Newton-Rhapson. Una vez solucionada deberá usarse el criterio de la segunda derivada para garantizar que el/los valores encontrados sean puntos mínimos.

Realizando lo discutido en el párrafo anterior se encontró que el valor minimo se da en $a_1 = 2,5$.

B.1.2. Coeficiente a_2

El calculo del coeficiente a_2 es muy similar al realizado para a_1 sin embargo a_2 es utilizado para el efecto de una métrica que mide la covarianza generada entre dos compuertas.

El modelo para calcular la covarianza entre dos compuertas lógicas se muestra a continuación.

$$\text{COV}(D_i, D_j) = \rho \frac{\kappa^2 C_{eff}^2}{\beta_i \beta_j} \times \sqrt{\Upsilon_i \Upsilon_j} \times \left(\frac{\delta_W}{(K_{eff}^2 F D_{eff}^2)^2} + \frac{\delta_L}{K_{eff}^2 F D_{eff}^2} + \frac{\delta_{T_{ox}}}{K_{eff}^2 F D_{eff}^2} \right) \quad (\text{B.9})$$

Reconociendo que la ecuación es muy similar a la encontrada para el σ_D^2 de una compuerta con dimensiones $K_{eff} = \sqrt{K_1 K_2}$ y carga $C_{eff} = \sqrt{C_{L1} C_{L2}}$ que no pose variaciones independientes ($\delta_{V_T} = 0$), por lo que puede usarse una métrica similar a B.6 para cuantificar la contricción.

$$M_C = \frac{C_{eff}}{K_{eff}^{a_2}} \quad (\text{B.10})$$

Derivando la expresión B.9 con respecto a K_{eff} se tiene

$$\frac{\partial \text{COV}(D_i, D_j)}{\partial K_{eff}} = \rho \frac{\kappa^2 C_{eff}^2}{\beta_i \beta_j} \times \sqrt{\Upsilon_i \Upsilon_j} \times \left(\frac{-4\delta_W}{(K_{eff}^5 F D_{eff}^4)} + \frac{-2\delta_L}{K_{eff}^3 F D_{eff}} + \frac{-2\delta_{T_{ox}}}{K_{eff}^3 F D_{eff}^2} \right) \quad (\text{B.11})$$

Comparando la métrica en B.10 y la raíz de la derivada calculada B.11 por medio de su grafico en unidades normalizadas.

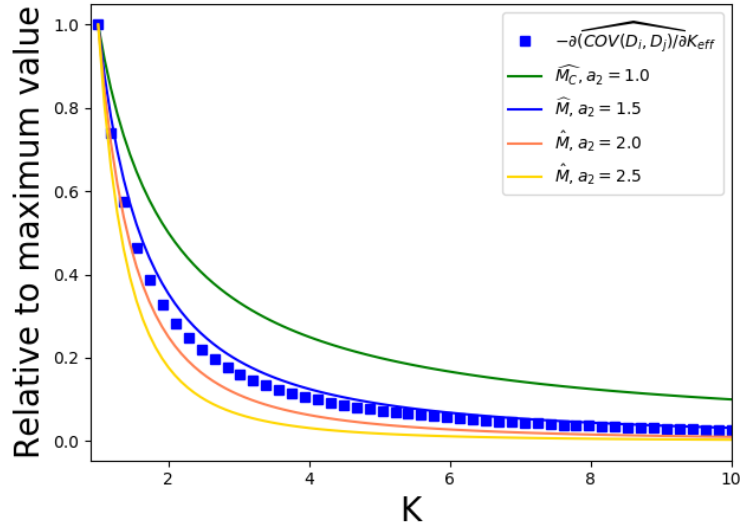


Figura B.5: Diferentes ajuste de curva por medio del parámetro a_2

Cabe destacar que es necesario utilizar la raíz debido a que la covarianza tiene unidades segundos cuadrados a diferencia de la desviación estándar que pose unidades lineales.

Se plantea ahora la función sumando cuadrático del error.

$$S(a_2) = \sum_{K=1}^{K_{max}} \left(-\frac{\partial \sqrt{\widehat{\text{COV}(D_i, D_j)}}}{\partial K_{eff}} - \widehat{M}_C(a_2) \right)^2 \quad (\text{B.12})$$

Tal que el valor mínimo de a_2 se puede encontrar al solucionar la ecuación:

$$\frac{d}{da_2} S(a_2) = 0 \quad (\text{B.13})$$

Dicha solución se da por medio de un algoritmo numérico como Newton-Rhapson, también es necesario comprobar que el resultado corresponda con un valor mínimo

y no un máximo para diferenciar dichos casos se puede usar el criterio de la segunda derivada en el punto de evaluación, al solucionar para la tecnología dada se obtiene que el mínimo se encuentra en el punto $a_2 = 1,5$.

Apéndice C

C.1. Aplicación de los multiplicadores de LaGrange

En este Apéndice se presentan ejemplos sencillos sobre la aplicación de los multiplicadores de LaGrange.

Como se comenta en el Capitulo 3 los multiplicadores de LaGrange permiten obtener la solución óptima para un sistema sujeto a restricciones.

C.1.1. Ejemplo 1: Minimizar la distancia desde el origen a una recta dada

Dada la recta

$$y = 2 - x \tag{C.1}$$

Y conociendo que la distancia desde el origen a un punto x, y se puede plantear como:

$$d = \sqrt{x^2 + y^2} \tag{C.2}$$

Si se quiere minimizar la distancia del origen a un punto de la recta se puede plantear el problema como:

$$\min(d) \quad s.a \quad g = 0 \tag{C.3}$$

Donde

$$g = y + x - 2 \tag{C.4}$$

En base a la teoría descrita en el Capitulo 3 dicho problema puede solucionarse si se encuentran las raíces del gradiente del Lagrangiano.

Empezando por definir el Lagrangiano.

$$\mathcal{L}(x, y, \lambda) = d - \lambda g = \sqrt{x^2 + y^2} - \lambda(y + x - 2) \quad (\text{C.5})$$

Tal que el gradiente del Lagrangiano se puede expresar como:

$$\nabla \mathcal{L}(x, y) = \ell_i \vec{i} + \ell_j \vec{j} + \ell_k \vec{k} \quad (\text{C.6})$$

Donde

$$\ell_i = \frac{\partial \mathcal{L}}{\partial x} = \frac{\partial d}{\partial x} - \frac{\partial \lambda g}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} - \lambda \quad (\text{C.7})$$

$$\ell_j = \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial d}{\partial y} - \frac{\partial \lambda g}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}} - \lambda \quad (\text{C.8})$$

$$\ell_k = \frac{\partial \mathcal{L}}{\partial \lambda} = \frac{\partial d}{\partial \lambda} - \frac{\partial \lambda g}{\partial \lambda} = 0 - g = -y - x + 2 \quad (\text{C.9})$$

La solución del problema de optimización se encuentra al resolver el sistema de ecuaciones.

$$\frac{x}{\sqrt{x^2 + y^2}} - \lambda = 0 \quad (\text{C.10})$$

$$\frac{y}{\sqrt{x^2 + y^2}} - \lambda = 0 \quad (\text{C.11})$$

$$-y - x + 2 = 0 \quad (\text{C.12})$$

Dicho sistema puede ser resuelto por métodos computacionales como Newton-Raphson.

Solucionando se encuentra que

$$x = y = 1 \quad (\text{C.13})$$

Por lo que gráficamente

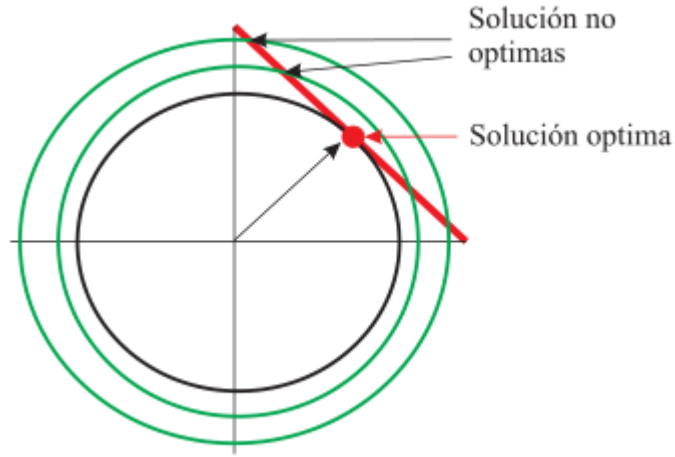


Figura C.6: Solución grafica del problema de optimización

C.1.2. Aplicación en la optimización de caminos lógicos

Con base al circuito mostrado

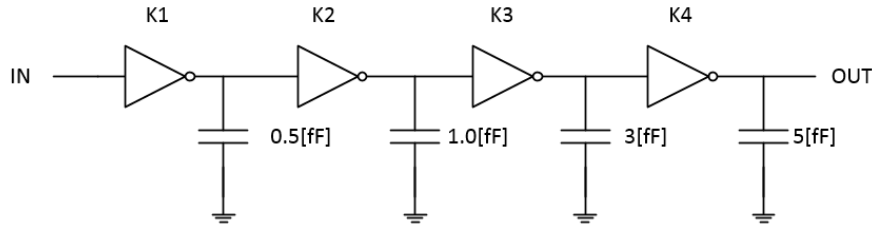


Figura C.7: Solución grafica del problema de optimización

Optimizar su retardo nominal usando el modelo analítico sujeto a una restricción de área $A_C = 25 * W_{min}L_{min}$.

