

Instituto Tecnológico de Costa Rica

Escuela de Ingeniería Electrónica



**DISEÑO DE UN AMPLIFICADOR OPERACIONAL
DE TRANSCONDUCTANCIA PARA LA
IMPLEMENTACIÓN DE FILTROS ANALÓGICOS
UTILIZADOS EN LA DETECCIÓN DE DISPAROS
DE ARMAS DE FUEGO**

Informe de Proyecto de Graduación para optar por el título
de Ingeniero en Electrónica con el grado académico de
Licenciatura

Frank Andrey Nicaragua Guzmán

Cartago, Junio 2010

INSTITUTO TECNOLÓGICO DE COSTA RICA

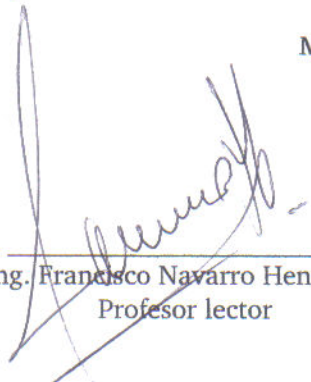
ESCUELA DE INGENIERÍA ELECTRÓNICA

PROYECTO DE GRADUACIÓN

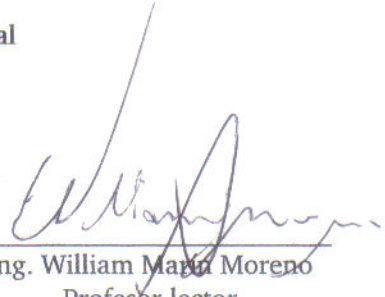
TRIBUNAL EVALUADOR

Proyecto de Graduación defendido ante el presente Tribunal Evaluador como requisito para optar por el título de Ingeniero en Electrónica con el grado académico de Licenciatura, del Instituto Tecnológico de Costa Rica.


Miembros del Tribunal



Ing. Francisco Navarro Henríquez
Profesor lector



Ing. William María Moreno
Profesor lector



Ing. Alfonso Chacón Rodríguez
Profesor asesor

Los miembros de este Tribunal dan fe de que el presente trabajo de graduación ha sido aprobado y cumple con las normas establecidas por la Escuela de Ingeniería Electrónica.

Cartago, Junio 2010

Declaración de Autenticidad

Declaro que el presente Proyecto de Graduación ha sido realizado enteramente por mi persona, utilizando y aplicando literatura referente al tema e introduciendo conocimientos propios.

En los casos en que he utilizado bibliografía, he procedido a indicar las fuentes mediante las respectivas citas bibliográficas.

En consecuencia, asumo la responsabilidad total por el trabajo de graduación realizado y por el contenido del correspondiente informe final.

22 de junio del 2010



Frank Andrey Nicaragua Guzmán

Cédula:1-1306-0314

Resumen

El presente documento describe el proceso de diseño e implementación de un OTA empleado en filtros analógicos de detección de disparos de armas de fuego. Esto forma parte del proyecto de construcción de una red de sensores para parques y reservas nacionales con el fin de auxiliar los esfuerzos de protección y conservación de recursos naturales.

El problema a combatir radica en la necesidad de corregir un error presentado en el circuito integrado de detección, dado que para obtener el comportamiento deseado era necesario aumentar el consumo de potencia causado por los ajustes en la corriente del circuito. Para corregir esto, se diseñó un nuevo OTA mediante cálculos manuales, pero que posteriormente fue optimizado mediante la utilización de una herramienta computacional basada en algoritmos genéticos. El OTA obtenido logró satisfacer todas las especificaciones solicitadas.

Palabras clave: algoritmos genéticos, filtros GmC, OTA, PESA, transconductancia.

Abstract

The aim of this project focused on the design and implementation of an OTA circuit, used in analog filters for gunshot detection. This work is part of the ITCR Electronics Engineering School project to develop a sensor network for national parks that will aid the efforts of protection and conservation of natural resources in Costa Rica.

The IC designed for detection presented undesired behavior. The current of the circuit was adjusted in order to establish proper operation, but this caused an increase in power consumption. In order to fix this, a new OTA was designed by hand-calculation methods and then it was properly optimized using computational software based in genetic algorithms. This OTA fulfilled every specification with improved characteristics.

Keywords: genetic algorithms, GmC filters, OTA, PESA, transconductance.

Dedicatoria

Todos los años de esfuerzo y dedicación reflejados en este trabajo los dedico a mi familia, la cual ha estado presente en los momentos buenos y difíciles de mi vida. A mi máma Damaris, por ser la principal columna de mi vida, la que siempre me ha alentado a seguir luchando por todos mis sueños y la que me ha mostrado lo lejos que se puede llegar cuando uno quiere.

A mi papá Francisco, por su deseo de darnos el mejor futuro y por su contante preocupación y apoyo durante todo este tiempo.

A mis hermanos Harold y Jaikel, quienes han sido mis confidentes más cercanos y para quienes espero haber sido un gran ejemplo de lucha y esfuerzo.

A mi abuela Leita, a mi tía Isabel y a mi prima Carolina, quienes han sido una familia que me ha mostrado muchísimo cariño y también han sido parte muy importante de todo este proceso.

A mis tíos, tías, abuelos y demás familiares, cuya constante admiración es motivo de esforzarme cada vez más y a seguir construyendo aún más logros.

A mis amigos y amigas, que en las diferentes etapas de mi vida impactaron de manera positiva y han dejado su huella en mi corazón.

A Dios Todopoderoso, porque sin Él, nada de esto hubiera sido posible.

Agradecimientos

Primeramente, al Dios Todopoderoso, por ser la razón de mi vida, quien me ha fortalecido y acompañado en todo tiempo y circunstancia, por lo que mi gratitud hacia Él nunca acabará.

A toda mi familia, porque su apoyo ha sido continuo y me han alentado a seguir creciendo aún mucho más. Gracias por su amor y su dedicación hacia mí.

Al profesor Alfonso Chacón Rodríguez por la oportunidad de realizar este proyecto, por la asesoría brindada y por el conocimiento transmitido, lo que ha permitido un mayor crecimiento personal y profesional en mi persona.

Al profesor Roberto Pereira, por la colaboración brindada a lo largo del proyecto y por la oportunidad de generar conocimiento para futuros proyectos.

A los compañeros del TEC, que son varios de enumerar, con los que vivimos alegrías y preocupaciones, proyectos interminables y muchos exámenes más allá de la comprensión de muchos. A mis compañeros de proyecto, Allan y Bryant, y a todos ellos les deseo los mejores éxitos en su carreras y en sus vidas.

Y a todos aquellos amigos, amigas, hermanos y personas que me conocieron más allá de lo que se ve, gracias de todo corazón por su tiempo, admiración y sacrificio en esta etapa de mi vida. Dios les bendiga!

Índice general

1. Introducción	1
1.1. Antecedentes	1
1.2. Problema existente e importancia de su solución	5
1.3. Solución seleccionada y requisitos de diseño	6
2. Marco teórico	7
2.1. Principios de funcionamiento de los transistores MOS	7
2.1.1. Estructura básica	7
2.1.2. Niveles de inversión	8
2.1.3. Efecto de cuerpo (Body effect)	10
2.1.4. Corriente de drenaje I_D	10
2.1.5. Curvas características	11
2.1.6. Convención de signos y símbolos para transistores NMOS y PMOS	15
2.2. Ecuaciones de DC de los transistores MOS	16
2.2.1. Parámetros dependientes de las características físicas del proceso	16
2.2.2. Ecuaciones de aproximación de los límites entre zonas de inversión	17
2.2.3. Ecuaciones de corriente para diversas regiones de operación	18
2.3. Características de pequeña señal de los transistores MOS	19
2.3.1. Parámetros de transconductancia	19
2.3.1.1. Modelo general y el cálculo de transconductancias	21
2.3.2. Parámetros de capacitancia	22
2.4. Principios básicos de los OTA	23
2.5. Esquemas básicos para el diseño de OTAs	24
2.5.1. OTA diferencial básico	24
2.5.2. Principio de degeneración de la fuente	25
2.5.3. Transistor con polarización en modo común	26
2.5.4. OTA con difusor simétrico	26
2.5.5. Par diferencial con fuentes de tensión flotante	28
2.6. Filtros GmC	31
2.7. Espejos de corriente para el escalamiento de transconductancias	32
2.7.1. Asociaciones serie-paralelo de transistores	32
2.7.2. Espejos de corriente utilizando asociaciones serie-paralelo de transistores	34
2.8. Herramientas de diseño y simulación de circuitos integrados	34
2.8.1. Simuladores SPICE y modelos MOS	34
2.8.2. Mentor Graphics	35

2.8.2.1. Design Architect-IC (DA-IC)	37
2.8.2.2. Eldo Simulator y EZwave waveform viewer	38
2.8.2.3. ICstation Layout	39
2.8.2.4. Calibre IC Verification	40
2.9. Algoritmos genéticos y el Frente de Pareto	43
2.9.1. El algoritmo PESA	44
3. Procedimiento metodológico	45
3.1. Descripción y síntesis del problema	45
3.2. Evaluación de alternativas	46
3.2.1. Sustitución del difusor simétrico	46
3.2.2. Software de optimización automática	47
4. Meta y objetivos	49
4.1. Meta	49
4.2. Objetivos	49
4.2.1. Objetivo general	49
4.2.2. Objetivos específicos	49
5. Descripción detallada de la solución	51
5.1. Selección del esquema de linealización	51
5.2. Análisis matemático del OTA con difusor simétrico	56
5.2.1. Análisis de polarización en DC	56
5.2.1.1. Polarización de M_1	57
5.2.1.2. Polarización de M_2	57
5.2.2. Análisis de pequeña señal	58
5.3. Diseño inicial del OTA	61
5.4. Implementación de la herramienta de optimización	63
5.4.1. Parametrización y simulación del circuito	63
5.4.2. Modificaciones a la herramienta de optimización	65
5.5. Pruebas de preselección de resultados	67
5.6. Resultados seleccionados y escalamiento del OTA	69
5.7. Gráficas de corriente, transconductancia y transitorios	74
5.8. Layouts trazados	79
6. Análisis de resultados	85
7. Conclusiones	89
8. Recomendaciones	91
Bibliografía	93
Apéndices	94
Apéndice A: Consideraciones sobre el slew rate	94
Apéndice B: Historial de simulaciones de la herramienta de optimización	95

Índice de figuras

1.1. Banco de filtros paralelos que equivalen a una descomposición continua de ondas. Tomado de [1].	2
1.2. Filtros GmC necesarios para cada etapa de la descomposición continua de ondas. Tomado de [1].	3
1.3. Amplificador operacional de transconductancia básico.	3
1.4. OTA linealizado con difusores simétricos.	4
1.5. Simulaciones del OTA de 137nS. Tomado de [1].	5
2.1. Estructura básica de un transistor NMOS	8
2.2. Conexión de un transistor NMOS para la evaluación de los niveles de inversión	9
2.3. Circuito de evaluación de la corriente de drenaje	11
2.4. Curva característica I_D vs V_{DS} para un transistor específico, con V_{GS} como parámetro y $V_{SB} = 0$	12
2.5. Curvas características I_D vs V_{DS} (escala lineal), con V_{GS} como parámetro y a) $V_{SB} = 0$ y b) $V_{SB} = 2,5V$. Gráficas obtenidas en LTSPICE IV para un transistor cuyo modelo corresponde a un proceso de $0.5\mu\text{m}$	13
2.6. Curvas de I_D (eje logarítmico) e $\sqrt{I_D}$ vs V_{GS} , alineadas con respecto al eje V_{GS}	15
2.7. Símbolos para los transistores a utilizar: a) NMOS ; b) PMOS	16
2.8. Circuito equivalente de pequeña señal para el transistor NMOS	20
2.9. Circuito equivalente de pequeña señal para un transistor NMOS, con capacitancias incluidas	22
2.10. a) Representación del OTA ideal. b) Circuito equivalente de pequeña señal.	23
2.11. OTA diferencial básico.	24
2.12. OTA linealizado por medio de un elemento resistivo.	25
2.13. OTA con transistor con polarización en modo común	26
2.14. OTA con difusor simétrico	27
2.15. Gráficas de I_{out} vs $V_{in}(V_d)$ y G_m vs V_{in} para las configuraciones: a) par diferencial básico. b) transistor en modo común. c) OTA con difusor simétrico. Tomado de [2].	28
2.16. Principio de suma constante de tensiones entre source y gate	28
2.17. Par diferencial con fuentes de tensión flotante	29
2.18. Implementación de Nedungadi del par diferencial con fuentes de tensión flotante.	30
2.19. Gráficas de i_{out} vs V_d para diferentes valores de n (1, 2, 5, 10 y 20, respectivamente). Tomado de [2]	31
2.20. Filtro GmC de primer orden	31

2.21. Configuración de transistores en paralelo y su equivalente.	33
2.22. Configuración de transistores en serie y su equivalente.	33
2.23. Espejo de corriente utilizando asociaciones de transistores serie-paralelo.	34
2.24. Estructura general del ambiente de desarrollo de Mentor Graphics. Tomado de [3].	36
2.25. Interfaz principal del ICstudio.	37
2.26. Interfaz del Design Architect.	38
2.27. Interfaz del EZwave waveform viewer.	39
2.28. Interfaz del ICstation Layout.	40
2.29. Calibre nmDRC.	41
2.30. Calibre nmLVS.	42
2.31. Calibre PEX.	42
2.32. Frente de Pareto. Tomado de [4]	43
3.1. OTA con transistor con polarización en modo común.	46
3.2. Arquitectura del optimizador de circuitos.	48
5.1. OTA simple con transistores PMOS	52
5.2. OTA con transistor en modo común, con transistores PMOS	52
5.3. OTA con difusor simétrico, implementado con transistores PMOS	53
5.4. OTA de Nedungadi utilizando transistores PMOS	53
5.5. Circuito de prueba para OTAs para la obtención de la transconductancia	54
5.6. Gráfica de G_m versus V_d para los cuatro esquemas de OTA a comparar: a) OTA simple; b) OTA con transistor en modo común; c) OTA con difusores simétricos y d) OTA de Nedungadi	54
5.7. Polarización DC de la etapa de entrada diferencial del OTA	56
5.8. Circuito equivalente de la etapa de entrada (análisis de pequeña señal)	59
5.9. OTA inicial implementado.	62
5.10. Prueba de DC para los OTAs.	64
5.11. Prueba de transitorio para los OTAs.	65
5.12. Diagrama de flujo para el método del cálculo del rango lineal	66
5.13. Frente de Pareto obtenido para una de las simulaciones realizadas.	68
5.14. Módulo del transistor M_1	69
5.15. Módulo del transistor M_2	70
5.16. Sub-módulo del transistor M_2	70
5.17. Módulo del transistor M_3	71
5.18. Módulo del transistor M_5	71
5.19. Esquemático del OTA de 16nS.	72
5.20. Esquemático del OTA de 32nS.	73
5.21. Esquemático del OTA de 64nS.	74
5.22. i_{out} vrs V_d para el OTA de 16nS.	75
5.23. i_{out} vrs V_d para el OTA de 32nS.	75
5.24. i_{out} vrs V_d para el OTA de 64nS.	76
5.25. G_m vrs V_d para el OTA de 16nS.	76
5.26. G_m vrs V_d para el OTA de 32nS.	77
5.27. G_m vrs V_d para el OTA de 64nS.	77
5.28. Gráficas de i_{out} y V_d en función del tiempo para el OTA de 16nS.	78

5.29. Gráficas de i_{out} y V_d en función del tiempo para el OTA de 32nS.	78
5.30. Gráficas de i_{out} y V_d en función del tiempo para el OTA de 64nS.	79
5.31. Layout del OTA de 16nS.	80
5.32. Layout del OTA de 32nS.	81
5.33. Layout del OTA de 64nS.	82
5.34. Gráficas de i_{out} y V_d en función del tiempo para el layout del OTA de 16nS.	83
5.35. Gráficas de i_{out} y V_d en función del tiempo para el layout del OTA de 32nS.	84
5.36. Gráficas de i_{out} y V_d en función del tiempo para el layout del OTA de 64nS.	84

Índice de tablas

1.1. Especificaciones mínimas para los OTAs.	6
1.2. Especificaciones de slew rate para los OTAs a diseñar.	6
2.1. Aproximaciones para la transconductancia en inversión fuerte y débil	21
5.1. Valores teóricos y prácticos de G_m para cada configuración de OTA.	55
5.2. Resumen de las características de los OTAs simulados	55
5.3. Parámetros de diseño del OTA inicial.	62
5.4. Especificaciones del OTA inicial.	62
5.5. Valores de aptitud asociados al OTA.	67
5.6. Preselección de los mejores resultados dados por la herramienta de optimización.	68
5.7. Dimensiones de los transistores unitarios empleados en los OTAs	69
5.8. Resumen de las características de los OTAs diseñados	79
5.9. Comparación entre las especificaciones de los OTAs obtenidos y las especificaciones iniciales.	79
5.10. Dimensiones físicas de los layouts generados	83

Capítulo 1

Introducción

El presente apartado introduce al lector a las circunstancias presentes que fomentaron la necesidad de desarrollar el proyecto en cuestión. Asimismo se presenta el problema suscitado debido a los acontecimientos anteriores y finalmente se muestra brevemente la solución implementada para la solución del mismo.

1.1. Antecedentes

La Escuela de Ingeniería Electrónica del Instituto Tecnológico de Costa Rica se encuentra trabajando en un proyecto para desarrollar una red de sensores inalámbricos que permitan detectar cierto tipo de anomalías en reservas naturales de Costa Rica, otorgándole así a las autoridades pertinentes un sistema de vigilancia eficiente que les permita actuar con mayor prontitud.

La red de sensores está compuesta por una serie de nodos o dispositivos remotos capaces de interpretar la información percibida para determinar cuáles eventos se están llevando a cabo, principalmente incendios, tala de árboles con motosierras y disparos de cazadores ilegales.

El procesamiento de señales a nivel de hardware es complejo debido a que el desarrollo de circuitos que realicen las funciones de detección, clasificación y localización requiere de más tiempo que el desarrollo de algoritmos que ejecuten las mismas operaciones sobre sistemas embebidos ya existentes.

No obstante, el desarrollo de un circuito integrado con una aplicación específica (ASIC, por sus siglas en inglés) presenta la ventaja de un menor consumo de potencia al destinar sus recursos a una labor en particular. Tanto la ejecución de los algoritmos como las diversas funciones que ejecutan los sistemas embebidos generan un consumo mucho más elevado de potencia con respecto a un ASIC. Este factor es crítico debido a que los nodos de la red cuentan únicamente con alimentación limitada, generalmente por medio de una pila y deben ubicarse en zonas de difícil acceso, por lo que para extender su tiempo de vida útil es necesario reducir al mínimo el consumo de energía.

El desarrollo de cada nodo se ha dividido en tres etapas: detección, clasificación y localización. Cada una de ellas requiere la implementación de un módulo diseñado de la manera más eficiente y compacta posible, que permitan su integración en un circuito integrado para la construcción final de los dispositivos de vigilancia. Las etapas de clasificación y localización son las que requieren mayor procesamiento, por lo que estas trabajan solo cuando el módulo de detección determina que un evento realmente ha ocurrido.

Para la etapa de detección, el Ing. Alfonso Chacón Rodríguez diseñó un circuito integrado de bajo consumo para la detección de disparos de armas de fuego. De entre varios métodos matemáticos que permiten la detección y clasificación de señales, se utilizó la transformada continua de onditas (CWT o wavelets), ya que poseen una implementación física sencilla al basarse en la utilización de filtros paralelos para obtener la transformación correspondiente.

Utilizando las mediciones de disparos realizadas por los profesores Néstor Hernández y Pablo Alvarado, se dividió el espectro tiempo-frecuencia que describe el disparo de un arma de fuego en ocho bandas o coeficientes por medio de un análisis de ondas discretas. Aplicando un análisis estadístico y una transformada continua de onditas equivalente (CWT o continuous wavelet transformation, por sus siglas en inglés), se determinó que con solo tres de esas bandas de frecuencia (las de coeficiente 3, 4 y 5) era posible obtener entre un 75 % y un 89 % de verdaderos positivos, es decir, que se detectó que la señal recibida era efectivamente un disparo. La Figura 1.1 muestra la implementación de la CWT calculada por medio de un banco de filtros paralelos:

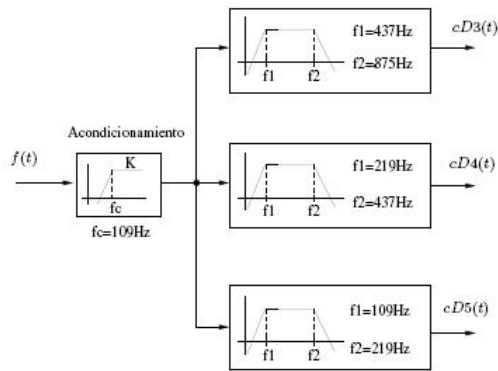


Figura 1.1: Banco de filtros paralelos que equivalen a una descomposición continua de onditas. Tomado de [1].

La implementación de estos filtros puede realizarse con transconductores y condensadores de manera sencilla (filtros GmC), reduciendo no solo la complejidad de la etapa de detección, sino también el consumo de energía al utilizar pocos componentes.

Estos filtros fueron construidos con tecnología CMOS estándar de $0.5\mu\text{m}$ junto con técnicas de diseño de bajo consumo. El diseño de los condensadores y transconductores de este proyecto presentaron dos desafíos principales:

1. La tecnología CMOS permite obtener condensadores en el orden de picofaradios, lo que implicó que para los rangos de frecuencia obtenidos fueran necesarias transconductancias en el orden de nano Siemens (o bien, resistencias en el orden de mega ohmios).
2. La implementación directa de resistencias en circuitos integrados requiere un espacio muy grande, por lo que se requieren mecanismos alternativos para obtener estos transconductores.

La Figura 1.2 muestra el banco de filtros GmC necesario para la obtención de la CWT, especificando el valor de los componentes a utilizar.

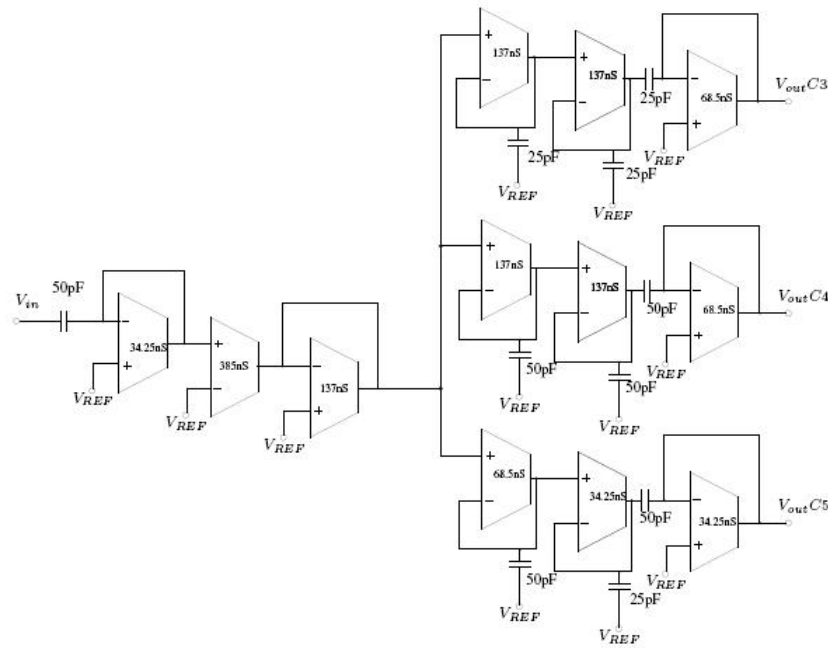


Figura 1.2: Filtros GmC necesarios para cada etapa de la descomposición continua de ondas. Tomado de [1].

Las transconductancias se obtuvieron mediante amplificadores operacionales de transconductancia (OTA) y la conexión realimentada mostrada en la Figura 1.2. Un OTA básico consiste en un amplificador diferencial cuya salida se da en términos de corriente, tal y como muestra la Figura 1.3. Estos dispositivos utilizan transistores en configuración diferencial para lograr un comportamiento en función de la diferencia entre tensiones de entrada aplicadas a estos transistores. En algunos casos se utilizan transistores MOSFET debido a que presentan alta impedancia de entrada, pueden utilizar baja potencia y tienen un comportamiento característico según la zona de operación donde se les ubique.

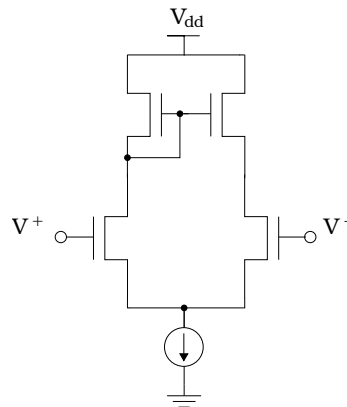


Figura 1.3: Amplificador operacional de transconductancia básico.

Los transistores MOSFET presentan tres zonas de operación: inversión débil, inversión moderada e inversión fuerte. En inversión débil, los transistores presentan una respuesta exponencial y sus rangos lineales son muy bajos (de algunos mV). En inversión fuerte, el transistor presenta una respuesta cuadrática, pero con la posibilidad de obtener rangos lineales en el orden de voltios inclusive, aunque con la desventaja de que presenta un consumo elevado de potencia. El modo de inversión moderada es el que presenta un compromiso entre consumo y linealidad, por lo que este modo de operación es el que se utiliza usualmente en aplicaciones de bajo consumo que requieran un buen desempeño sin elevar los requerimientos de potencia de los circuitos, según se sugiere en [5].

El problema principal que presenta el OTA básico es que su curva de transconductancia no es lineal debido a las características de los transistores. Para lograr una respuesta más lineal, se coloca una resistencia entre los transistores de entrada para que la característica lineal de la resistencia sea la que domine frente a las propiedades de los transistores.

Como se mencionó anteriormente, la colocación de resistencias en circuitos integrados no es deseable debido al gran tamaño de las mismas y es por esta razón que para emular la presencia de un elemento resistivo, se utiliza una variación del circuito de la Figura 2 conocida como OTA linealizado con difusores simétricos. El difusor simétrico es un arreglo de transistores operando en la zona de triodo para la obtención de un comportamiento resistivo controlado, como una forma de contrarrestar los efectos no lineales del par diferencial. El difusor debe tener un tamaño de al menos siete veces el de los transistores de entrada, según recomienda Krummenacher [6], quien postuló esta variante del OTA. La transconductancia obtenida es por tanto mucho más estable en un rango más amplio de tensiones de entrada.

La Figura 1.4 muestra el OTA finalmente utilizado, el cual representó un patrón en el diseño, ya que mediante el uso de espejos de corriente es posible escalar los valores de transconductancia para obtener todos los que se especifican en el diseño de la Figura 1.2.

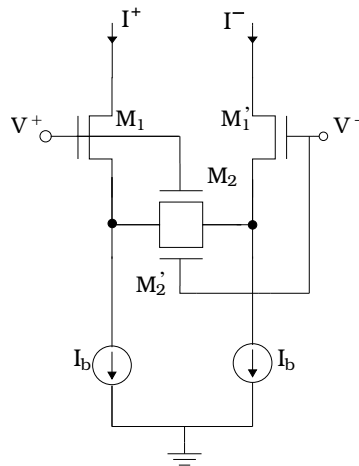


Figura 1.4: OTA linealizado con difusores simétricos.

La Figura 1.5 muestra los resultados obtenidos mediante simulaciones en Mentor Graphics para el OTA de 137nS, siendo este el OTA patrón.

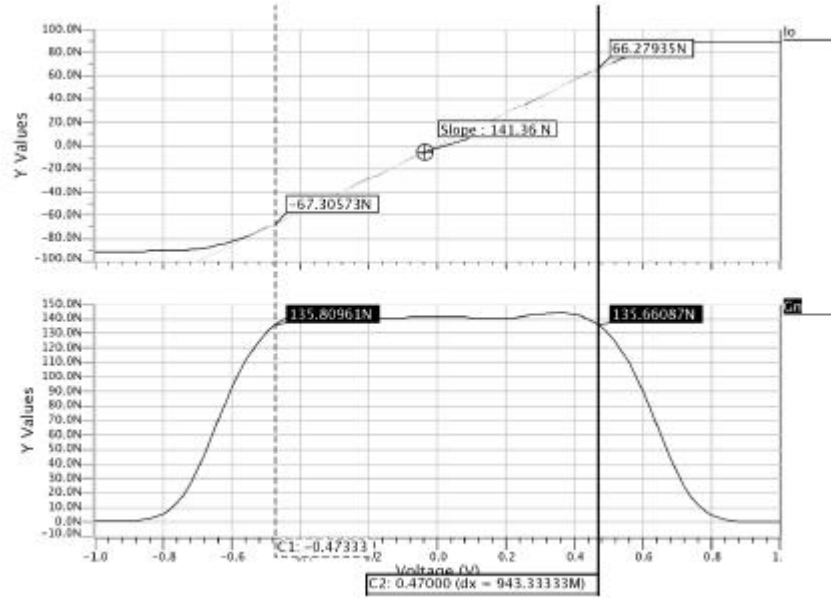


Figura 1.5: Simulaciones del OTA de 137nS. Tomado de [1].

1.2. Problema existente e importancia de su solución

Los OTAs dentro de los filtros GmC implementados (Figura 1.2) presentaron una serie de inconvenientes que desviaron su comportamiento esperado. Entre estos se encuentran una capacitancia parásita de entrada muy alta que modificaba el polo de los filtros, un slew rate insuficiente para las frecuencias de interés de cada filtro, problemas de *matching* que agravaron el corrimiento de los polos y un excesivo consumo de potencia debido a que el ajuste del circuito se realizaba por medio de la corriente de polarización.

A pesar de que el ajuste por medio de la corriente de polarización reacomodó los filtros en la posición correcta, esto disparó el consumo de potencia del dispositivo. Por las condiciones en las que trabajará el dispositivo de detección, este consumo es inaceptable y es por ello que obtener un OTA cuyas especificaciones cumplan o superen las necesidades del dispositivo con un menor consumo de energía permitiría la obtención de un dispositivo más eficiente y cuyo funcionamiento podría prolongarse por más tiempo.

Este logro además representaría un paso más cerca de llegar a obtener un nodo completo por medio de un ASIC, que como se mencionó anteriormente, es la alternativa de implementación de la red de sensores más eficiente en cuanto a consumo de potencia y cuyo legado representaría un gran avance en las intenciones de conservación y protección de los recursos naturales de Costa Rica.

1.3. Solución seleccionada y requisitos de diseño

Para la obtención del OTA definitivo (G_m), se empleó inicialmente una evaluación de algunos esquemas existentes para utilizar el más eficiente de ellos. Con ello se determinaría que el OTA con difusor simétrico continuaría siendo la mejor opción. A partir de este se diseñaría un circuito inicial que permitiría la ejecución de una herramienta de optimización de circuitos, con lo que se obtendría un conjunto de parámetros que representarían los valores más óptimos según las especificaciones solicitadas.