Utilizando la ecuación 2.98 para definir el retardo del camino.

$$\mu_P = \sum_{i=1}^n \mu_i = \mu_1 + \mu_2 + \mu_3 + \mu_4 \quad (C.14)$$

Donde con base a la ecuación 2.87.

$$\mu_i = \frac{\kappa}{\beta_0} * \frac{C_L}{K} * \frac{\chi}{FD} \quad (C.15)$$

Debido a que el retardo nominal es independiente de la transición se puede simplificar la ecuación anterior a:

$$\mu_i = \frac{\kappa}{\beta_n} * \frac{C_{L_i}}{K_i} \quad (C.16)$$

Si se desprecia la contribución de C_{self} en C_L para cada etapa se puede expresar C_L en terminos de la carga intrínseca del nodo y el FANOUT.

$$C_{L_i} = C_i + C_{in_{i+1}} \quad (C.17)$$

Usando la ecuación 2.9 se tiene

$$C_{L_i} = C_i + K_{i+1} * 3W_{min}L_{min}C_{ox} \quad (C.18)$$

Expresando el retardo nominal del camino

$$\mu_P(K_1, K_2, K_3, K_4) = \frac{\kappa}{\beta_n} * \left(\frac{0,5 + \gamma K_2}{K_1} + \frac{1 + \gamma K_3}{K_2} + \frac{3 + \gamma K_4}{K_3} + \frac{5}{K_4} \right) \quad (C.19)$$

Donde

$$\gamma = 3W_{min}L_{min}C_{ox} \quad (C.20)$$

Esta es la función objetivo a optimizar.

Para la restricción se puede expresar el área del camino con base a las ecuaciones 3.9 y 3.12.

$$A_P = \sum_{i=1}^n A_i \quad (C.21)$$

Al existir puras compuertas inversoras:

$$A_i = 3W_{min}L_{min} * K_i \quad (C.22)$$

Sustituyendo

$$A_P = 3W_{min}L_{min} * (K_1 + K_2 + K_3 + K_4) \quad (C.23)$$

Por lo que la función de restricción queda como

$$A_P = A_C \rightarrow g = A_P - A_C = (3(K_1 + K_2 + K_3 + K_4) - 25)W_{min}L_{min} \quad (C.24)$$

Planteando el Lagrangiano.

$$\begin{aligned} \mathcal{L} = \frac{\kappa}{\beta_n} * \left(\frac{0,5 + \gamma K_2}{K_1} + \frac{1 + \gamma K_3}{K_2} + \frac{3 + \gamma K_4}{K_3} + \frac{5}{K_4} \right) \\ - \lambda * W_{min}L_{min}(3(K_1 + K_2 + K_3 + K_4) - 25) \end{aligned} \quad (C.25)$$

La solución optima podrá calcularse al solucionar el sistema de ecuaciones

$$\frac{\partial \mathcal{L}}{\partial K_1} = 0 \quad (\text{C.26})$$

$$\frac{\partial \mathcal{L}}{\partial K_2} = 0 \quad (\text{C.27})$$

$$\frac{\partial \mathcal{L}}{\partial K_3} = 0 \quad (\text{C.28})$$

$$\frac{\partial \mathcal{L}}{\partial K_4} = 0 \quad (\text{C.29})$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \quad (\text{C.30})$$

Dicha solución devuelve

$$K_1 = 1,45 \quad (\text{C.31})$$

$$K_2 = 1,59 \quad (\text{C.32})$$

$$K_3 = 2,41 \quad (\text{C.33})$$

$$K_4 = 2,87 \quad (\text{C.34})$$

Con un retardo

$$\mu_P = 63,28[\text{ps}] \quad (\text{C.35})$$

Apéndice D

La herramienta SSTA tienen como objetivo analizar y optimizar un camino lógico dentro de un circuito digital, esta aplicación se caracteriza por que es muy accesible a usuarios nuevos debido a su GUI amigable e interactiva.

A lo largo de este apéndice se mostrara y explicara el funcionamiento de la herramienta.

D.1. Archivo de entrada

La herramienta cuenta con su propio lenguaje de descripción, el cual esta inspirado en el formato original utilizado por el simposio internacional de circuitos y sistemas para describir los circuitos combinacionales de referencia en 1985. Aunque también en la herramienta se encuentran integrados dos compiladores de lenguajes externos como lo son Verilog y SPICE, mas adelante en este apéndice se explicara como usar estos compiladores.

Un ejemplo de un archivo nativo en formato ISCAS se muestra a continuación.

```

1  -----
2  -- Archivo:   ISCAS DE PRUEBA 01
3  -- Ingeniero: Greg Rifka
4  -- Descripcion: Ejemplo
5  -----
6  Start   Gate
7  --Todas las compuertas empiezan con X,
8  --sin coma entre las asignaciones
9  X1  NOT   IN1   G1           K=1   X=100 Y=0
10 X2  NOT   IN3   G2           K=1   X=200 Y=200
11 X3  NAND2 G1    IN2   Y1     K=2   X=300 Y=400
12 X4  NAND2 IN2   G2    Y2     K=2   X=1000 Y=200
13 X5  NOR2  Y1    Y2     OUT    K=4   X=500 Y=1000
14 End     Gate
15
16 Start   Capacitor
17 --Todos los capacitores empiezan con C
18 C1    OUT    5E-15
19 End     Capacitor
20
21 Start   Options
22 NAME:           Generic_Logic
23 Language:       English
24 SpacialCorrelation: OFF
25 ShowCap:       ALL
26 VarVT          ON
27 VarW           ON
28 VarL           ON
29 VarTOX         ON
30 ShowPath:      ALL
31 Accuracy:      High
32 IN_NODES:      IN1 IN2 IN3
33 OUT_NODES:     OUT
34 End   Options
35
36 EOF

```

En el ejemplo anterior se muestran tres secciones bien definidas que inician por la palabra reservada Start * y finalizan con la palabra End *, estas secciones son siempre obligatorias y especifican:

1. El circuito a nivel compuerta lógica, las compuertas soportadas por la herramienta son NAND2, NOR2 e inversores, se puede establecer un nombre personalizado a cada compuerta pero es de carácter obligatorio que el primer carácter del nombre sea la letra *X* para el resto del nombre cualquier combinación de 10 caracteres en forma de números y letras es aceptable, el orden de los nodos siempre va de entrada(s) a salida.
2. Capacitancias externas a las compuertas con referencia a tierra en cualquier nodo del circuito para simular efectos eléctricos ajenos a los circuitos, principalmente la capacitancia no despreciable de los metales de interconexión.
3. Opciones del análisis sobre que tipo de variaciones incluir y que caminos lógicos analizar (poner todas las opciones es de carácter obligatorio, sin embargo el orden es irrelevante), así como otras configuraciones como el idioma de reportes textuales.

Dentro y fuera de estas secciones se pueden especificar comentarios auxiliares para hacer mas legible el netlist, los comentarios siempre deben empezarse al principio de la linea por medio dos guiones altos —, la herramienta SSTA ignora todos los comentarios, lineas en blanco y espacios entre descripciones por lo que puede aprovecharse esto para dar un formato legible a los archivos de entrada.

D.2. Interfaz Primaria

Una vez definido el archivo de entrada se puede ejecutar la herramienta SSTA, para lograr la ejecución se debe tener instalado la versión de Python 3.3 (como mínimo) y el interprete anaconda el cual cargara automáticamente las librerías necesarias a excepción del visualizador de esquemáticos esta ultima librería deberá ser instalada manualmente por medio del comando **pip install SchemDraw** en la terminal de anaconda, este proceso solo es necesario para configurar el equipo para el primer uso.