El OTA obtenido ha de escalarse en dos OTAs adicionales ($G_m/2$ y $G_m/4$), para su posterior obtención de filtros con diferentes frecuencias de corte (ver Figura 1.1). La Tabla 1.1 muestra las especificaciones necesarias que deben cumplir todos los OTAs obtenidos. En la medida de lo posible, siempre será deseable que el valor de la transconductancia pueda minimizarse, ya que esto implicaría una reducción del condensador requerido para formar el filtro (ver sección 2.6) y también la capacitancia parásita, ya que reduciría el corrimiento del polo del filtro. Un aumento en el rango lineal es deseable, aunque no indispensable puesto que se asegura que las señales de entrada no superarán este rango (ver [1]).

Tabla 1.1: Especificaciones mínimas para los OTAs.

Transconductancias (nS)	Rango lineal (mV)	Capacitancia parásita de entrada (pF)
134, 68.5 y 34.25 nS	± 500	0.2

Otra consideración de diseño importante es el slew rate de los OTAs obtenidos. La Tabla 1.2 muestra los requerimientos de slew rate que deben cumplirse, los cuales fueron calculados empleando la ecuación A.6 del Apéndice A para las frecuencias de corte que se indican en la Figura 1.1.

Tabla 1.2: Especificaciones de slew rate para los OTAs a diseñar.

Frecuencia de corte f_c (Hz)	OTA	Slew rate (mV/ μ s)
875	G_m	2.75
437	$G_m/2$	1.37
219	$G_m/4$	0.69

Capítulo 2

Marco teórico

2.1. Principios de funcionamiento de los transistores MOS

2.1.1. Estructura básica

La Figura 2.1 muestra la estructura básica de un transistor NMOS. Este se compone del *sustrato* (terminal B), el cual es una base de silicio sobre la cual se introducen átomos aceptores. Estos átomos, al contar con tres electrones de valencia, requieren un electrón adicional para formar los enlaces covalentes de la red que forman, el cual toman de los átomos del silicio y dejan un *hueco* (ausencia de carga negativa) en el material. Debido a la carga positiva que cada hueco representa, se dice que el conjunto de la base junto con los átomos aceptores reciben el nombre de sustrato tipo p.

Las regiones denominadas como *fuentes* (terminal S) y *drenaje* (terminal D) están compuestas por material tipo n, el cual contiene átomos donadores en lugar de aceptores a los que más bien se les asocia una carga negativa. La concentración de impurezas en estas dos regiones es mucho mayor que en el sustrato y por ello poseen una menor resistividad.

La región entre la fuente y el drenaje es llamada *canal*. Este se caracteriza por sus dimensiones (ancho (W) y largo (L)) y pueden variar desde algunos nanómetros hasta cientos de micrómetros según la tecnología que utilicen. La terminal de compuerta (G) se localiza sobre el canal, pero se encuentra aislada del mismo por una capa de material dieléctrico, lo que impide una conexión eléctrica directa entre ambas regiones y genera una alta impedancia de entrada para este tipo de transistores. Los procesos de manufactura modernos utilizan silicio policristalino (conocido también en los procesos de fabricación como *poly*) para la terminal de compuerta y dióxido de silicio para la capa aislante.

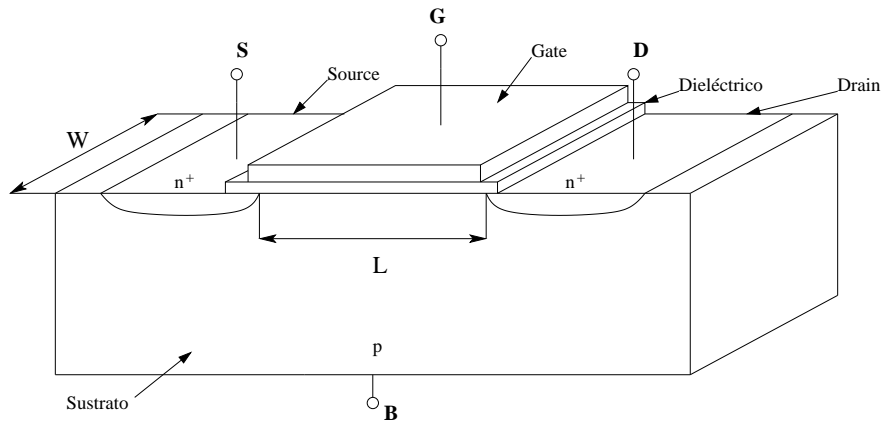


Figura 2.1: Estructura básica de un transistor NMOS

Cuando el potencial de la compuerta es mucho más positivo que las demás terminales, se genera una atracción de electrones por debajo de la superficie del material aislante. Esta acumulación de electrones viene a formar un canal entre las regiones de fuente y drenaje, permitiendo así un flujo de electrones dependiendo de la polarización que las terminales en cuestión presenten¹. La tensión aplicada en la compuerta determina la concentración de electrones en el canal y por consiguiente, cuanta corriente atraviesa el dispositivo.

Los transistores MOS (metal-óxido-semiconductor) derivan su nombre de las diversas capas que los componen. En la literatura también pueden encontrarse otros acrónimos derivados en su mayoría del idioma inglés, tales como MOST (MOS transistors), MOSFET (MOS field-effect transistor o transistores MOS de efecto de campo) e IGFET (insulated gate field-effect transistor o transistor de efecto de campo con compuerta aislada). Este último es el término más general ya que no especifica los materiales que componen cada una de las capas, aunque no es común utilizarlo. En el presente documento se utilizarán las notaciones NMOS y PMOS para indicar los transistores MOS de canal n y canal p a utilizar, respectivamente.

La Figura 2.1 muestra que las regiones de fuente y drenaje forman uniones pn con el sustrato. Si estas uniones se polarizan de manera directa, se generarían corrientes no deseadas que interferirían con la acción de transistor buscada. Por ello, debe asegurarse que tales uniones no sean polarizadas de manera directa. En un transistor NMOS este se logra al asegurar las siguientes condiciones:

$$V_{SB} \geq 0$$

$$V_{DB} \geq 0 \quad (2.1)$$

2.1.2. Niveles de inversión

Considérese el circuito de la Figura 2.2, en el cual se tiene una tensión entre drenador y fuente $V_{DS} = 0V$ y una tensión entre fuente y sustrato de magnitud positiva $V_{SB} = V_{SB1}$, la

¹La región con menor potencial es denominada fuente y la de mayor potencial drenaje, por lo que la convención pasiva de signos indicaría que el flujo positivo de corriente iría de drenaje a fuente para un potencial mayor en la terminal de drenaje.

cual garantiza una polarización inversa de ambas uniones pn . La región sombreada debajo de estas uniones se conoce como la región de agotamiento, debido a que la tensión positiva atrae los electrones de la región n hacia su superficie, mientras que el potencial negativo “atrae” los huecos hacia el terminal de sustrato, lo que genera un espacio en el cual no hay portadores libres de carga y no puede generarse una corriente eléctrica. Debido a que la concentración de impurezas es mayor en la región n que en el sustrato, la región de agotamiento es menor que la que se forma en el sustrato (la neutralidad de cargas debe conservarse).

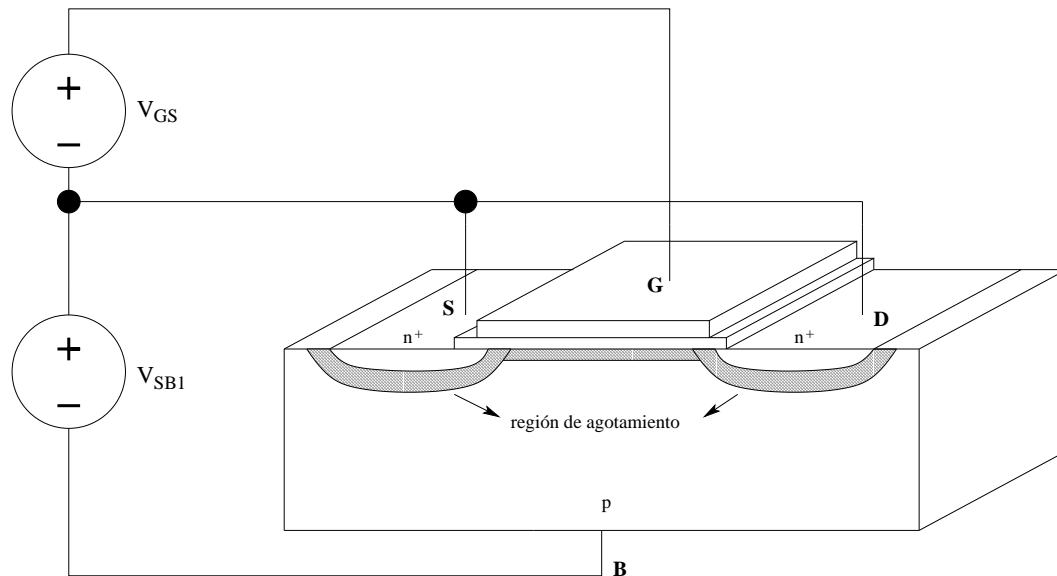


Figura 2.2: Conexión de un transistor NMOS para la evaluación de los niveles de inversión

Si la tensión entre compuerta y fuente V_{GS} es lo suficientemente negativa, esta puede atraer huecos a la superficie del canal, generando una condición conocida como acumulación. Cuando esta tensión alcanza un valor en el cual el proceso de acumulación finaliza (es decir, se atraen todas las cargas positivas disponibles) se obtiene un sustrato con una carga neutral y se dice que se alcanza la condición de *banda plana* (flat-band condition). Asimismo se habla de una neutralidad eléctrica a nivel macroscópico.

Si por el contrario se aplica una tensión V_{GS} positiva, se produce una repulsión de los huecos del sustrato, generando una zona de agotamiento en el canal, cuya profundidad depende de la intensidad de la tensión aplicada. Esta condición es conocida como agotamiento.

Cuando la tensión V_{GS} es lo suficientemente positiva como para generar una profundidad de la zona de agotamiento en el canal similar a la producida en las regiones n , se aprecia que el potencial en los terminales de drenaje y fuente y en la superficie del canal son iguales, lo que produce que los electrones de las regiones n encuentren viable desplazarse a lo largo del canal, a pesar de que el sustrato sea tipo p. Los electrones cercanos a la superficie del canal forman lo que se denomina la capa de inversión y se habla de que la superficie está invertida.

Los niveles de inversión pueden ser controlados por medio de la tensión V_{GS} y se habla de tres niveles de inversión: débil, moderada y fuerte. La concentración de electrones puede variar hasta en varios órdenes de magnitud entre una región y otra. En inversión débil y moderada, la región de agotamiento necesita extenderse ampliamente, ya que los electrones disponibles no

son suficientes para equiparar las cargas acumuladas en la terminal de compuerta por lo que es necesario obtener cargas adicionales al ensanchar la región de agotamiento. En inversión fuerte, existe mucha facilidad para el paso de los electrones entre las regiones n , ya que la abundancia de electrones es tal que tal zona de inversión puede considerarse como otra región n .

2.1.3. Efecto de cuerpo (Body effect)

La tensión V_{SB} permite definir el tamaño de la región de agotamiento que se forma alrededor de las regiones n . Si esta se hace más positiva, entonces dicha región se ensancha, por cuanto hay una mayor atracción de electrones en la región n y de huecos en el sustrato hacia los correspondientes terminales. Esto incide en que para alcanzar un nivel de inversión previo, sea necesario aumentar el valor de la tensión V_{GS} para lograrlo (la región de agotamiento debajo del canal debe aumentar también). Este fenómeno causado por V_{SB} es lo que se conoce como efecto de cuerpo y puede llegar a ser problemático si el dopado del sustrato es muy alto, ya que incrementos de V_{SB} pueden deteriorar la conducción del canal.

2.1.4. Corriente de drenaje I_D

Se asume ahora la inserción de una fuente de tensión positiva entre drenaje y fuente V_{DS} tal y como muestra la Figura 2.3. En esta ocasión, la tensión de la unión pn del lado del drenaje aumenta a $V_{DB} = V_{DS} + V_{SB1}$, lo que implica que la región de agotamiento es mayor en el drenaje. Esto genera que la condición de inversión sea más difícil de satisfacer en el lado del drenaje. En la vecindad de la región de drenaje, los electrones son atraídos fácilmente debido a que existe un potencial mayor en esta terminal. Los electrones de la fuente ingresan a la zona de agotamiento de canal, se difunden a través de este y son atraídos por el terminal de drenaje. El gradiente de concentración de electrones entre ambas regiones produce un flujo de electrones de la fuente al drenaje, causado por el fenómeno de *difusión* y se presenta cuando el nivel de inversión es débil. Es importante aclarar que bajo estas condiciones, el nivel de inversión se define de acuerdo al grado de inversión del extremo de la fuente específicamente, ya que por la diferencia de concentraciones es posible que cada extremo se encuentre en una zona de inversión diferente.

Conforme se ingresa a la inversión moderada, la capa de inversión se desarrolla y esta presenta características similares a las de los materiales resistivos. Los electrones fluyen debido a la diferencia de potencial entre ambos terminales, generando una *corriente de arrastre (drift current)*. En este modo de inversión, la corriente es producida tanto por los fenómenos de difusión y de arrastre. Finalmente, con un mayor incremento de V_{GS} se ingresa a inversión fuerte, en donde la capa de inversión se comporta prácticamente como un resistor no lineal y la corriente es predominantemente de arrastre.

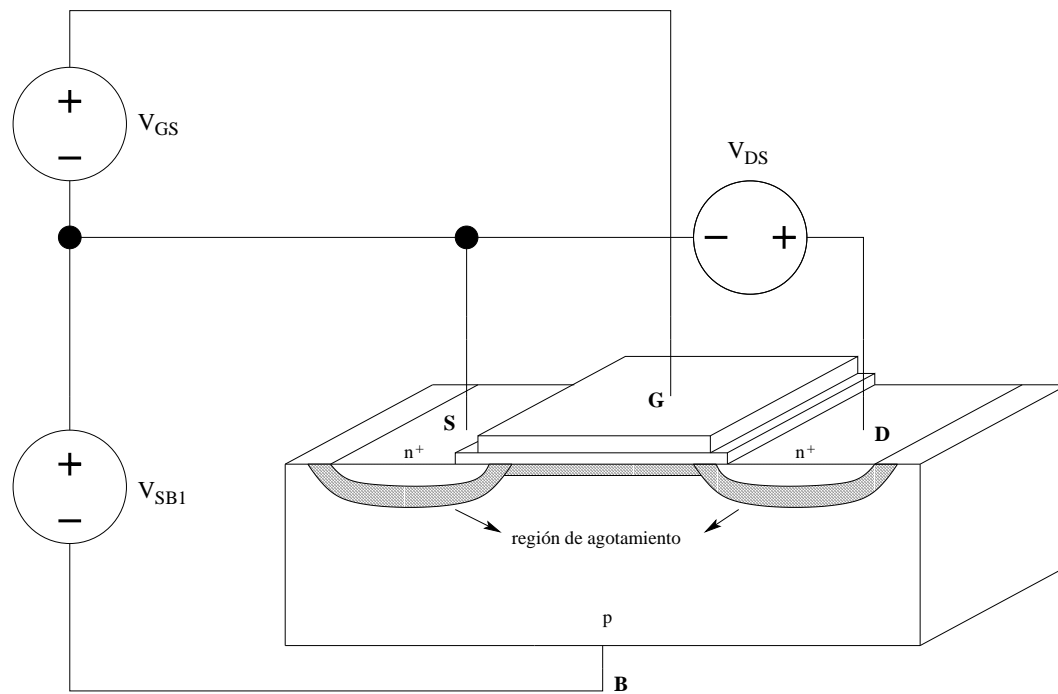


Figura 2.3: Circuito de evaluación de la corriente de drenaje

Dentro de cada nivel de inversión, es posible definir también dos regiones características dependiendo del efecto de V_{DS} en el funcionamiento del transistor. Cuando esta tensión es lo suficientemente alta, se alcanza un estado en el que los incrementos de esta no incrementan la corriente que fluye, puesto que la concentración de electrones es mínima y su velocidad de desplazamiento es alta, generando una saturación en la corriente. En esta circunstancia se dice que el extremo del drenaje se encuentra estrangulado (“pinch off”) y el transistor trabaja en modo de saturación. Asimismo, se dice que el transistor trabaja en la zona de no saturación o en la región de triodo cuando aún no alcanza la condición de estrangulamiento. El límite entre ambas regiones se definirá de manera cuantitativa en la sección 2.2.