Con el equipo configurado se puede acceder a la interfaz principal de la herramienta la cual es mostrada a continuación.

La interfaz principal puede ser dividida en dos partes

1. Barra de Menú

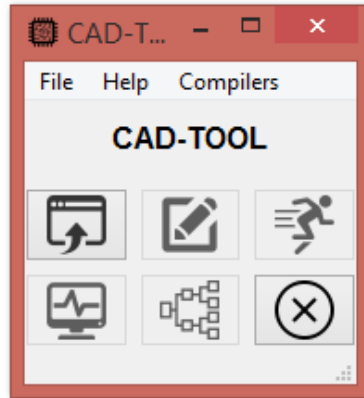


Figura D.8: Interfaz principal de la herramienta SSTA

2. Zona de trabajo

Se describirá cada una de estas partes por separado

D.2.1. Barra de Menú

La barra de menú se encuentra en el borde superior de la ventana de la herramienta y esta se subdivide a la vez en:

- **File:** Al dar click en esta parte se abre un submenú con dos opciones:
 1. Open: Permite abrir un archivo ISCAS y cargarlo, se puede usar la combinación de acceso rápido Ctrl-O para usar esta opción.
 2. Exit: Cierra la herramienta, Ctrl-E para acceso rápido.
- **Help:** Al dar click en esta parte se abre un submenú con dos opciones:
 1. About the tool: Explica de forma breve el propósito de la herramienta.
 2. Credits: Despliega el nombre del autor de esta tesis.
- **Compilers:** Al dar click en esta parte se abre un submenú con dos opciones:
 1. SPICE: Abre la interfaz de compilación, configurada para traducir código SPICE.

2. Verilog: Abre la interfaz de compilación, configurada para traducir código Verilog.

Las interfaces de compilación son prácticamente idénticas para el lenguaje Verilog y SPICE, dichas interfaces se muestran a continuación.

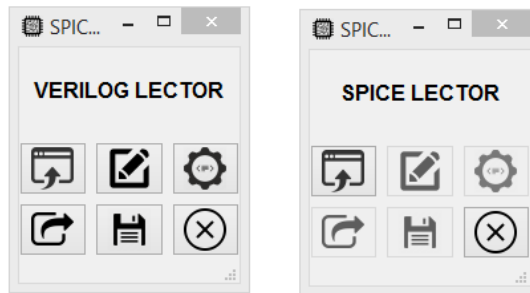


Figura D.9: Interfaces de compilación

Ambas interfaces se componen de los mismos botones por lo que su uso es idéntico, los botones del cuerpo son inteligentes y se activan progresivamente conforme se usa la herramienta, en el momento que aparece la interfaz solo el botón superior izquierdo e inferior derecho pueden ser pulsados, el botón inferior derecho sirve para cerrar la herramienta, mientras que el superior izquierdo corresponde con la acción de abrir un netlist del lenguaje a compilar, cargado el archivo se habilitaran los dos botones restantes de la fila, el primero de estos botones habilitados se denomina editar y abre el archivo a compilar en el editor de texto por defecto del S.O, el segundo de estos se denomina verificador y como es de esperarse verifica la sintaxis del archivo a compilar.

Después de verificar la sintaxis se habilitan el botón inferior izquierdo el cual compila el archivo de forma interna en la memoria RAM hecho esto finalmente se habilita el botón guardar el cual exporta el archivo interno a un directorio físico del disco duro y el cual puede ser cargado desde la interfaz principal.

Para utilizar código Verilog se debe describir el circuito en términos de compuertas lógicas y no mediante bloques de construcción especializados (ejemplo: bloques always o assign).

Un ejemplo de código Verilog sintetizable es mostrado a continuación.

```

1 module c17 (N1,N2,N3,N6,N7,N22,N23);
2 input N1,N2,N3,N6,N7;
3 output N22,N23;
4 wire N10,N11,N16,N19;
5
6 NA2 NAND2_1 ( .Q(N10), .A(N1), .B(N3) );
7 NA2 NAND2_2 ( .Q(N11), .A(N3), .B(N6) );
8 NA2 NAND2_3 ( .Q(N16), .A(N2), .B(N11) );
9 NA2 NAND2_4 ( .Q(N19), .A(N11), .B(N7) );
10 NA2 NAND2_5 ( .Q(N22), .A(N10), .B(N16) );
11 NA2 NAND2_6 ( .Q(N23), .A(N16), .B(N19) );
12
13 endmodule

```

Aun cuando la herramienta solo pueda trabajar con compuertas NAND2, NOR2 e inversores, se puede escribir el circuito en el código Verilog en términos de compuertas NAND,NOR,AND,OR,XOR de múltiples entradas aparte de las compuertas nativas, el compilador convertirá estas estructuras en su equivalente lógico compuesto de compuertas NAND2, todas las compuertas sintetizadas tendrán $K = 1$.

Para utilizar el compilador de SPICE es necesario describir el circuito usando los subcircuitos de la librería personalizada **PRIMITIVAS_SPICE.tanner**, (poner link de acceso) esta librería esta pensada para usarse por medio del programa S-Edit, ya que no solo contiene las directivas para generar el netlist adecuado sino los símbolos esquemáticos para permitir la captura del camino a analizar.

Una vez generado el esquemático se puede usar la opción de S-Edit de exportar netlist, el cual genera un archivo .sp que puede ser leído por el compilador.

D.2.2. Zona de trabajo

La zona de trabajo se compone de seis botones los cuales se activan progresivamente mientras se utiliza la herramienta, al abrir la herramienta por primera vez solo están habilitados dos botones, el de la esquina superior izquierda llamado **open** y el de la esquina inferior derecha llamado **close**.

El funcionamiento del botón **close** es muy directo y consiste en cerrar la

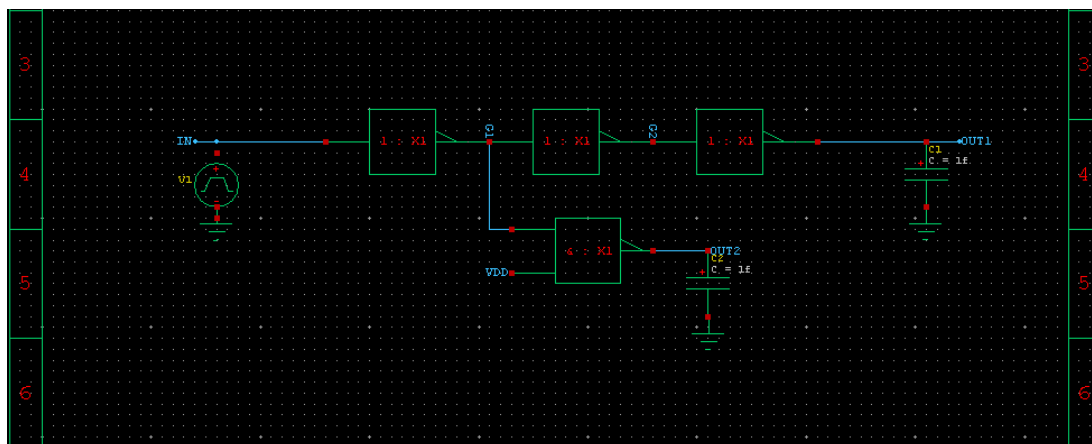


Figura D.10: Utilización de la librería **PRIMITIVAS_SPICE.tanner** en S-Edit

herramienta SSTA, por comodidad se puede usar la combinación de teclas Ctrl-E para realizar esta misma acción.

El botón **open** tiene el mismo comportamiento que la acción open de la barra de menú y dicho comportamiento consiste en cargar un archivo ISCAS por medio del explorador de archivos del S.O utilizado, una vez cargado el archivo se habilitaran todos los botones de la fila superior, el primero de estos botones habilitados se denomina **edit** y el segundo **run** ambos botones tienen simbolos distintivos que permiten su identificación rápidamente.

El botón **edit** como su nombre lo dice permite editar en por medio del editor de texto por defecto el archivo ISCAS actualmente cargado, es decir si se presiona este botón se abrirá comúnmente el Notepad con el archivo para su edición, si se salva cualquier cambio del archivo estos cambios se reflejaran en tiempo real en la herramienta.

El botón **Run** lleva un análisis pre-eliminar que permite reconocer todos los caminos lógicos en el circuito (para circuitos con un numero de nodos mayor que 16 solo se reconocen los caminos mas críticos), aparte de esto bajo este análisis se identifican las capacitancias efectivas en cada nodo, una vez corrido este análisis se habilitan los dos botones restantes de la fila inferior.

El primero de estos botones habilitados se denomina **Basic Report** y genera un archivo de texto en el idioma especificado (Ingles o Español), en dicho archivo se muestran los caminos identificados junto a su retardo de primer orden y las cargas efectivas de cada nodo, el tamaño de la lista de caminos mostrados puede

ser controlado por medio de la sección opciones en el formato ISCAS.

Un ejemplo de este tipo de archivo se muestra a continuación en su versión en inglés.

```

1 Logical paths and capacitance report
2 Analysis hour: 18:06:39
3 Circuit under test
4   Name: Cicuito_Sintetizado
5   Number of gates: 4
6   Number of nodes: 7
7 Input nodes
8   IN
9 Output nodes
10  OUT
11 Logic paths
12 Path: 1
13   Input node: IN
14   Output node: OUT
15   Logic Depth: 4
16   Area: 38.60[Wmin*Lmin]
17   Delay: 166.5[ps]
18   Sequence of nodes: ['IN', 'N2', 'N3', 'N4', 'OUT']
19   Electric effort: 16.2
20   Gates in path:
21     X1
22     X2
23     X3
24     X4
25 Nodes capacitance
26   Node: OUT
27   Value: 9.78[fF]
28   Node: N4
29   Value: 7.99[fF]
30   Node: N3
31   Value: 4.93[fF]
32   Node: N2
33   Value: 2.18[fF]
34   Node: GND
35   Value: 1.04[fF]
36   Node: VDDN3
37   Value: 0.77[fF]
38   Node: IN
39   Value: 0.56[fF]
40 End of report

```

El segundo botón habilitado se denomina como **Path Selector** y muestra una tabla ordenada de todos los caminos reconocidos, este flujo es mosrado a continuación.

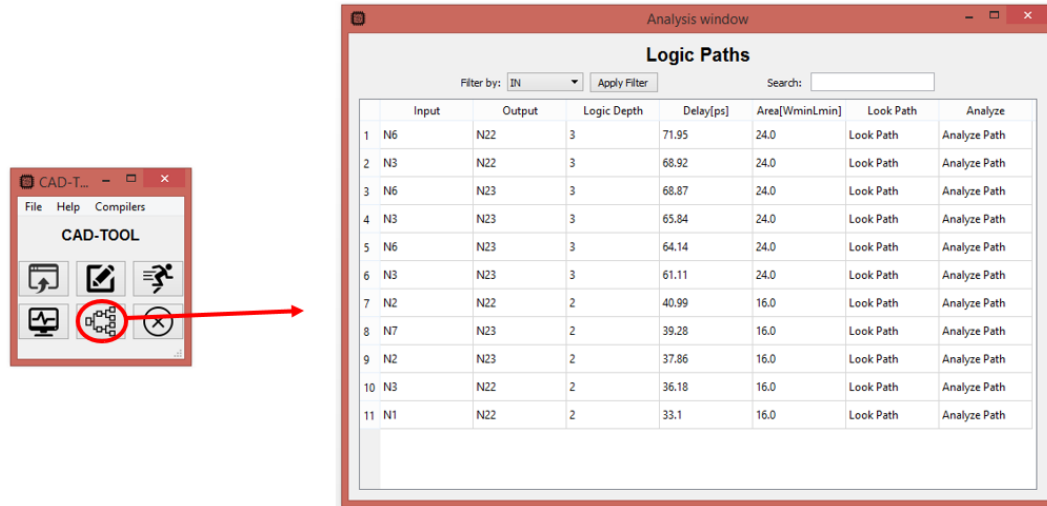


Figura D.11: Funcionamiento del **Path Selector**

D.3. Interfaz de selectora de caminos

La interfaz abierta por el botón **Path Selector** se denomina formalmente como interfaz de selectora de caminos y la función principal de esta misma es permitir la selección de un camino lógico para su posterior análisis y optimización.

La interfaz discutida es mostrada a continuación:

En el área de trabajo la interfaz muestra una lista de todos los caminos identificados, con sus propiedades mas importantes, estos caminos siempre serán ordenados con respecto a su retardo nominal de primer orden.

En la parte superior de la interfaz se encuentra integrado un filtro, el cual puede utilizarse para discriminar caminos en función de su profundidad (Logic Depth), terminal de entrada (IN) o terminal de salida (OUT), para discriminar los caminos basta con seleccionar uno de los criterios mencionados y escribir el valor buscado para dicho criterio en el cuestionario de búsqueda (search).

	Input	Output	Logic Depth	Delay[ps]	Area[WminLmin]	Look Path	Analyze
1	N6	N22	3	71.95	24.0	Look Path	Analyze Path
2	N3	N22	3	68.92	24.0	Look Path	Analyze Path
3	N6	N23	3	68.87	24.0	Look Path	Analyze Path
4	N3	N23	3	65.84	24.0	Look Path	Analyze Path
5	N6	N23	3	64.14	24.0	Look Path	Analyze Path
6	N3	N23	3	61.11	24.0	Look Path	Analyze Path
7	N2	N22	2	40.99	16.0	Look Path	Analyze Path
8	N7	N23	2	39.28	16.0	Look Path	Analyze Path
9	N2	N23	2	37.86	16.0	Look Path	Analyze Path
10	N3	N22	2	36.18	16.0	Look Path	Analyze Path
11	N1	N22	2	33.1	16.0	Look Path	Analyze Path

Figura D.12: Interfaz de selectora de caminos

Para cualquier camino lógico detectado la herramienta puede visualizar su esquemático, para hacer esto solo se debe dar doble click en la sección **Look Path** del camino a visualizar.

La opción **Analyze Path** aísla el camino lógico seleccionado para un análisis y/u optimización mas detallada, todas las compuertas fuera del camino que afecten a los nodos internos del camino seleccionado serán moldeadas como capacitores de carga para estos análisis detallados, al dar doble click a esta sección abrirá una ventana que preguntara al usuario que motor de optimización usar (LaGrange o método Heurístico), dicho proceso es mostrado a continuación.

En dicha ventana la opción de la izquierda corresponde con usar la metodología de los multiplicadores de LaGrange, mientras que la derecha permite usar el método Heurístico basado en métricas, el seleccionar una opción abrirá la ventana correspondiente de acuerdo motor de optimización seleccionado.

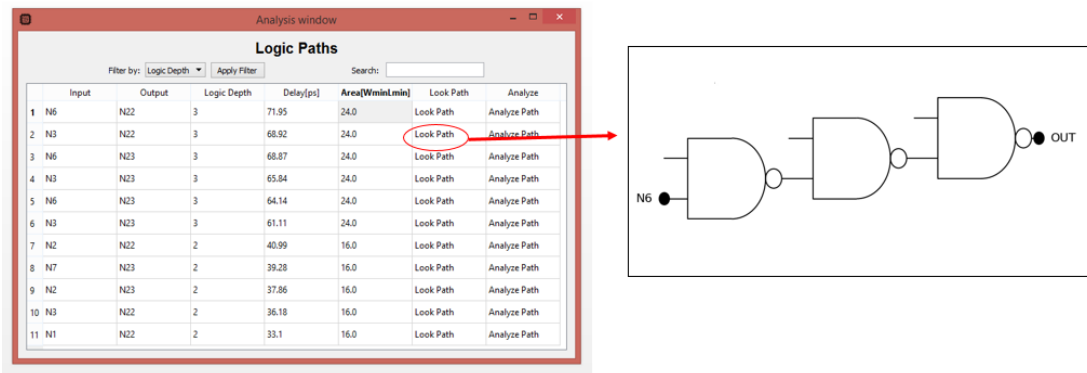


Figura D.13: Interfaz de selectora de caminos: **Look Path**

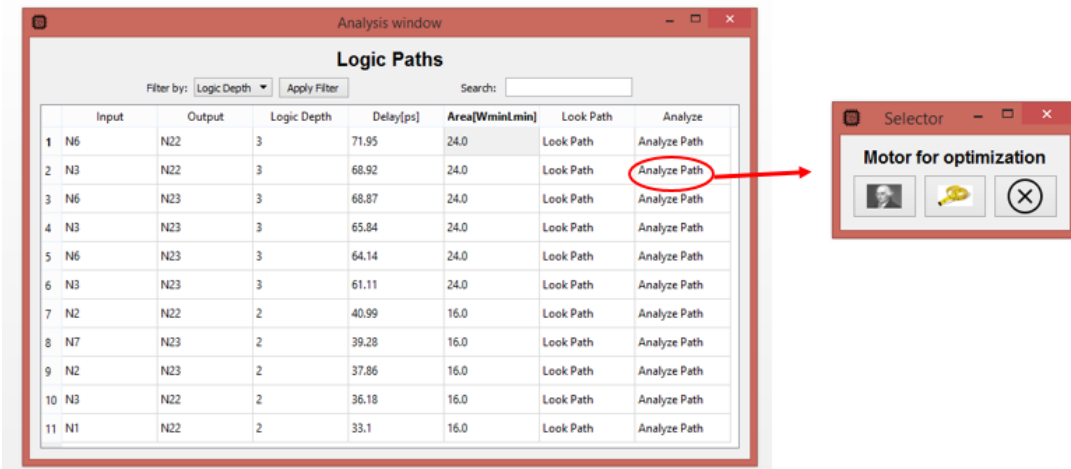


Figura D.14: Interfaz de selectora de caminos: **Analyse Path**

D.4. Interfaz de Optimización de Lagrange

Si se selecciona la opción de Lagrange en la ventana de **Analyse Path**, se abrirá la siguiente interfaz.