El aumento de V_{DS} en la región de saturación causa que la capa de inversión cercana al drenador se angoste, causando el estrangulamiento mencionado y desplazando este efecto a lo largo del canal en dirección a la fuente del transistor. Sobre esta región estrecha se produce una caída de tensión que afecta la zona de agotamiento alrededor del drenador, lo que a su vez causa que la capa de inversión vista hacia la fuente modifique su tamaño para compensar la variación alrededor del drenador, ocasionando un fenómeno conocido como *modulación del ancho del canal* (ver [7] y [8]).

2.1.5. Curvas características

El comportamiento descrito en las secciones anteriores puede apreciarse en las curvas características de un transistor MOS. Estas gráficas muestran la relación entre las tensiones y corriente del dispositivo, permitiendo identificar el comportamiento del mismo bajo diversos parámetros.

Por ejemplo, la Figura 2.4² muestra una serie de gráficas de la corriente de drenador I_D en función de la tensión entre drenador y fuente V_{DS} , variando el valor de V_{GS} para así obtener diversas curvas.

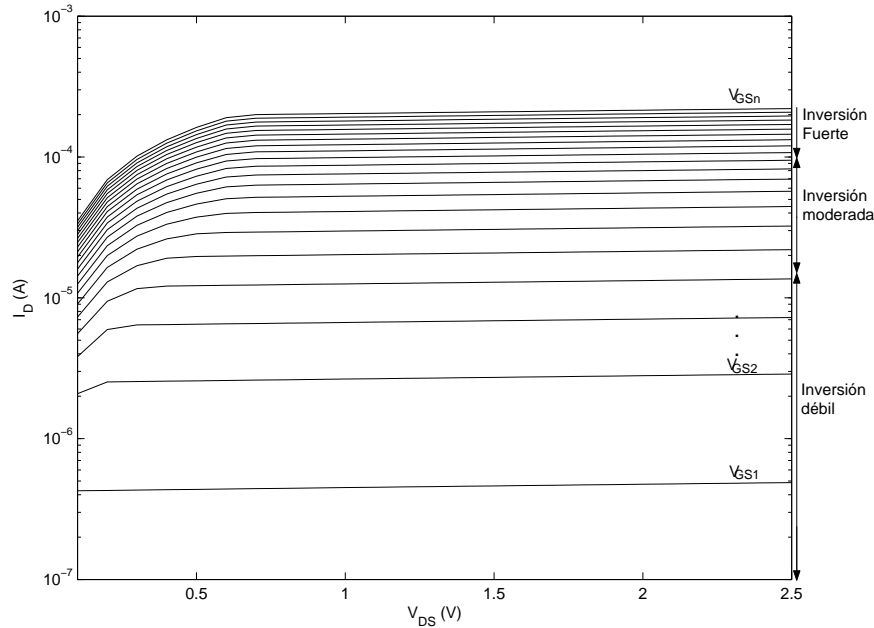


Figura 2.4: Curva característica I_D vs V_{DS} para un transistor específico, con V_{GS} como parámetro y $V_{SB} = 0$.

El eje de la corriente tiene una escala logarítmica, la cual permite apreciar la presencia de la zona de inversión débil, ya que la ubicación de estas regiones se da por medio de grandes cambios en el orden de magnitud de I_D . De hecho, la región de inversión débil se descubrió mucho después que las otras regiones, al confundirse inicialmente con corrientes de fuga dentro de los transistores, lo que llevó a los investigadores a analizar el fenómeno y así encontrar esta zona que en la actualidad permite el desarrollo de dispositivos de muy bajo consumo y operación a baja tensión ([8]).

En la Figura 2.4, las zonas de inversión se han trazado en términos de V_{GS} , lo cual ya había sido explicado previamente por su efecto en el desarrollo de la zona de agotamiento. En la sección 2.2 se mostrará cuantitativamente la separación de las regiones en términos de esta tensión.

El límite entre la saturación y la no saturación tiende a presentarse en el punto de inflexión de la curva, en donde el valor final de la corriente tiende a mantenerse constante ante variaciones de V_{DS} . También se verá que es posible establecer la relación matemática que delimita ambas regiones. En inversión débil, este límite se ubica en el mismo valor de V_{DS} sin importar la variación de V_{GS} ; no así sucede en las otras dos regiones.

La principal consecuencia del efecto de cuerpo (aumento de V_{SB}) está en la reducción de la concentración de electrones en la capa de inversión, produciendo una reducción en la corriente

²Esta curva se obtuvo con el modelo en Matlab que proporciona Berkley Wireless Research Center (BWRC) como material de apoyo de [9] para un transistor fabricado con un proceso de $0.25\mu\text{m}$.

del dispositivo si se mantienen las mismas condiciones para V_{DS} y V_{GS} . En la Figura 2.5 puede apreciarse como en a) el valor final de la curva con parámetro V_{GSn} es aproximadamente 5.9mA mientras que en b) es cercano a 2.3mA.

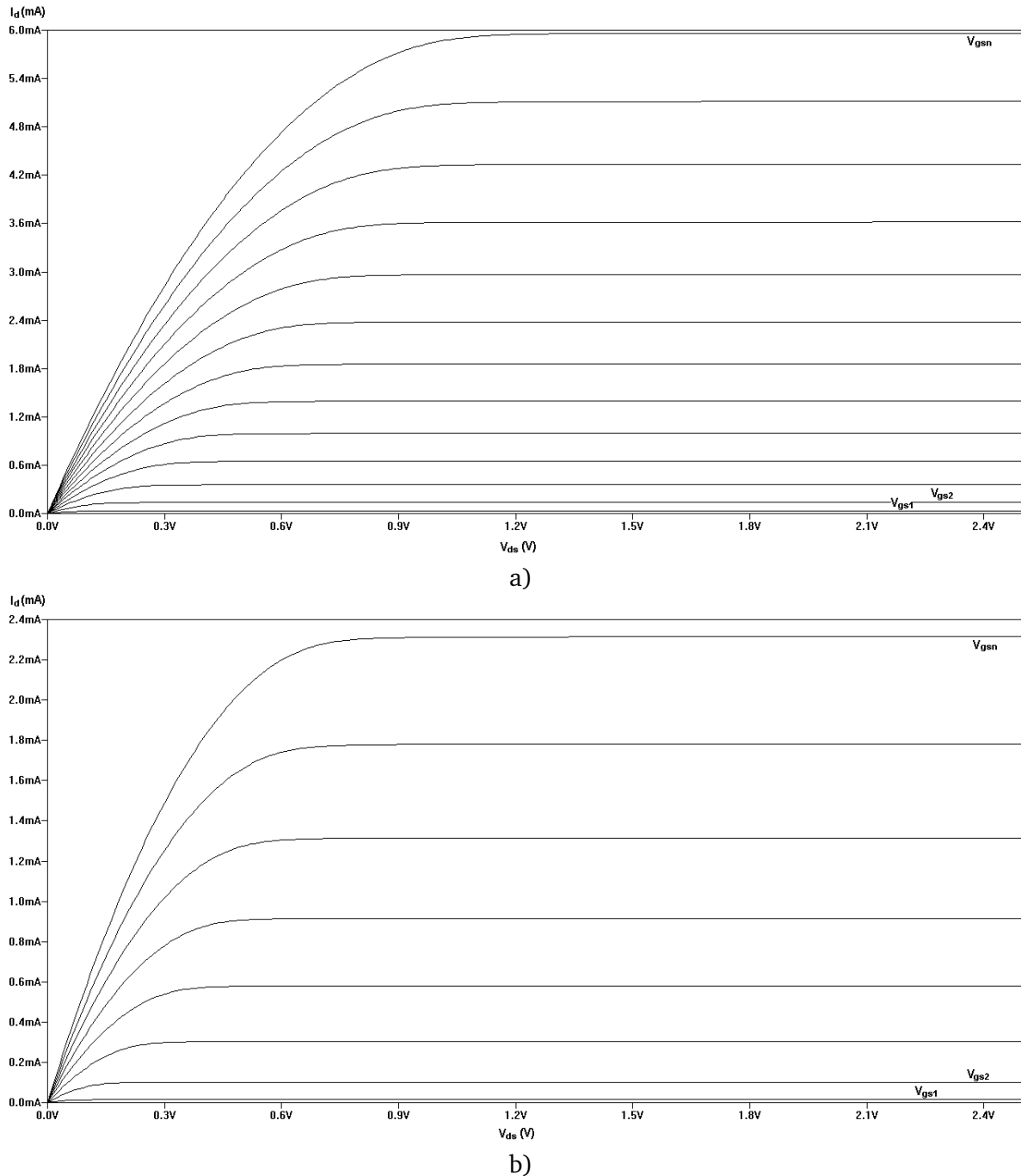


Figura 2.5: Curvas características I_D vs V_{DS} (escala lineal), con V_{GS} como parámetro y a) $V_{SB} = 0$ y b) $V_{SB} = 2,5V$. Gráficas obtenidas en LTSPICE IV para un transistor cuyo modelo corresponde a un proceso de $0,5\mu\text{m}$.

Otras curvas que permiten visualizar la localización de las zonas de inversión son la de I_D en función de la tensión V_{GS} , con un valor fijo para V_{DS} . Si se grafica la corriente en un eje logarítmico, como la gráfica superior de la Figura 2.6, se obtiene una región con un comportamiento lineal entre las tensiones V_K y V_M , siendo esta la zona de inversión débil. Al tener un comportamiento lineal en una gráfica con eje logarítmico, se aprecia que en inversión débil el comportamiento de la corriente de drenador en función de V_{GS} es exponencial, lo que se corroborará cuando las ecuaciones que definen estas curvas sean establecidas. La región por debajo de V_K carece de comportamiento exponencial debido a que las corrientes de fuga a esos niveles de tensión interfieren significativamente en el comportamiento del transistor.

La curva inferior de la Figura 2.6 utiliza $\sqrt{I_D}$ para obtener un comportamiento lineal para valores superiores a V_H , siendo este valor el límite inferior de la zona de inversión fuerte. Gracias a este comportamiento es que se dice que el transistor MOS es un dispositivo de ley cuadrática cuando opera en inversión fuerte. De esta curva se extrae otro parámetro importante para el modelado y diseño de circuitos MOS, al extrapolar la extensión lineal de dicha curva sobre el eje de V_{GS} . La intersección producida en el punto V_T se denomina tensión de umbral extrapolada³ y esta representa la facilidad de “encender” el canal, es decir, pasar de inversión débil a fuerte. Este parámetro depende de V_{SB} debido al efecto de cuerpo y de las propiedades del aislante utilizado, ya que este influye en el campo eléctrico que incide en el canal. Un valor de referencia para algunas ecuaciones se obtiene cuando $V_{SB} = 0$ y se denota como V_{T0} .

Cuando un transistor NMOS posee un valor positivo para V_{T0} , se dice que este es un dispositivo en *modo de enriquecimiento o incremental* puesto que se requiere un valor positivo de V_{GS} para alcanzar la inversión fuerte. Por el contrario, si V_{T0} es un valor negativo, el transistor se denomina en *modo de agotamiento o decremental*. La analogía opuesta aplica para transistores PMOS.

La zona de inversión moderada, ubicada dentro de los límites V_M y V_H , no presenta ninguna de las características exponenciales o cuadráticas de las regiones vecinas. Esto se debe a que los efectos de deriva y difusión previamente mencionados son igualmente importantes para el comportamiento de la corriente en esta zona, por lo que el manejo transistores dentro de esta región requiere el uso de expresiones adicionales.

³Usualmente conocida solo como tensión de umbral

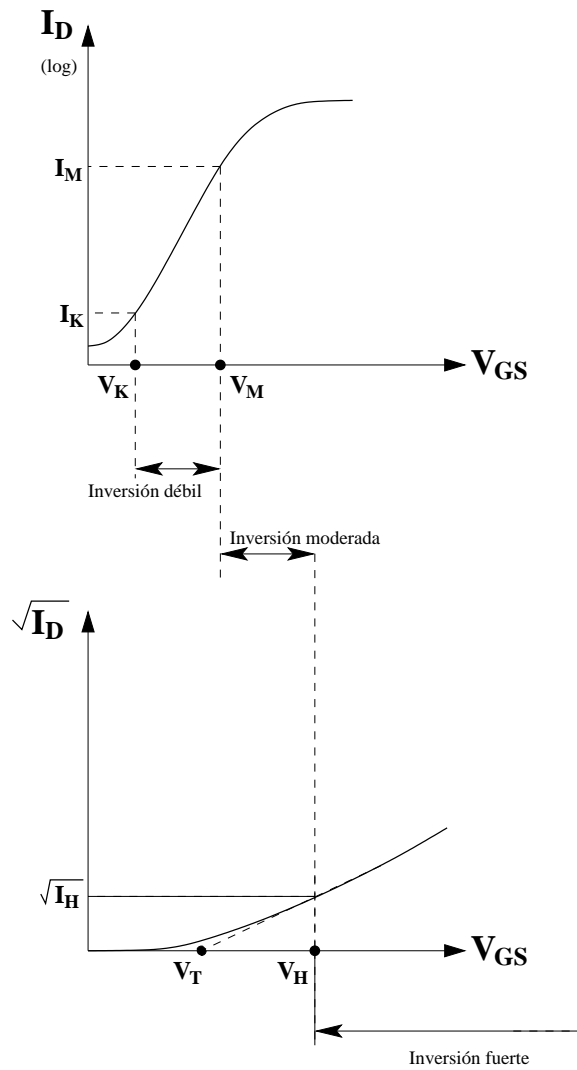


Figura 2.6: Curvas de I_D (eje logarítmico) e $\sqrt{I_D}$ vs V_{GS} , alineadas con respecto al eje V_{GS}

2.1.6. Convención de signos y símbolos para transistores NMOS y PMOS

A pesar de existir diversos símbolos para la identificación de los transistores MOS (ver [8]), en el presente documento se usaran los que se muestran en la Figura 2.7.

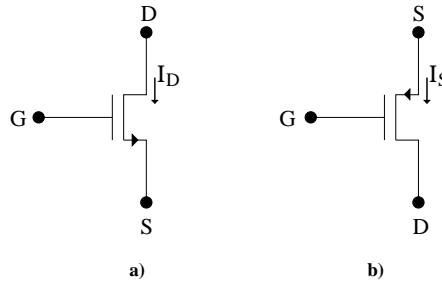


Figura 2.7: Símbolos para los transistores a utilizar: a)NMOS ; b)PMOS

La notación de corrientes (I_D para NMOS e I_S para PMOS) se hace de manera que esta sea positiva en dirección hacia la terminal de entrada respectiva. Los símbolos de la Figura 2.7 están dispuestos de tal manera que pueda apreciarse la analogía en el comportamiento de la corriente.

2.2. Ecuaciones de DC de los transistores MOS

La presente sección busca enmarcar las características descritas anteriormente dentro de ecuaciones que relacionen los diferentes parámetros mencionados y analizados, así como generar un punto de referencia para el posterior diseño de circuitos. Estas ecuaciones se incluyen en diversas secciones dependiendo de la naturaleza de las mismas y están escritas para transistores NMOS. Cabe recalcar que para los identificadores de corriente y tensión en DC se utilizan subíndices con letras mayúsculas, mientras que magnitudes en AC utilizan subíndices con letras minúsculas para su debida diferenciación. Las ecuaciones de DC presentadas en esta sección son las que desarrolla Tsividis [8].

2.2.1. Parámetros dependientes de las características físicas del proceso

Las siguientes expresiones permiten el cálculo de algunos parámetros necesarios en ecuaciones posteriores. Estos parámetros destacan por depender de las características del proceso y la tecnología de fabricación:

$$\phi_t = 25,9mV \times \frac{T}{300K} \quad (2.2)$$

$$K' = \mu C_{ox} \quad (2.3)$$

$$C_{ox} = \frac{\epsilon_{ox}}{d_{ox}} \quad (2.4)$$

$$\gamma = \frac{\sqrt{2q\epsilon_r N_B}}{C_{ox}} \quad (2.5)$$

$$V_A = \phi_A \times \frac{L}{1\mu m} \times \sqrt{\frac{N_C}{10^{16}cm^{-3}}} \quad (2.6)$$

En las ecuaciones anteriores, ϕ_t representa el voltaje térmico, T la temperatura de operación del transistor, μ la movilidad de los portadores en el canal, ϵ_{ox} es la permitividad de la capa aislante, d_{ox} el espesor de la capa aislante, N_B es la concentración de dopado en el sustrato, N_C es la concentración de dopado en el canal, ϵ_r representa la permitividad del conductor, L indica el largo del canal y ϕ_A es un valor de compromiso que depende del proceso de fabricación (para cálculos a mano se utiliza un valor típico de 16V).

2.2.2. Ecuaciones de aproximación de los límites entre zonas de inversión

Las ecuaciones mostradas a continuación permiten establecer los límites entre las zonas de inversión (V_M , V_H y sus respectivas corrientes, I_M e I_H , según [8]) mostrados en la Figura 2.5, así como el cálculo de la tensión de umbral V_T .

$$n = 1 + \frac{\gamma}{2\sqrt{V_{SB} + \phi_0}} \quad (2.7)$$

$$\alpha = 1 + \frac{\gamma}{2\sqrt{V_{SB} + \phi_0 + \phi_a}} \quad (2.8)$$

$$V_T = V_{T0} + \gamma \left(\sqrt{V_{SB} + \phi_0} - \sqrt{\phi_0} \right) \quad (2.9)$$

$$V_M = V_T - c \cdot n \cdot \phi_t \quad (2.10)$$

$$I'_M = K' (n - 1) \phi_t^2 \quad (2.11)$$

$$V_H \approx V_T + 6n\phi_t \quad (2.12)$$

$$I'_H = 20K'n\phi_t^2 \quad (2.13)$$

Las ecuaciones 2.7 y 2.8 representan dos cantidades dependientes principalmente de V_{SB} y que representan parámetros dentro de otras ecuaciones. También involucran un nuevo parámetro, ϕ_0 , que usualmente tiene un valor de 0.7V.

V_H suele ser un parámetro más flexible para el proceso de diseño, ya que se define en una proporción por encima de V_T . Es por ello que la ecuación 2.12 muestra una aproximación típica para V_H . No obstante, V_M sí suele establecerse con mayor precisión mediante la ecuación 2.10, aunque depende de c , el cual es un parámetro que no se conoce con precisión, por lo que Tividis [8] recomienda que el proceso de diseño no dependa de la exactitud de este parámetro.

I'_M e I'_H identifican la corriente de drenador que se presenta en los límites establecidos por V_M y V_H , respectivamente. La utilidad de estas ecuaciones (2.11 y 2.13) está en la verificación de que la corriente obtenida mediante las ecuaciones de la sección siguiente se encuentre en la región adecuada.

Los fabricantes de circuitos integrados proporcionan a sus clientes los datos de K' , γ , V_{T0} , n y α directamente para complementar el diseño de circuitos a mano. No obstante, de acuerdo con las políticas del proceso y derechos de propiedad sobre los mismos, esta información puede ser de carácter confidencial o no.

2.2.3. Ecuaciones de corriente para diversas regiones de operación

Para que un transistor opere en la región de inversión débil, debe cumplir las dos siguientes condiciones:

$$V_{GS} \leq V_M$$

$$I_{D,SAT} \leq I_M$$

en donde $I_{D,SAT}$ representa la corriente de polarización obtenida para la región de inversión débil (ecuación 2.14):

$$I_D = I_M \exp\left(\frac{V_{GS} - V_M}{n\phi_t}\right) \left[1 - \exp\left(-\frac{V_{DS}}{\phi_t}\right)\right] \quad (2.14)$$

$$I_M = \frac{W}{L} I'_M \quad (2.15)$$

En la ecuación 2.14, se aprecia el comportamiento exponencial descrito anteriormente para la corriente dentro de esta región. En la ecuación anterior, si $V_{DS} \gg 5\phi_t$, se considera que el transistor opera en saturación y todo el segundo paréntesis puede aproximarse a 1. De lo contrario, opera en la región de triodo y la ecuación completa debe considerarse en el cálculo de I_D .

Un transistor que opera en la región de inversión fuerte debe cumplir las siguientes condiciones:

$$V_{GS} \geq V_H$$

$$I_{D,SAT} \geq I_H$$

Con I_H definido como:

$$I_H = \frac{W}{L} I'_H \quad (2.16)$$

En este caso, $I_{D,SAT}$ representa la corriente de drenador calculada con ecuaciones para la región de inversión fuerte. La ecuación de corriente debe elegirse de acuerdo a si se trabaja en saturación o no y el límite entre ambos casos está dado por V'_{DS} y se define como:

$$V'_{DS} = \frac{V_{GS} - V_T}{\alpha} \quad (2.17)$$

Para que el transistor opere en la zona de no saturación debe cumplirse que $V_{DS} \leq V'_{DS}$ y la ecuación en este caso sería:

$$I_D = K \left[(V_{GS} - V_T) V_{DS} - \frac{1}{2} \alpha V_{DS}^2 \right] \quad (2.18)$$

$$K = \frac{W}{L} K' \quad (2.19)$$

Si opera en la zona de saturación ($V_{DS} > V'_{DS}$), la corriente se define entonces como:

$$I_D = \frac{k}{2} (V_{GS} - V_T)^2 \quad (2.20)$$

$$k = \frac{W}{L} k' \quad (2.21)$$

$$k' = \frac{K'}{\alpha} \quad (2.22)$$

en donde la ecuación 2.20 muestra el comportamiento cuadrático de la corriente en función de la tensión V_{GS} descrito anteriormente. Si se desea modelar el efecto de la tensión V_{DS} en el cálculo de I_D en la región de inversión fuerte con saturación, se emplea la siguiente ecuación:

$$I_D \approx I'_D \left(1 + \frac{V_{DS}}{V_A} \right) \quad (2.23)$$

El parámetro V_A (presentado en la ecuación 2.6 de la sección 2.2.1) es el análogo de la tensión de Early⁴ para transistores MOS. No obstante la aproximación de la ecuación 2.23 es un estimado de baja precisión puesto que el fenómeno de modulación de ancho del canal debido a V_{DS} en la región de saturación es complejo de modelar (ver [8]), por la cantidad y tipo de efectos que se producen en esas condiciones.

Tsividis [8] menciona que para el caso de inversión moderada, debe aplicarse la ecuación general para todas las regiones (débil, moderada y fuerte, saturación y no saturación), de la cual se derivan las ecuaciones anteriores mediante la aproximación de algunos de sus términos. La ecuación general es:

$$I_D = I_Z \left\{ \ln^2 \left[1 + \exp \left(\frac{V_{GS} - V_T}{2n\phi_t} \right) \right] - \ln^2 \left[1 + \exp \left(\frac{V_{GS} - V_T - nV_{DS}}{2n\phi_t} \right) \right] \right\} \quad (2.24)$$

$$I_Z = 2Kn\phi_t^2$$

Si un transistor trabaja en la zona de saturación, el segundo logaritmo de la ecuación 2.24 es despreciable y la expresión puede reducirse a:

$$I_D = I_Z \ln^2 \left[1 + \exp \left(\frac{V_{GS} - V_T}{2n\phi_t} \right) \right] \quad (2.25)$$

2.3. Características de pequeña señal de los transistores MOS

2.3.1. Parámetros de transconductancia

Los modelos de pequeña señal para los transistores MOS permiten el análisis del comportamiento del dispositivo ante variaciones en las tensiones del circuito. Si estas variaciones son

⁴El efecto de Early en transistores BJT (ver [10]) se debe a las variaciones del ancho de la base causadas por la tensión base-colector aplicada al transistor, afectando el cálculo de la corriente de emisor por un factor $(1 + \frac{V_{CE}}{V_A})$, siendo V_{CE} la tensión entre colector y emisor y V_A el voltaje de Early, en honor a su descubridor, James M. Early.

pequeñas, producen una variación en la corriente en una cantidad proporcional. Primeramente debe asumirse que el transistor ha sido polarizado en un punto de operación Q , con valores fijos para V_{GSQ} , V_{DSQ} y V_{SBQ} , respectivamente.

Para cada una de las cantidades anteriores, puede definirse la variación de la corriente I_D ante variaciones de cada una de ellas. Esto permite la obtención de tres parámetros de transconductancia:

1. Transconductancia de compuerta g_m :

$$g_m = \left. \frac{\partial I_D}{\partial V_{GS}} \right|_Q \quad (2.26)$$

2. Transconductancia de sustrato g_{mb} :

$$g_{mb} = \left. \frac{\partial I_D}{\partial V_{BS}} \right|_Q \quad (2.27)$$

3. Conductancia entre fuente y drenador g_{ds} :

$$g_{ds} = \left. \frac{\partial I_D}{\partial V_{DS}} \right|_Q \quad (2.28)$$

en donde la transconductancia respectiva debe evaluarse en el punto de operación Q .

La variación de corriente ΔI_D producida por la variación de las tensiones es una combinación de las mismas, tal y como lo muestra la ecuación 2.29:

$$\Delta I_D = g_m \Delta V_{GS} + g_{mb} \Delta V_{BS} + g_{ds} \Delta V_{DS} \quad (2.29)$$

El circuito equivalente obtenido para el transistor MOS y que implementa la ecuación 2.29 se muestra en la Figura 2.8.

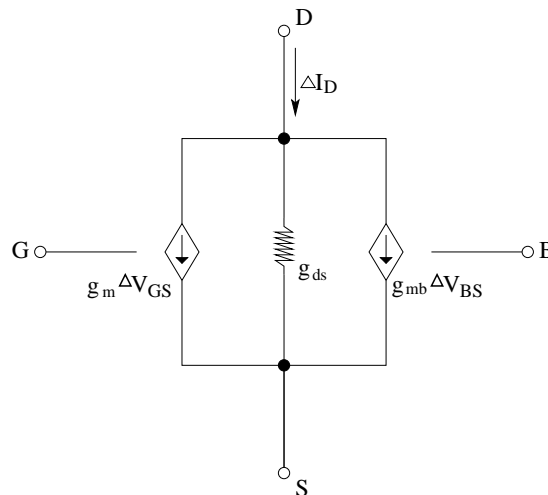


Figura 2.8: Circuito equivalente de pequeña señal para el transistor NMOS

La ecuación 2.29 puede reescribirse en términos de variables en AC como:

$$i_d = g_m \cdot v_{gs} + g_{mb} \cdot v_{bs} + g_{ds} \cdot v_{ds} \quad (2.30)$$

Utilizando las definiciones de I_D presentadas en la sección 2.2 y las ecuaciones de transconductancia (ecuaciones 2.26, 2.27 y 2.28) es posible obtener expresiones en función de los valores de polarización y características físicas del transistor. La Tabla 2.1 muestra las ecuaciones derivadas para los casos de inversión fuerte y débil.

Tabla 2.1: Aproximaciones para la transconductancia en inversión fuerte y débil

	Inversión débil	
	No saturación	Saturación
g_m		$\frac{I_D}{n\phi_t}$
g_{mb}		$(n-1)g_m$
g_{ds}	$\frac{I_D}{\phi_t} \exp\left(-\frac{V_{DS}}{\phi_t}\right)$	$\frac{I_D}{V_A}$
	Inversión fuerte	
	No saturación	Saturación
g_m	$\frac{KV_{DS}}{\gamma}$	$\frac{\sqrt{2kI_D}}{\gamma}$
g_{mb}	$\frac{\gamma}{2\sqrt{V_{SB}+\phi_0+0,4V_{DS}}}g_m$	$\frac{\gamma}{2\sqrt{V_{SB}+\phi_0+0,4V'_{DS}}}g_m$
g_{ds}	$K(V_{GS} - V_T - \alpha V_{DS})$	$\frac{I_D}{V_A}$

2.3.1.1. Modelo general y el cálculo de transconductancias

Las ecuaciones para las transconductancias, válidas para todas las regiones de inversión en saturación, se derivan del modelo general de la ecuación 2.25 y son especialmente útiles para trabajar en inversión moderada, por cuanto no se dispone de expresiones más específicas para esta región. Las ecuaciones para las transconductancias mostradas en el circuito equivalente de la Figura 2.8 son:

$$g_m = \frac{I_D}{n\phi_t} \frac{1}{f(x)} \quad (2.31)$$

$$g_{mb} = (n-1)g_m \quad (2.32)$$

$$g_{ds} = \frac{I_D}{V_A} \quad (2.33)$$

$$f(x) = \sqrt{1 + 0,5\sqrt{x} + x} \quad (2.34)$$

$$x = \frac{I_D}{I_Z} \quad (2.35)$$

$$I_Z = 2\frac{W}{L}K'n\phi_t^2 \quad (2.36)$$

en donde el parámetro x de la función $f(x)$ es un indicativo de la región de inversión en la cual opera el transistor⁵ y depende esencialmente de la corriente de polarización y de las dimensiones del elemento, tal y como lo muestra la ecuación 2.34 y sus ecuaciones auxiliares (2.35 y 2.36).

2.3.2. Parámetros de capacitancia

Las variaciones de tensión aplicadas causan variaciones similares en la concentración de cargas q en el transistor. Si las variaciones de tensión aplicadas son lo suficientemente rápidas, la relación dq/dt puede llegar a ser significativa, ocasionando una corriente capacitiva que se añade a las corrientes de transconductancia presentadas anteriormente. Estas corrientes pueden afectar el desempeño de circuitos en los cuales el funcionamiento en frecuencia es crucial, ya que la capacitancia equivalente para estas corrientes altera las funciones de transferencia de los sistemas en donde se emplean estos transistores.

El circuito equivalente de pequeña señal completo se muestra en la Figura 2.9, el cual incluye las diversas capacitancias que se forman entre las diferentes capas que componen el transistor.