En esta interfaz puede reconocerse tres partes aisladas.

1. Analizador del camino.
2. Configuración y generación de optimización.
3. Resultados de optimización.

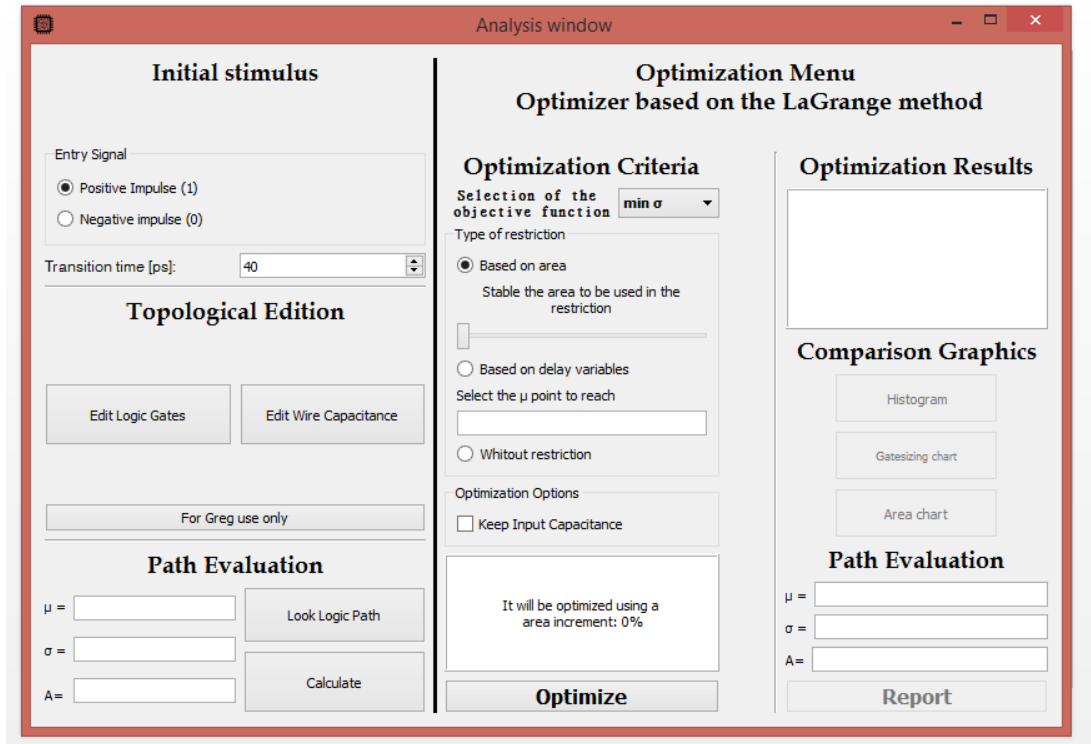


Figura D.15: Interfaz de selectora de caminos

D.4.1. Analizador del camino

Esta sección esta compuesta a la vez por tres elementos principales.

- **Initial stimulus:** Selecciona si el camino se va estudiar para una conmutación generada por un impulso negativo o positivo, y permite describir el tiempo de transición para dicha señal.
- **Topological Edition:** Esta sección se divide a la vez en dos partes
 1. **Edit Logic Gate:** Este botón abrirá una ventana donde se podrán editar las escalas (K) de las compuertas del camino en tiempo real.
 2. **Edit Wire Capacitance:** Este botón abrirá una ventana que permite agregar, eliminar o modificar capacitancias externas a los nodos internos del camino.
- **Path Evaluation:** Esta sección se divide a la vez en dos partes
 1. **Look Logic Path:** Muestra el esquemático del camino bajo análisis.

2. **Calculate:** Cacula el retardo nominal, su desviación estándar y el área ocupada por el camino, los resultados se muestran en la parte izquierda de esta sección.

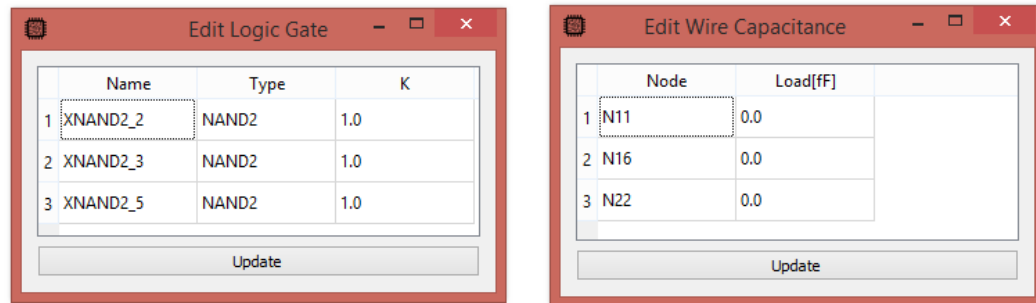


Figura D.16: Ventanas **Edit Logic Gate** y **Edit Wire Capacitance**

D.4.2. Configuración y generación de optimización

Esta sección esta compuesta de cuatro elementos principales, los cuales serán descritos a continuación.

Selection of the objective function

Este es un menú desplegable que permite seleccionar una de las cinco funciones objetivo a minimizar.

Dichas funciones objetivo son:

- La desviación estándar del retardo (σ)
- El valor nominal del retardo (μ)
- La variabilidad del retardo (σ/μ)
- El peor caso ($\mu + 3\sigma$)
- El área del camino lógico (Area)

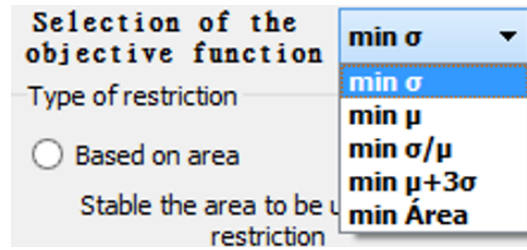


Figura D.17: Selection of the objective function

Sección Type of restriction y Optimization Options

La sección **Type of restriction** permite escoger la restricción de optimización, es una interfaz inteligente que se adapta a la F.O objetivo seleccionada en la sección anterior.

Generalmente deja escoger entre dos tipos de restricciones:

- Basada en área: Especificando el área disponible para la optimización (A_C) por medio de un selector deslizante
- Basada en una variable de retardo especificada por medio de un Traget (T_D).

Como se menciona es una interfaz inteligente que puede o no desplegar todas sus opciones de acuerdo a la F.O considerada, por ejemplo esta interfaz no dejara seleccionar una restricción de área para la F.O Area.

La sección **Optimization Options** permite configurar si la capacitancia de entrada del camino C_{in_1} permanecerá inalterada por la optimización.

Para guiar al usuario en la parte inferior se muestra una ventana inteligente que especifica los detalles de la optimización configurada.

Type of restriction

☐ Based on area
Stable the area to be used in the restriction

☒ Based on delay variables
Select the μ point to reach

☐ Whitout restriction

Figura D.18: **Type of restriction**

Optimization Options

☐ Keep Input Capacitance

Figura D.19: **Optimization Options**

Una vez configurada la optimización esta podrá llevarse acabo presionando el botón **Optimize**, los resultados de la optimización se mostraran del lado derecho de la herramienta.