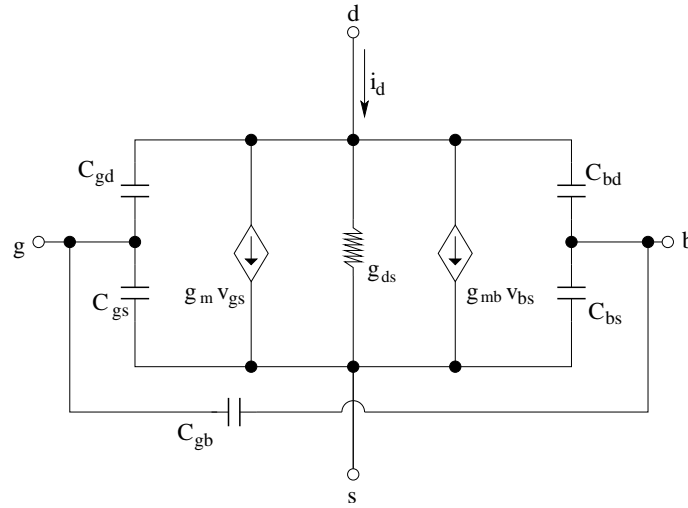


Figura 2.9: Circuito equivalente de pequeña señal para un transistor NMOS, con capacitancias incluidas

Para el modelo general en saturación se tiene el siguiente conjunto de ecuaciones para el cálculo de las transconductancias mostradas:

$$C_{gs} = WLC_{ox} \left[\frac{3}{2} + \frac{f(x)}{x} \right]^{-1} \quad (2.37)$$

$$C_{bs} = WLC_{ox} (n-1) \left[2 + \frac{f(x)}{x} \right]^{-1} \quad (2.38)$$

$$C_{gb} = WLC_{ox} \frac{n-1}{n} \left[1 - \left(\frac{3}{2} + \frac{f(x)}{x} \right)^{-1} \right] \quad (2.39)$$

⁵ x es análoga a la corriente de inversión i_f mostrada en [1]. Una discusión asociada a esta corriente de inversión se muestra al final de la sección 5.2.2 como parte de la aproximación del rango lineal.

$$C_{gd} = C_{bd} = 0 \quad (2.40)$$

2.4. Principios básicos de los OTA

Los OTA son circuitos que producen una corriente de salida en función a una tensión de entrada diferencial. Su comportamiento ideal se caracteriza por una impedancia de entrada infinita y una impedancia de salida también infinita, tal y como se muestra en la Figura 2.10. La función de transferencia ideal de este circuito vendría a ser la transconductancia equivalente, denotada por G_m . Como se mencionó anteriormente, la alta impedancia de entrada de los transistores MOS es la que propiedad que convierte a este tipo de transistores como los predilectos a utilizar en el diseño de OTAs, así como el bajo consumo de energía en los circuitos construidos con estos dispositivos y la posibilidad de obtener transconductancias en el orden de nanoSiemens incluso picoSiemens.

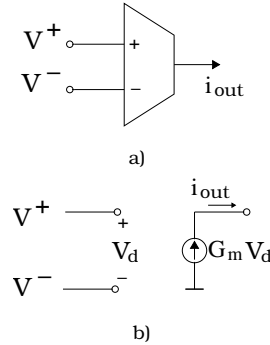


Figura 2.10: a) Representación del OTA ideal. b) Circuito equivalente de pequeña señal.

Las ecuaciones ideales que describen a un OTA son (ver [11]):

$$i^+ = i^- = 0 \quad (2.41)$$

$$i_{out} = G_m V_d \quad (2.42)$$

La definición formal para la transconductancia del OTA se basa en el conocimiento del comportamiento de la corriente de salida en función de la tensión diferencial de entrada, para lo cual se aplica una derivada parcial según se plantea en [12] y se muestra en la ecuación 2.43:

$$G_{m(V_d)} = \frac{\partial i_{out}}{\partial V_d} \quad (2.43)$$

La estructura principal de un OTA se compone esencialmente de dos etapas. La primera de ellas es un amplificador de entrada diferencial, el cual produce variaciones de corriente como respuesta a cada una de las entradas de tensión (V^+ y V^-). La segunda etapa se compone usualmente de espejos de corriente que combinan estas fluctuaciones para formar una corriente de salida y al mismo tiempo suprimir la corriente de polarización de DC. En la Figura 1.3 pueden

apreciarse ambas etapas, en donde los transistores inferiores representan el par diferencial de entrada y los superiores el espejo de corriente.

La respuesta de transconductancia del par diferencial no es una función lineal de la tensión de entrada diferencial. Esto se debe a que las ecuaciones de la sección 2.2 muestran que la región de operación define el tipo de respuesta de la corriente del transistor. Dado que en inversión débil se presenta una respuesta de tipo exponencial y en inversión fuerte aparecen expresiones cuadráticas, el diseño de OTAs se ha abocado a la utilización de la zona de inversión moderada, ya que esta presenta un comportamiento mucho más lineal que en las zonas mencionadas y presenta un mejor compromiso entre los parámetros de diseño (consumo, linealidad, distorsión, etc) de acuerdo con [5].

Para lograr una respuesta mucho más lineal de la transconductancia de entrada, se pueden aplicar algunas mejoras al par diferencial. La sección 2.5 tratará algunos de los esquemas propuestos en [6], [8] y [13] y vistos en el material de Dualibe [2].

2.5. Esquemas básicos para el diseño de OTAs

Las configuraciones aquí presentadas para la estabilización del valor de la transconductancia corresponden a la etapa de entrada del OTA. La obtención de la transconductancia en los casos siguientes se da para cuando este opera en inversión fuerte, ya que el análisis matemático es más sencillo en esta región, aunque un análisis exhaustivo se presentará posteriormente para el caso implementado.

2.5.1. OTA diferencial básico

El caso general, visto en la Figura 2.11, contiene básicamente dos componentes: los transistores NMOS de entrada y la fuente de corriente de polarización.

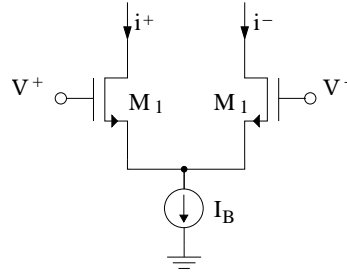


Figura 2.11: OTA diferencial básico.

Las ecuaciones que describen el comportamiento de este circuito se establecen a continuación:

$$i_{out} = i^+ - i^- = \sqrt{\mu C_{ox} (W/L)_{M1} I_B} V_d \sqrt{1 - \frac{\mu C_{ox} (W/L)_{M1} V_d^2}{4 I_B}}$$

$$G_m \approx \sqrt{\mu C_{ox} (W/L)_{M1} I_B} - \frac{3}{8} \sqrt{\frac{(\mu C_{ox} (W/L)_{M1})}{4 I_B}} V_d^2 \quad (2.44)$$

Como se aprecia en la ecuación 2.44, el valor final de la transconductancia del circuito va a ser modificado por un factor cuadrático de la tensión de entrada diferencial $V_d = V^+ - V^-$. Este factor indica que el rango lineal del circuito va a ser muy limitado (unos pocos mV). La mayor ventaja que presenta este circuito es que emplea una cantidad mínima de componentes. En el diseño de configuraciones del par diferencial, los transistores de entrada son simétricos, es decir, poseen las mismas dimensiones.

2.5.2. Principio de degeneración de la fuente

Tanto este esquema como los posteriores se basan en la colocación de un elemento con comportamiento resistivo entre los terminales de source de ambos transistores del par diferencial. Esto permite lograr un mejor balance de las corrientes del circuito, ya que la corriente que atraviesa este componente viene dada directamente por las variaciones de la tensión diferencial (ver [2]). El circuito que utiliza este principio se muestra en la Figura 2.12.

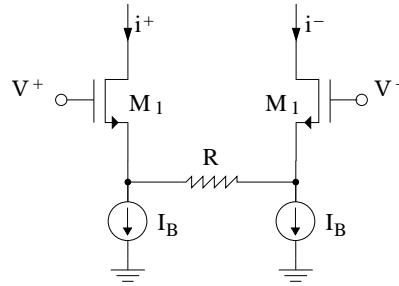


Figura 2.12: OTA linealizado por medio de un elemento resistivo.

Las ecuaciones que describen el comportamiento de este circuito se describen a continuación:

$$g_m = \sqrt{2\mu C_{ox} (W/L)_{M1} I_B}$$

$$n = 1 + g_m \cdot R$$

$$i_{out} = \left(\frac{\sqrt{2\mu C_{ox} (W/L)_{M1} I_B}}{n} \right) V_d \sqrt{1 - \left(\frac{V_d}{2n(V_{GS} - V_T)} \right)^2}$$

$$G_m \approx \frac{1}{R}; \text{ si } V_d \ll 2n(V_{GS} - V_T) \quad (2.45)$$

La transconductancia equivalente del circuito puede aproximarse al valor inverso del elemento resistivo para valores de tensión de entrada inferiores a la condición mostrada en la ecuación 2.45. Esto limita el rango lineal del amplificador, aunque va a presentar una transconductancia estable dentro del mismo, a diferencia del OTA básico. Sin embargo, la obtención de un elemento resistivo dentro de circuitos integrados es una labor poco práctica ya que estos requieren de mucho espacio.

2.5.3. Transistor con polarización en modo común

Para simular el elemento resistivo del principio anterior, se puede emplear un transistor operando en la zona de triodo (no saturación). Por medio de la tensión de gate es posible regular la transconductancia entre los terminales de drain y source que este transistor representa. La Figura 2.13 muestra el circuito correspondiente a este esquema.

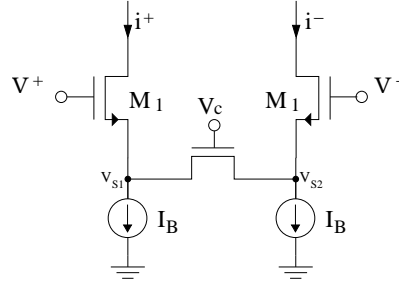


Figura 2.13: OTA con transistor con polarización en modo común

Las ecuaciones que describen el comportamiento de este circuito se muestran a continuación:

$$V_{GS3} = V_C - V_{S3} = V_C - V_{S1}$$

$$V_{DS3} = V_{S2} - V_{S1}$$

$$g_{ds} = \mu C_{ox} (W/L)_{M3} (V_{GS3} - V_T - V_{DS3})$$

$$G_m \approx g_{ds} \approx \mu C_{ox} (W/L)_{M3} (V_C - V_{S1} - V_T - V_{DS3}) \quad (2.46)$$

En este caso, las tensiones de entrada son las que ocasionan los fenómenos de no linealidad para el valor resultante de la transconductancia. En régimen de pequeña señal, las variaciones de la tensión de entrada afectan proporcionalmente las tensiones de los nodos V_{S1} y V_{S2} , modificando así el valor de V_{DS3} del transistor M_3 . Estas mismas variaciones se reflejan dentro de la ecuación 2.46 directamente en el valor de V_{S1} . Esto hace que el circuito sea muy sensible a entradas de modo común.

Para contrarrestar parcialmente estos efectos, se utiliza una tensión V_C mucho más grande que las demás tensiones dentro de la ecuación 2.46, para que de esta manera las variaciones que se presenten no sean tan comparables al valor de V_C y afecten en lo más mínimo la transconductancia resultante.

2.5.4. OTA con difusor simétrico

El esquema propuesto por Krummenacher [6] emplea dos transistores en paralelo, cuya tensión V_{gs} varía de manera proporcional a la misma tensión de los transistores del par de entrada diferencial. Con esta configuración, el difusor simétrico (nombre que reciben los dos transistores en paralelo) busca estabilizar el valor de la transconductancia a partir de las variaciones de la tensión de entrada. El circuito de la Figura 2.14 muestra el circuito para el modelo descrito.

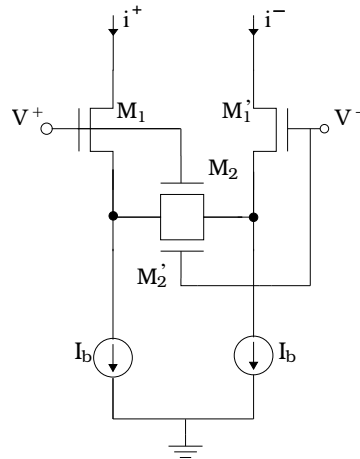


Figura 2.14: OTA con difusor simétrico

Las ecuaciones que describen el comportamiento del circuito presentado son:

$$k_i = \frac{\mu C_{ox} (W/L)_{M_i}}{2}$$

$$r_{ds_i} = \frac{1}{2k_i (V_{gs_i} - V_T)}$$

$$r_{s_i} = r_{ds_i}$$

$$G_m = \frac{1}{r_{s1} + r_{s1'} + r_{ds2} \parallel r_{ds2'}} = \frac{k_1 \sqrt{I_B}}{\left(1 + \frac{k_1}{4k_2}\right) \sqrt{k_1}} \quad (2.47)$$

La ecuación anterior muestra que el valor de transconductancia depende esencialmente de las dimensiones de los transistores y la corriente de polarización. No obstante, un OTA real sí va a presentar un comportamiento variable en función de la tensión de entrada, debido a los efectos no lineales de los transistores junto con el hecho de que las ecuaciones presentadas en esta sección aplican para transistores en inversión fuerte. Sin embargo, esta configuración presenta uno de los rangos lineales más amplios. La Figura 2.15 muestra una comparativa de tres de las estructuras presentadas anteriormente. Tanto en el esquema básico (a) como en el del transistor en modo común (b) se aprecia un comportamiento cuadrático de la transconductancia en función de V_d . Para el caso del difusor simétrico (c), se obtiene una curva con un rango lineal limitado, pero que supera al que puede obtenerse con las configuraciones anteriores [2].

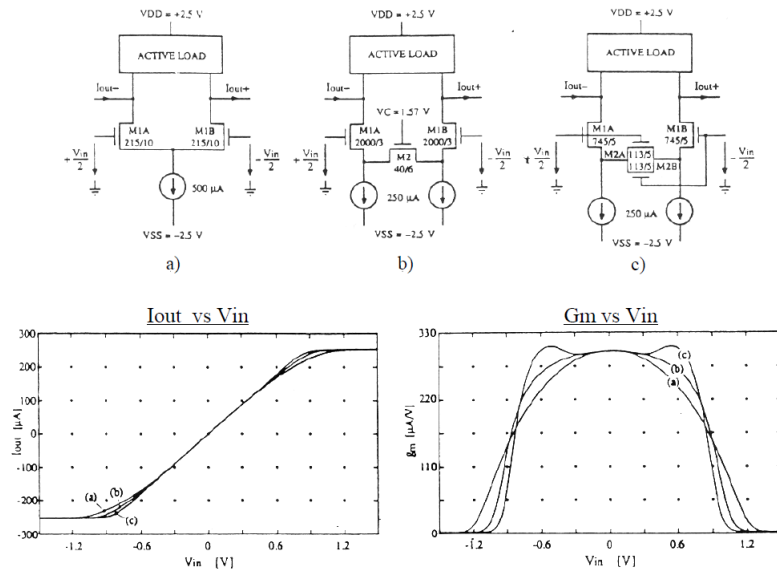


Figura 2.15: Gráficas de I_{out} vs $V_{in}(V_d)$ y G_m vs V_{in} para las configuraciones: a) par diferencial básico. b) transistor en modo común. c) OTA con difusor simétrico. Tomado de [2].

2.5.5. Par diferencial con fuentes de tensión flotante

La estructura propuesta por Nedungadi [13] se basa en el principio de la suma constante de tensiones source y gate, el cual se ilustra en la Figura 2.16. Este principio establece que si la suma de las tensiones v_{gs} de los transistores del par diferencial es constante, entonces la diferencia de corrientes $i^+ - i^-$ es linealmente proporcional a $v_{gs1} - v_{gs2}$, según puede apreciarse en la ecuación 2.48.

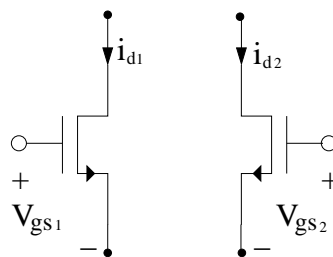


Figura 2.16: Principio de suma constante de tensiones entre source y gate

Matemáticamente, este principio se describe de la siguiente forma:

$$i_{d1} = k_1 (v_{gs1} - V_T)^2$$

$$i_{d2} = k_1 (v_{gs2} - V_T)^2$$

$$i_{out} = i_{d1} - i_{d2} = k_1 (v_{gs1} - V_T)^2 - k_2 (v_{gs2} - V_T)^2$$

$$i_{out} = k_1 (v_{gs1} + v_{gs2} - 2V_T) (v_{gs1} - v_{gs2}) \quad (2.48)$$

La suma de las tensiones v_{gs1} y v_{gs2} pueden mantenerse constante mediante el uso de fuentes de tensión flotante, las cuales se colocan de manera complementaria (el terminal positivo se conecta a la compuerta y el negativo al terminal de source pero del transistor opuesto), como lo indica la Figura 2.17.

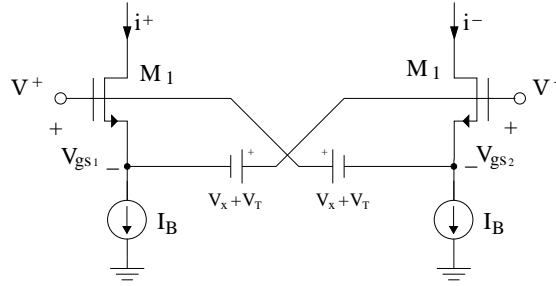


Figura 2.17: Par diferencial con fuentes de tensión flotante

Al aplicar la ley de tensiones de Kirchoff (LKT) en ambas mallas se obtiene:

$$V^+ - v_{gs1} + V_x + V_T = V^- \quad (2.49)$$

$$V^- - v_{gs2} + V_x + V_T = V^+ \quad (2.50)$$

Al efectuar la suma de las ecuaciones 2.49 y 2.50 se obtiene que:

$$v_{gs1} + v_{gs2} = 2(V_x + V_T) \quad (2.51)$$

La expresión anterior muestra que la suma de las tensiones gate-source de los transistores es constante y permite aplicar el principio recién mostrado. Por otra parte, al efectuar la resta de las ecuaciones 2.49 y 2.50 se demuestra que:

$$v_{gs1} - v_{gs2} = 2(V^+ - V^-) \quad (2.52)$$

Al sustituir las ecuaciones 2.51 y 2.52 en la ecuación 2.48, se obtiene una expresión para la corriente de la configuración anterior y de ella puede derivarse la transconductancia del circuito:

$$i_{out} = kV_x (V^+ - V^-) = kV_x V_d$$

$$G_m = 4kV_x \quad (2.53)$$

La transconductancia obtenida es prácticamente lineal ya que no se utilizan aproximaciones ni suposiciones en la derivación de la misma, como si sucedía para los casos anteriores. No

obstante, la obtención de estas fuentes flotantes no es una tarea fácil y mucho menos lo es la precisión con la que se debe generar la tensión $V_x + V_T$ para lograr la simplificación mostrada.

La propuesta de Nedungadi para obtener estas fuentes flotantes se basa en el uso de transistores que se conectan al potencial de alimentación y a las terminales de gate y source de los transistores del par diferencial de manera complementaria, tal y como se aprecia en la Figura 2.18. Los transistores adicionales utilizan dimensiones escaladas por un factor de proporcionalidad n (con respecto a los transistores del par diferencial), el cual debe tener un valor alto para que las variaciones debidas a las tensiones de entrada sobre los transistores escalados no afecten la suma constante de tensiones en gran manera y así lograr una alta linealidad en el circuito.

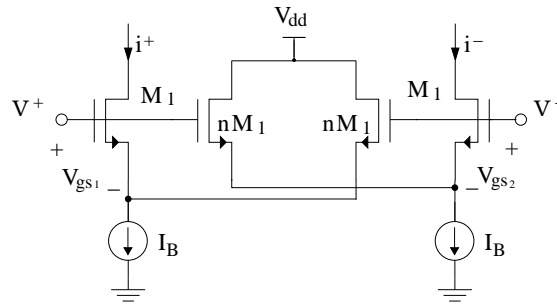


Figura 2.18: Implementación de Nedungadi del par diferencial con fuentes de tensión flotante.

Las ecuaciones que describen el circuito anterior son:

$$i_o = \left(\frac{n}{n+1} \right) 4\sqrt{k_1 I_B} V_d$$

$$G_m = \left(\frac{n}{n+1} \right) 4\sqrt{k_1 I_B} \quad (2.54)$$

La gráfica 2.19 muestra la respuesta del circuito de Nedungadi para diferentes valores de n . En esta se aprecia que la respuesta presenta un comportamiento mucho más lineal conforme se incrementa el valor de este parámetro, especialmente por encima de $n=5$. El compromiso que presenta este circuito está en su alto consumo de energía, ya que las fuentes deben suministrar una corriente total de $2(n+1)I_B$, por lo que el grado de linealidad deseado impacta directamente en la potencia que este circuito emplea.

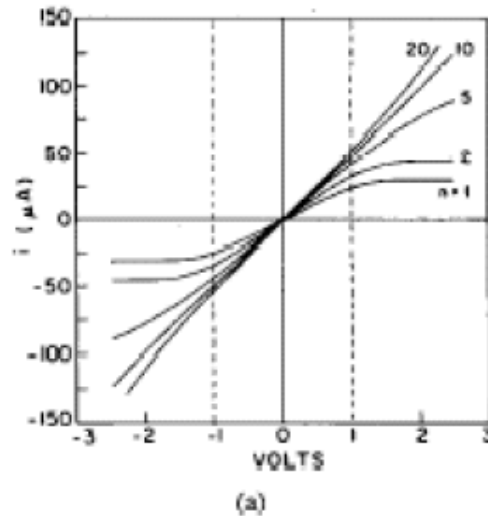


Figura 2.19: Gráficas de i_{out} vrs V_d para diferentes valores de n (1, 2, 5, 10 y 20, respectivamente). Tomado de [2]

2.6. Filtros GmC

Una de las principales aplicaciones de los OTAs radica en el diseño de filtros de grandes constantes de tiempo que requieren ya sea una alta resistencia y/o una baja capacitancia. Un filtro GmC, como su nombre lo indica, utiliza una transconductancia y un condensador en una configuración similar a la de los filtros RC para la obtención de filtros de diversos tipos y órdenes. Un filtro GmC de primer orden se presenta en la Figura 2.20.

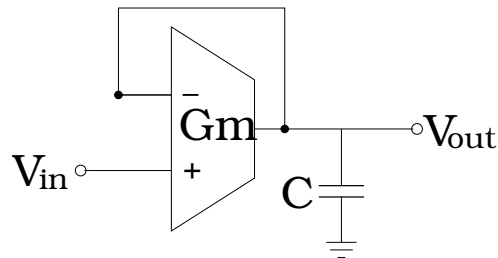


Figura 2.20: Filtro GmC de primer orden

La función de transferencia del circuito anterior viene dada por:

$$\frac{V_{out}}{V_{in}} = \frac{1}{\left(1 + s\frac{C}{G_m}\right)} \quad (2.55)$$

La frecuencia de corte (f_c) del filtro pasabajos anterior se localiza en:

$$\omega_c = \frac{G_m}{C}$$

$$f_c = \frac{G_m}{2\pi C} \quad (2.56)$$

Este tipo de filtros son los que se utilizaron como base en la implementación de la CWT mostrada en la Figura 1.2, en donde el escalamiento de la transconductancia permitió generar filtros con frecuencias de corte diferentes.

2.7. Espejos de corriente para el escalamiento de transconductancias

La búsqueda de transconductancias cada vez más pequeñas ha producido el desarrollo de muchas estructuras de OTAs distintas. El problema radica que entre más compleja sea una estructura, más complejos serán sus problemas de ruido y *mismatch*, así como el área efectiva y los compromisos de diseño entre los parámetros. Arnaud [14] emplea la técnica del uso de asociaciones serie-paralelo de transistores en espejos de corriente para escalar transconductancias y mantener grandes rangos lineales de entrada. También menciona que esta técnica representa una excelente solución en términos de los compromisos entre las características del OTA y permite agilizar el proceso de diseño de este tipo de circuitos.

2.7.1. Asociaciones serie-paralelo de transistores

En la Figura 2.21 se utilizan n -transistores idénticos conectados en paralelo, cuyas terminales G, D y S están conectadas entre sí. La corriente I_D que atraviesa este arreglo de transistores se divide equitativamente entre los mismos, por lo que el transistor equivalente puede verse ya sea como un transistor por el que atraviesa una corriente n -veces equivalente la corriente que atraviesa uno de los transistores individuales (con dimensiones W_P/L_P), o bien, como un transistor cuyo ancho es n -veces el ancho de uno de los transistores individuales. Las ecuaciones que muestran el principio anterior son:

$$I_D = nI_{D_n} \quad (2.57)$$

$$\frac{W_{eq}}{L_{eq}} = n \frac{W_P}{L_P} \quad (2.58)$$

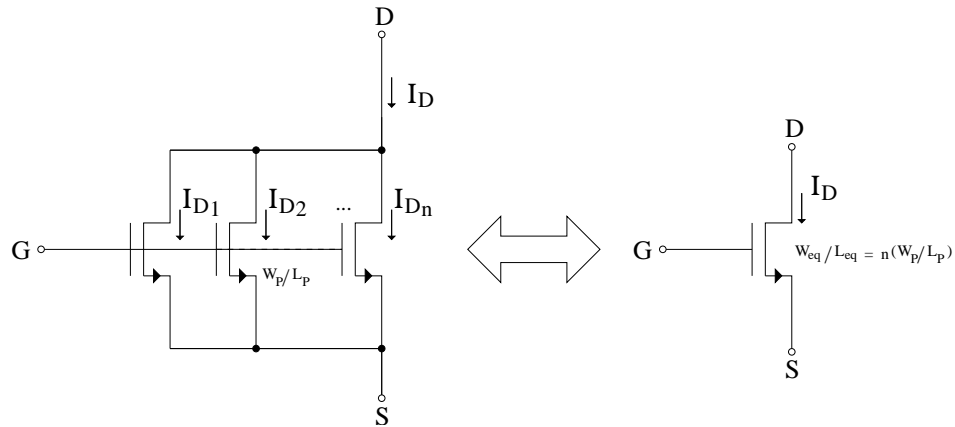


Figura 2.21: Configuración de transistores en paralelo y su equivalente.

En la Figura 2.22 se emplean n -transistores idénticos conectados en serie, uno tras otro con la terminal G interconectada entre ellos. La misma corriente nI_D atraviesa a todos los transistores, por lo que cada uno de ellos presenta tensiones V_{GS} y V_{DS} equitativamente reducidas entre sí. El transistor equivalente presenta entonces una corriente n -veces reducida a la que recorre el arreglo, o bien, representa un transistor n -veces más largo que uno de los transistores individuales. Matemáticamente lo anterior se representa como:

$$I_D = \frac{I_{Dn}}{n} \tag{2.59}$$

$$\frac{W_{eq}}{L_{eq}} = \frac{1}{n} \frac{W_P}{L_P} \tag{2.60}$$

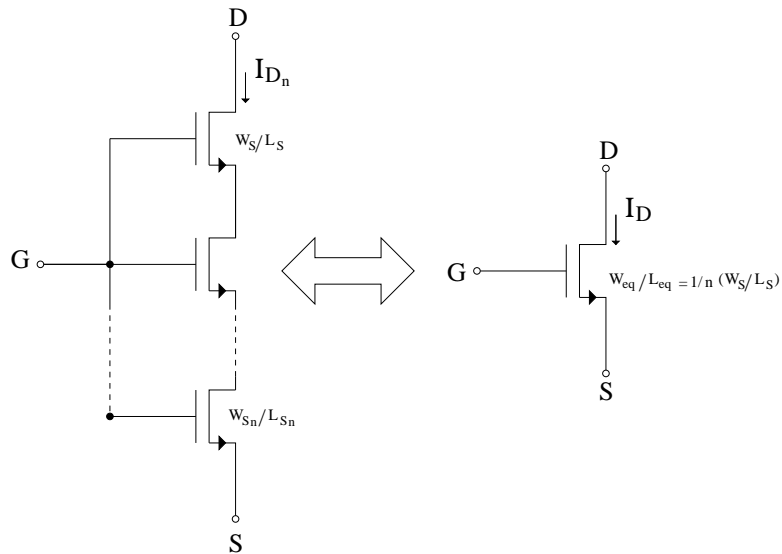


Figura 2.22: Configuración de transistores en serie y su equivalente.

2.7.2. Espejos de corriente utilizando asociaciones serie-paralelo de transistores

Combinando los esquemas en paralelo y serie de transistores (Figuras 2.21 y 2.22, respectivamente) dentro de la configuración básica de un espejo de corriente, es posible escalar el valor de la corriente. La Figura 2.23 muestra la configuración de este espejo de corriente.

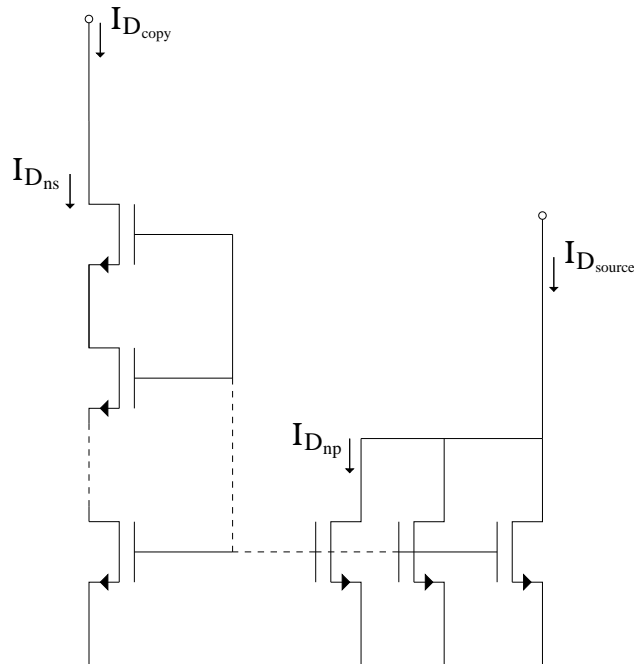


Figura 2.23: Espejo de corriente utilizando asociaciones de transistores serie-paralelo.

La corriente de entrada $I_{D_{source}}$ pasa a través de un arreglo de n_p -transistores en paralelo, lo que produce que la corriente que se va a copiar sea $I_{D_{np}} = I_{D_{source}}/n_p$. Del lado del arreglo de n_s -transistores en serie se realiza la copia de la corriente, es decir, $I_{D_{ns}} = I_{D_{np}}$, lo que produce que la corriente equivalente del arreglo en serie según la ecuación 2.59 sea $I_{D_{copy}} = I_{D_{ns}}/n_s$. Combinando los hechos anteriores puede verse que la corriente copiada por el espejo respresenta una proporción de la cantidad de transistores colocados en serie y en paralelo tal y como muestra la siguiente ecuación:

$$I_{D_{copy}} = \frac{I_{D_{source}}}{n_p \cdot n_s}$$

en donde para los efectos del presente informe el producto de la cantidad de transistores en serie y paralelo se denomina factor de escalamiento y se denota como $m = n_p \cdot n_s$.

2.8. Herramientas de diseño y simulación de circuitos integrados

2.8.1. Simuladores SPICE y modelos MOS

El lenguaje SPICE (Simulation Program with Integrated Circuits Emphasis, por sus siglás en inglés) es el estándar internacional para la simulación de circuitos electrónicos. Permite la

implementación de circuitos que utilicen todo tipo de componentes pasivos, activos, fuentes y demás. Actualmente la versión en uso es la SPICE3. Su código es completamente abierto y está disponible en los recursos de la Universidad de Berkley, California, lugar en donde fue desarrollado en 1975.

En la actualidad, se han desarrollado muchos programas y aplicaciones que implementan gráficamente este lenguaje. Estos programas interpretan los circuitos que el usuario traza y genera un archivo llamado *netlist*, el cual contiene las instrucciones e instancias equivalentes en SPICE. La variedad de pruebas que se pueden realizar abarcan desde análisis de puntos de operación, así como análisis de transitorios en AC, respuesta en frecuencia, FFTs y muchas más.