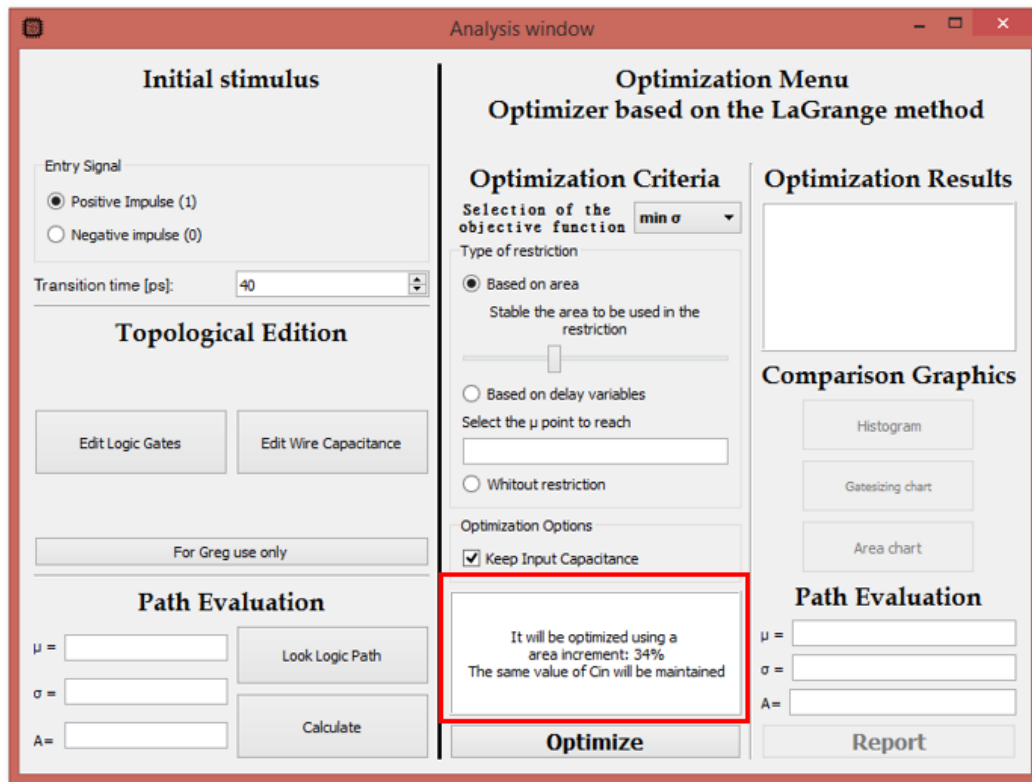


Figura D.20: Mensaje de optimización configurada

D.4.3. Resultados de optimización

Esta sección esta compuesta de cuatro elementos principales, los cuales serán descritos a continuación.

Optimization Results

Cuando se lleva acabo una optimización se mostrara el resultado de la corrida en la parte superior derecha del área de trabajo.

Esta ventana informara si la optimización se llevo correctamente o si hubo algún inconveniente dentro de la misma (por ejemplo un target inalcanzable).

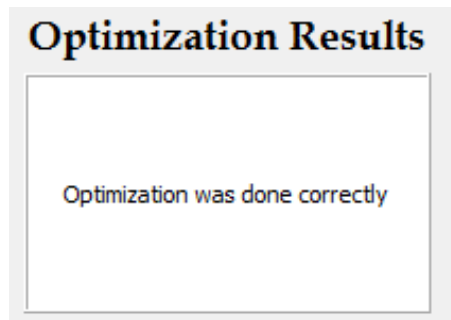


Figura D.21: Optimization Results

Comparison Graphics

Esta parte de la herramienta esta dedicada a generar graficos de comparación entre el camino antes y después de la optimización.

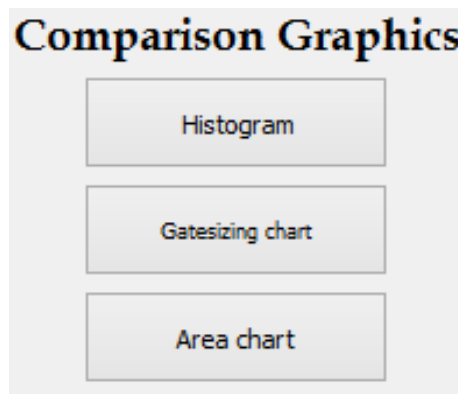


Figura D.22: Comparison Graphics

Los gráficos que pueden ser generados se discuten a continuación:

- **Histogram:** Este botón permite generar un histograma para el camino antes y después de la optimización. En el eje Y del histograma se tiene la probabilidad de obtener un retardo específico, mientras que en el eje X se encuentran los valores de retardo.
- **Gatesizing Chart:** Permite comprar los tamaños de las compuertas individuales antes y después de la optimización. En el eje Y se tiene el valor de escala de la compuerta y en el X cada compuerta específica del camino.

- **Area Chart:** Permite comprar el área total del camino antes y después de la optimización.

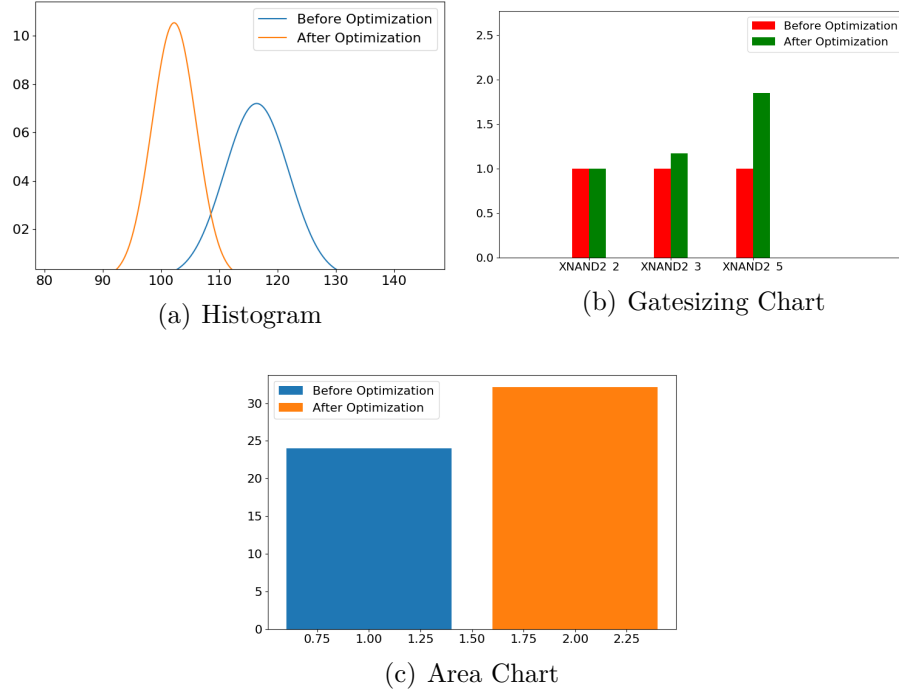


Figura D.23: Comparison Graphics

Botón Report

Permite generar un archivo de texto con los resultados de la optimización.

En este archivo se muestra:

1. Datos básicos del camino lógico analizado, profundidad, nodo de entrada, nodo de salida y compuertas dentro del camino.
2. Comparativa del σ_P , D_P y área del camino antes y después de la optimización.
3. Para cada compuerta se muestra:

- Escala (K),

- Retardo nominal (D_P)
- σ_D de la compuerta
- Área utilizada por la compuerta
- Capacitancia de entrada
- Auto capacitancia,
- Capacitancia de carga
- Contribución de cada variación al valor del σ_D .

Esto es mostrado para antes y después de la optimización

4. La raíz cuadrada de la covarianza entre todo par de compuertas adyacentes debido a la pendiente de entrada, esto es mostrado para antes y después de la optimización
5. Capacitancias total en los nodos del camino antes y después de la optimización
6. Matriz de correlación espacial en caso de incluirse en el análisis.

El reporte puede generarse en ingles o español por medio de la sección opciones del archivo ISCAS.

Un ejemplo de este reporte se muestra a continuación.

```

1
2 Optimization report
3 Analysis time 05:45:43
4 The input signal in the path is a logic 0
5 Path under test
6   Input node: N6
7   Output node: N22
8   Logic Depth: 3
9   Squence of nodes: ['N6', 'N11', 'N16', 'N22']
10  Gates in path:
11    Name: XNAND2_2
12    Type: NAND2
13    Name: XNAND2_3
14    Type: NAND2
15    Name: XNAND2_5
16    Type: NAND2
17 Data of delay in the path
18 Before optimization
19   Nominal Delay: 116.45 [ps]
```

```

20     Standar deviation of delay: 5.54 [ps]
21     Area: 24.0[W*L]
22 After optimization
23     Nominal Delay: 102.28 [ps]
24     Standar deviation of delay: 3.79 [ps]
25     Area: 32.16[W*L]
26 Data about the logic gates
27 Gate: 1
28     Name: XNAND2_2
29     Type: NAND2
30     Input node: N6
31     Output node: N11
32     Layout position:
33         X: 0
34         Y: 0
35 Data about the delay in gate
36     Before optimization
37         Scale factor[K]: 1.0
38         Nominal delay: 28.55 [ps]
39         Standar deviation of delay: 1.58 [ps]
40         Area: 8.0[W*L]
41         Input Capacitance: 0.6 [fF]
42         Self Capacitance: 0.68 [fF]
43         Fan out Capacitance: 1.2 [fF]
44         Independent contribution of process variation in the
standard deviation:
45             VT contribution: 1.28 [ps]
46             W contribution: 0.29 [ps]
47             L contribution: 0.75 [ps]
48             TOX contribution: 0.46 [ps]
49             SR contribution: 0.0 [ps]
50     After optimization
51         Scale factor[K]: 1.0
52         Nominal delay: 29.13 [ps]
53         Standar deviation of delay: 1.61 [ps]
54         Area: 8.0[W*L]
55         Input Capacitance: 0.6 [fF]
56         Self Capacitance: 0.68 [fF]
57         Fan out Capacitance: 1.27 [fF]
58         Independent contribution of process variation in the
standard deviation:
59             VT contribution: 1.3 [ps]
60             W contribution: 0.3 [ps]
61             L contribution: 0.77 [ps]
62             TOX contribution: 0.47 [ps]
63             SR contribution: 0.0 [ps]

```

```

64  Gate: 2
65  .....
66  .....
67  .....
68  Gate: 3
69  .....
70  .....
71  .....
72  Contribution to the standard deviation produce by the input
    variation
73  Before the optimization
74  Square root of the covariance contribution in between
    adjacent gates: XNAND2_2 y XNAND2_3 : 1.35 [ps]
75  Square root of the covariance contribution in between
    adjacent gates: XNAND2_3 y XNAND2_5 : 3.03 [ps]
76  After the optimization
77  Square root of the covariance contribution in between
    adjacent gates: XNAND2_2 y XNAND2_3 : 1.4 [ps]
78  Square root of the covariance contribution in between
    adjacent gates: XNAND2_3 y XNAND2_5 : 1.66 [ps]
79
80  Spacial correlation matrix
81  |1.00    0.99    0.99|
82  |0.99    1.00    1.00|
83  |0.99    1.00    1.00|
84
85  Nodal capacitnace comparation
86  Before the optimization
87  Node: N11
88  Value: 1.88 [fF]
89  Node: N16
90  Value: 1.88 [fF]
91  Node: N22
92  Value: 5.68 [fF]
93  After the optimization
94  Node: N11
95  Value: 1.95 [fF]
96  Node: N16
97  Value: 2.34 [fF]
98  Node: N22
99  Value: 6.13 [fF]
100 End of report

```

D.5. Interfaz del Método Heurístico

Si se selecciona la opción de Método Heurístico en la ventana de **Analyse Path**, se abrirá la siguiente interfaz.

The screenshot shows the 'Analysis window' for the Heuristic Method. It features a top bar with the title 'Analysis window'. The main area is divided into several functional sections:

- Initial stimulus:** Includes 'Entry Signal' with radio buttons for 'Positive impulse (1)' (selected) and 'Negative impulse (0)'. Below it is a 'Transition time [ps]' input field set to 40.
- Topological Edition:** Contains buttons for 'Edit Logic Gates', 'Edit Wire Capacitance', and a 'For Greg use only' button.
- Path Evaluation:** Includes input fields for μ , σ , and A , along with 'Look Logic Path' and 'Calculate' buttons.
- Optimization Menu:** Features the title 'Optimizer based on the Heuristic method'.
- Optimization Criteria:** Includes checkboxes for 'Keep Input Capacitance' and 'Stable the area to be used in the restriction'. It also has input fields for '# Max iterations' (1000) and '# Max number of gates to be modified' (1000).
- Optimization Results:** A large empty box for displaying results.
- Graficos de Comparación:** Includes buttons for 'Histogram', 'Gatesizing chart', and 'Area chart'.
- Path Evaluation (bottom right):** Includes input fields for μ , σ , and A , and a 'Report' button.
- Central Status:** A box indicating 'It will be optimized using a area increment: 0%' and an 'Optimize' button at the bottom center.

Figura D.24: Interfaz del Método Heurístico

Como puede apreciarse esta interfaz es prácticamente idéntica a la interfaz de optimización de Lagrange, las únicas diferencias notables son encontradas en la sección de en medio, bajo la designación **Optimization Criteria**.

A continuación se explican los parametros de la sección **Optimization Criteria**.

- **Keep Input Capacitance:** Mantiene la capacitancia de entrada del camino constante durante el proceso de optimización.
- **Max Iterations:** Numero de iteraciones máximas por el método Heurístico.
- **Max number of gates to be modified:** Establece el numero de compuertas factibles que el algoritmo puede incrementar, si el numero especificado es mayor que el numero de compuertas en el camino, todas las compuertas se consideran factibles.
- **Delta:** Incremento en el valor de K para la compuerta seleccionada en cada iteración.
- **Stable the area to be used in restriction:** Especifica el área maxima a la cual puede crecer el camino (A_C), este proceso se realiza por medio del selector deslizante, para guiar al usuario en la parte inferior se muestra una ventana inteligente que especifica los detalles de la optimización configurada.
- **Heuristic metric selection:** Permite escoger como métrica de la compuerta la derivada de la propia compuerta o la métrica propuesta en [11].

El resto de opciones de esta interfaz son idénticas a las encontrada en la interfaz de LaGrange, por lo que no se explicaran en esta sección.

Índice de figuras

1.1. Ley de Moore	5
1.2. Trade-off del escalamiento tecnológico	5
1.3. Impacto de las variaciones de proceso [17]	6
1.4. Efecto LER [5]	7
1.5. Variaciones Globales [5]	9
1.6. Variaciones Correlacionadas [17]	10
1.7. Numero de dopantes para diferentes nodos tecnológicos [16]	11
1.8. Esquinas de diseño [5]	12
2.1. Modelo eléctrico de una compuerta lógica	17
2.2. Inversor CMOS	18
2.3. Esquemático de la compuerta NAND2	20
2.4. Circuito de prueba para obtener la resistencia equivalente del transistor	22
2.5. Layout sencillo de un MOSFET	26

2.6. Modelo distribuido de una compuerta lógica	31
2.7. Stack de tres transistores nMOS mostrando todas las capacitancias parásitas de la red	43
2.8. Evolución temporal de los nodos intermedios	44
2.9. Evolución temporal de los valores de V_{DS}	45
2.10. Estructura NAND2 y su nodo intermedio	47
2.11. Formas de onda en una estructura NAND2	47
2.12. Comparación entre el modelo analítico del σ_D y SPICE, en el Inversor	52
2.13. Error en función de C_{self} en el σ_D para el Modelo Analítico, probado en un inversor con $K=2$	53
2.14. Ejemplo de camino lógico	53
2.15. Comportamiento de la correlación entre compuertas [17]	55
2.16. Camino de validación de la covarianza	57
2.17. Validación de la covarianza	57
3.1. Familia de curvas de la desviación estándar en un inversor obtenidas por el modelo analítico propuesto	60
3.2. Comparación del σ_D en las compuertas básicas con $C_L=4[\text{fF}]$. . .	60
3.3. Comparación entre la derivada del σ_D del modelo analítico y SPICE, en un inversor con $C_L = 4[\text{fF}]$	62
3.4. Ejemplo de camino lógico	66
3.5. Diagrama de flujo del proceso de optimización	68

3.6. Espacio de optimización para el camino de prueba con $A_C = 27$.	69
3.7. Resultados de optimización base	70
3.8. Comparación del σ_D entre la optimización base y discreta	72
3.9. Espacio de optimización con correlación espacial	73
3.10. Resultados de optimización base	74
3.11. Circuito bajo prueba	76
3.12. Espacio de optimización considerado	77
3.13. Resultados de optimización para el circuito bajo prueba	78
3.14. Comparación entre los dos circuitos especificados en la Tabla 3.9 .	79
3.15. Espacio de optimización conservando el valor de C_{in} de la primera compuerta	80
3.16. Diagrama de flujo para la optimización conservando el valor de C_{in} de la primera compuerta	81
3.17. Resultados de optimización conservando el valor de C_{in}	82
3.18. Repetición de la optimización anterior con un criterio de paro mas fino	83
4.1. Comparación del σ_D con respecto a la escala de la compuerta, en un inversor con $C_L = 4[\text{fF}]$	86
4.2. Comparación del σ_D con respecto a la carga de la compuerta, en un inversor con $K = 1$	86
4.3. Comparación del σ_D con respecto al Stack con $C_L = 2[[\text{fF}]$	87
4.4. Valores de la derivada de la desviación estándar del retardo para una compuerta lógica	89

4.5. Vista en cortes de la Figura 4.4	89
4.6. Comparación entre la tendencia de la métrica y la derivada completa, con $a = 2,5$	90
4.7. Cadena de inversores bajo redimensionamiento	91
4.8. Comportamiento del camino lógico anterior bajo un dimensionamiento indiscriminado de K_3	92
4.9. Covarianza y coeficiente de correlación para dos compuertas inversoras [5]	93
4.10. Impacto de la profundidad en el σ_P bajo condición de covarianza .	94
5.1. Diagrama de flujo compacto de la herramienta CAD	100
5.2. Gráficos generados por la herramienta CAD	102
5.3. Diseño factorial de dos y tres factores	105
5.4. Diseño factorial con dos factores	106
5.5. Superficie de respuesta	108
5.6. Metodología de caracterización de cantidades nominales	111
5.7. Resultados de la metodología de caracterización nominal, los valores numéricos son únicamente con fin de ilustración	112
5.8. Metodología de caracterización de la sensibilidad	113
5.9. Comparación entre los polinomios caracterizados y SPICE en el Inversor	116
5.10. Comparación entre los polinomios caracterizados y SPICE en la NAND2	117

5.11. Comparación entre los polinomios caracterizados y SPICE en la NOR2	118
5.12. Camino básico para validación	119
5.13. Camino compuesto para validación	120
5.14. El gradiente es un campo vectorial cuyos elementos siempre apuntan en torno al máximo de la función	123
5.15. Diagrama de bloques de proceso de optimización	124
5.16. Ejemplo de camino lógico, donde $C_1 = 2\text{fF}$, $C_2 = 2\text{fF}$, $C_3 = 4\text{fF}$, $C_4 = 2\text{fF}$, $C_5 = 5\text{fF}$	128
5.17. Comparación de resultados	129
5.18. Yield devuelto por las diferentes funciones objetivos bajo un incremento de área pequeño.	130
5.19. Histograma para optimización con incremento de área considerable	131
5.20. Yield para optimización con incremento de área considerable . . .	132
5.21. Histograma para optimización sin restricciones	133
5.22. Yield para optimización sin restricción	133
5.23. Ejemplo de optimización minimizando el área	135
5.24. Diagrama de flujo para optimización heurística	137
5.25. Ejemplo de un circuito digital y algunos de sus caminos lógicos . .	138
5.26. Grafo dual del circuito digital anterior	139
A.1. Comparación entre el modelo analítico del σ_D y SPICE, en la compuerta NAND2	157

A.2. Comparación entre el modelo analítico del σ_D y SPICE, en la compuerta NOR2	158
A.3. Proceso completo de la transformación de Cholesky	159
B.4. Diferentes ajuste de curva por medio del parámetro a_1	162
B.5. Diferentes ajuste de curva por medio del parámetro a_2	164
C.6. Solución grafica del problema de optimización	168
C.7. Solución grafica del problema de optimización	168
D.8. Interfaz principal de la herramienta SSTA	174
D.9. Interfaces de compilación	175
D.10.Utilización de la librería PRIMITIVAS_SPICE.tanner en S-Edit	177
D.11.Funcionamiento del Path Selector	179
D.12.Interfaz de selectora de caminos	180
D.13.Interfaz de selectora de caminos: Look Path	181
D.14.Interfaz de selectora de caminos: Analyse Path	181
D.15.Interfaz de selectora de caminos	182
D.16.Ventanas Edit Logic Gate y Edit Wire Capacitance	183
D.17.Selection of the objective function	184
D.18. Type of restriction	185
D.19. Optimization Options	185

D.20.Mensaje de optimización configurada	186
D.21.Optimization Results	187
D.22.Comparison Graphics	187
D.23.Comparison Graphics	188
D.24.Interfaz del Método Heurístico	192

Índice de tablas

2.1. Factor de dimensionamiento en las compuertas digitales básicas. .	21
2.2. Parámetros para la tecnología de 65 nm	21
2.3. Comparación del parámetro eléctrico de la resistencia de canal entre el modelo desarrollado y SPICE	22
2.4. Comparación del parámetro eléctrico de la capacitancia de entrada para un inversor entre el modelo desarrollado y SPICE	24
2.5. Número de transistores en el nodo de salida para las topologías básicas.	27
2.6. Comparación del parámetro eléctrico de la auto-capacitancia para un inversor entre el modelo desarrollado y SPICE	28
2.7. Comparación de la resistencia de canal predicha por el modelo analítico y SPICE	48
2.8. Valor de ξ para las topologías NAND	48
3.1. Proceso de optimización iterativo	69
3.2. Comparación de valor optimo bajo restricción de área sin considerar correlación espacial	70
3.3. Proceso de optimización iterativo con paso discreto	71

3.4. Comparación entre la optimización base y discreta	72
3.5. Proceso de optimización iterativo con correlación espacial	74
3.6. Comparación entre la optimización base y con correlación espacial	75
3.7. Comparación de valor optimo bajo restricción de área considerando correlación espacial	75
3.8. Proceso de optimización del circuito bajo prueba	78
3.9. Comparación de valor óptimo para circuito de prueba	79
3.10. Proceso de optimización conservando el valor de C_{in}	81
3.11. Comparación de valor optimo conservando el valor de C_{in}	82
3.12. Repetición de la comparación de valor optimo conservando el valor de C_{in} con un criterio de para mas fino	83
5.1. Ejemplo de diseño factorial	105
5.2. Parámetros a caracterizar	110
5.3. Variables a examinar	111
5.4. Comparación en el camino básico	119
5.5. Comparación del camino lógico compuesto	121
5.6. Comparación de resultados	128
5.7. Cambios entre las variables de retardo para las optimizaciones bajo restricción de área $A_C = 1,1A_{P0}$	129
5.8. Comparación de resultados	131
5.9. Comparación de resultados	132

5.10. Comparación de resultados	138
---	-----

Bibliografía

- [1] Ivan E. Sutherland, Bob F. Sproull, David L. Harris Logical Effort: Designing Fast CMOS Circuits 1ed. Morgan Kaufman 1999
- [2] Neil H. E. Weste, David Money Harris. CMOS VLSI Design A Circuits and Systems Perspective 4ed. Addison-Wesley 2011
- [3] John P. Uyemura. CMOS Logic Circuit Design 1ed. Kluwer Academic Publishers 2002
- [4] Behzad Razavi. Fundamentals of Microelectronics 2ed. Wiley 2014
- [5] Victor Champac, Jose Garcia Gervacio Timing Performance of Nanometer Digital Circuits Under Process Variations 1ed. Springer 2018
- [6] David Comer Diseño de circuitos electrónicos 1ed. Limusa Wiley 2000
- [7] William Stallings Organización y arquitectura de computadoras 7ed. Pearson 2005
- [8] Thomas L. Floyd Dispositivos Electrónicos 8ed. Pearson 2008
- [9] Manfred Dietrich Joachim Haase Process Variations and Probabilistic Integrated Circuit Design 1ed. Springer 2012
- [10] Yannis Tsividis Colin McAndrew Operation and Modeling of the MOS Transistor 3ed. OXFORD UNIVERSITY PRESS 2011
- [11] V. Champac H. Villacorta Z. Perez Circuit performance optimization for local intra-die process variations using a gate selection metric. Electronics letters
- [12] Douglas C. Montgomery Design and Analysis of Experiments 8ed. McGraw-Hill Education 2013
- [13] Douglas C. Montgomery Applied Statistics and Probability for Engineers 6ed. WILEY 2013

- [14] James Stewart Calculus: Early Transcendentals 6ed. Thomson Brooks/Cole 2008
- [15] Alejandra Nichte-Ha Reyes Flores Análisis y Optimización de circuitos digitales nano-métricos en presencia de variaciones de proceso INAOE 2014
- [16] Martin Wirnshofer Variation-Aware Adaptive Voltage Scaling for Digital CMOS Circuits Springer 2014
- [17] Andrés Felipe Gómez Chacón Reliability Analysis and Improvement of Nanoscale CMOS Digital Circuits INAOE 2017