LTSPICE IV es uno de estos simuladores creado por la compañía Linear Technology, la cual se especializa en el diseño y fabricación de circuitos analógicos y de señal mixta de alto desempeño. Este simulador es de acceso libre a través de la página web de la compañía y es uno de los simuladores gratuitos más completos.

Una de las particularidades de los simuladores SPICE es la posibilidad de utilizar diferentes modelos para los componentes empleados en los circuitos. Así por ejemplo, se pueden dar instrucciones para que el simulador utilice los parámetros de un transistor obtenidos para un proceso de fabricación en particular y se pueden variar los modelos de ecuaciones que se utilizan para la simulación de los circuitos. Algunos de estos modelos de ecuaciones válidos, según la documentación del LTSPICE IV, son:

1. Modelo de Shichman-Hodges [15]: es uno de los modelos más básicos para los transistores MOS. Sus ecuaciones son muy similares a las presentadas para la región de inversión fuerte (ver sección 2.2) y son utilizadas en el cálculo básico de circuitos.
2. Modelos BSIM [16]: los modelos BSIM son desarrollados por la Universidad de Berkley, California. El diseño de circuitos utiliza ampliamente las diversas versiones de estos modelos, los cuales se caracterizan por emplear más de 100 parámetros en los cálculos, lo que lo hace muy complejo para el diseño a mano, pero si muy precisos para las herramientas computacionales. Compañías como Intel, IBM, TI, HP y Motorola utilizan estos modelos para el diseño de sus productos, según menciona la página oficial del grupo de desarrollo de estos modelos. La versión más moderna es la BSIM4, la cual modela los efectos físicos de los transistores MOSFET por debajo del régimen de los 100nm.
3. Modelo EKV: el modelo propuesto por Enz, Krummenacher y Vittoz (ver [17]) es reconocido por su mayor precisión con respecto al modelo de Shichman-Hodges y no es tan complejo como los modelos BSIM que incluso permiten ser utilizados en el diseño a mano de circuitos con relativa precisión. Este modelo representa la base del modelado de transistores MOS dado por Tsividis [8] y visto en las ecuaciones de las secciones 2.2 y 2.3.

2.8.2. Mentor Graphics

En el campo de las herramientas automatizadas de diseño electrónico (*EDA tools*, por sus siglas en inglés) existen grandes compañías que dominan este sector. Abate [18] reporta que para el 2008, Mentor Graphics, Synopsis y Cadence dominan un 64 % del mercado de este tipo. Estas compañías, conocidas como las Tres Grandes (the Big Three) se han formado a partir de la constante adquisición de pequeñas compañías que desarrollan productos que estas compañías no posean y que esten requiriendo sus clientes (ver [19]).

Una herramienta EDA es un software especializado para el diseño de circuitos integrados y circuitos impresos. Usualmente son ambientes o *suites* de desarrollo completos que involucran una serie de programas con funciones más específicas. Así por ejemplo, algunos programas se encargan de aspectos propios del diseño, otros sobre la verificación de los circuitos, simulación y visualización de resultados, etc.

ICstudio es el ambiente de desarrollo de circuitos integrados (*IC design*) de Mentor Graphics. En la Figura 2.24 se aprecia la relación entre los diversos programas que forman este *suite*, los cuales se utilizan según la etapa del proceso de diseño.

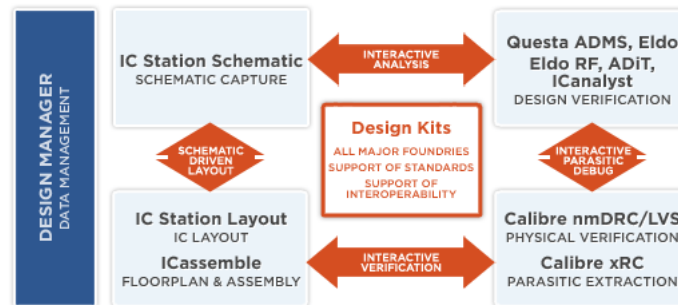


Figura 2.24: Estructura general del ambiente de desarrollo de Mentor Graphics. Tomado de [3].

La interfaz principal del ICstudio se muestra en la Figura 2.25⁶. En ella se aprecia la esquematización de un proyecto mediante librerías que contienen los módulos de los circuitos y todos aquellos componentes que se quieran diseñar. Los módulos a su vez están conformados por vistas, las cuales son las diferentes representaciones de un circuito en particular. Las vistas más frecuentes son:

1. Esquemático: como su nombre lo indica, es la representación física del circuito mediante componentes (transistores, resistencias, etc).
2. Símbolo: cuando un módulo se instancia dentro de un circuito más grande, se utiliza la vista símbolo para representar dicha instancia, como si fuera un componente nuevo. Esta vista puede editarse a conveniencia y usualmente se traza después de haber generado un esquemático, un archivo SPICE o por medio del lenguaje Verilog.
3. Verilog: los circuitos digitales pueden implementarse directamente en lenguaje Verilog sin la necesidad de trazar primeramente un esquemático del mismo. El software también es capaz de generar un símbolo adecuado según los pines que el circuito posea.
4. Layout: esta vista muestra el plano físico del circuito, es decir, los componentes a nivel de los materiales que los componen (sustratos, pozos, metales, etc).
5. SPICE: la vista SPICE usualmente se obtiene luego de generar el *netlist* del circuito, ya sea previo a la simulación del esquemático o como un archivo *post-layout* que permite simular con mayor precisión el comportamiento que llegará a tener el circuito una vez que este sea construido.

⁶Las imágenes que muestran pantallas e interfaces fueron tomadas a los diferentes programas que integran el ICstudio, todos ellos propiedad de Mentor Graphics Corporation.

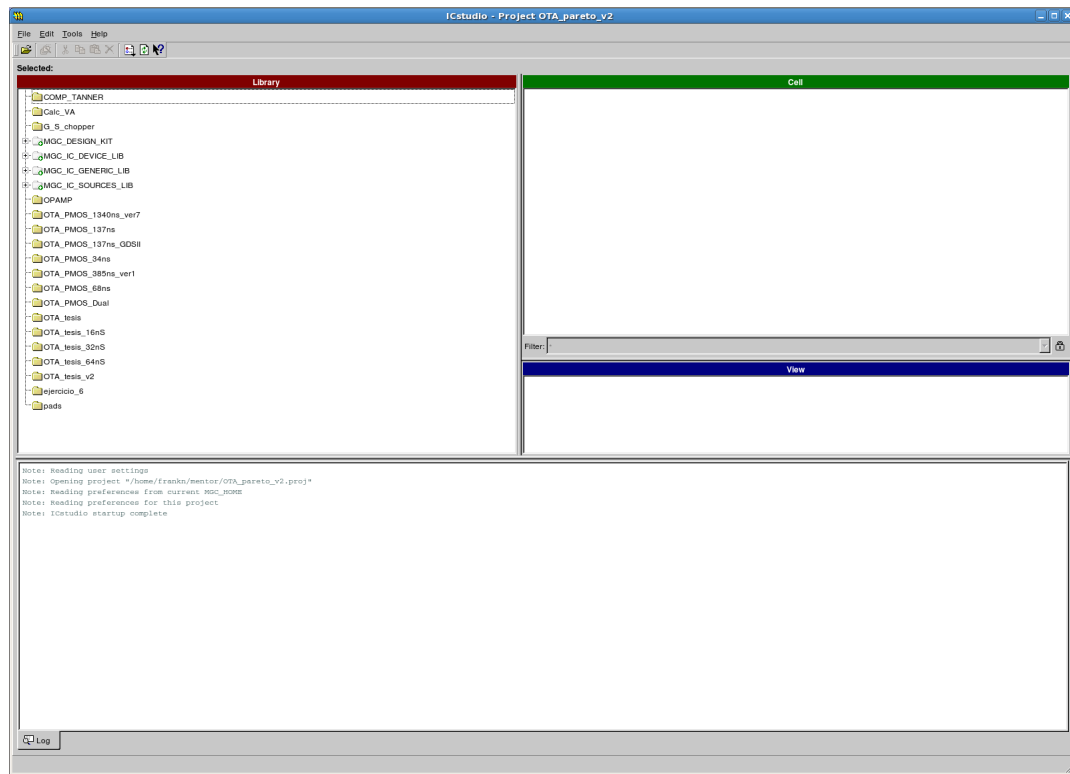


Figura 2.25: Interfaz principal del ICStudio.

Las secciones siguientes listan las herramientas del ICStudio que serán utilizadas a lo largo del proyecto.

2.8.2.1. Design Architect-IC (DA-IC)

Este programa se ejecuta al crear una vista esquemática y permite la creación de circuitos a nivel de componentes. Su interfaz es la que se aprecia en la Figura 2.26 y sus funcionalidades permiten el instanciamiento de módulos previos, utilización de componentes básicos (fuentes, elementos pasivos y activos, etc) y la modificación de sus propiedades. Desde este programa, mediante el comando Simulation se ingresa al ambiente de simulación donde pueden configurarse las pruebas a realizar y las señales que se han de analizar, entre varias opciones adicionales.

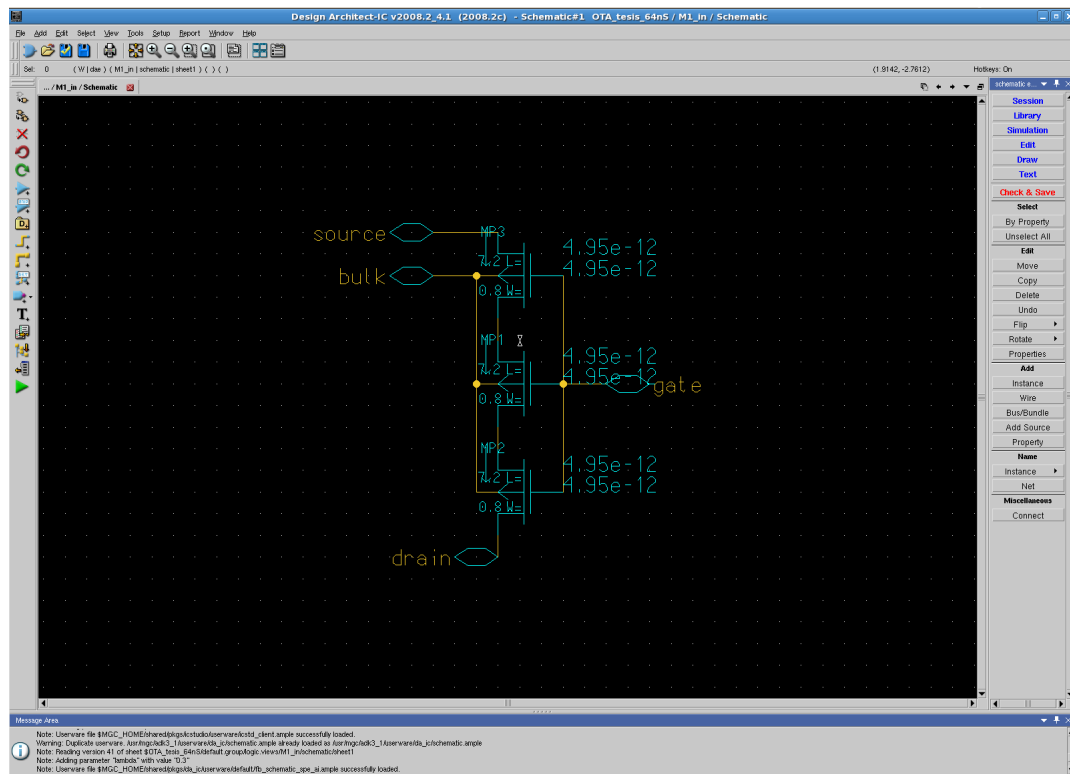


Figura 2.26: Interfaz del Design Architect.

2.8.2.2. Eldo Simulator y EZwave waveform viewer

Eldo es un motor de simulación analógica que emplea el lenguaje SPICE. Entre sus características destaca su alta velocidad de análisis con respecto a otros simuladores SPICE, su capacidad para analizar circuitos de gran escala (300 000 o más transistores), el uso de modelos de transistores más precisos, como el BSIM, PSP o HiSIM (ver [3]) y la inclusión de análisis de gran capacidad y variedad. Esta utilidad se encuentra integrada dentro del ambiente de desarrollo.

EZwave es una GUI (interfaz gráfica de usuario, por sus siglas en inglés) que permite la visualización de bases de datos de formas de ondas de análisis analógicos, digitales o de señal mixta. De esta manera, Eldo genera un archivo de simulación con los resultados de los análisis efectuados, el cual es interpretado por EZwave y despliega una pantalla como la de la Figura 2.27. Esta herramienta también permite la obtención de valores importantes de una curva (máximos, mínimos, valor rms, pendiente, etc) y operaciones matemáticas sobre estas gráficas.

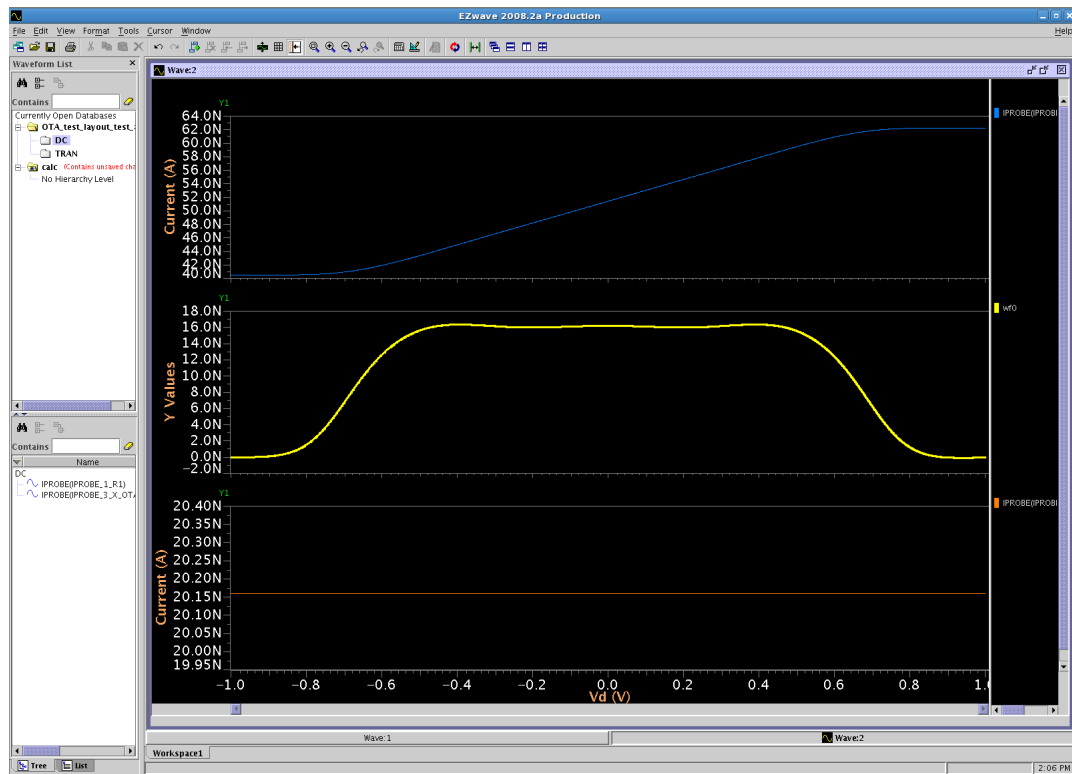


Figura 2.27: Interfaz del EZwave waveform viewer.

2.8.2.3. ICstation Layout

Esta herramienta permite el trazado y edición de layouts. Integra las funcionalidades del DA-IC en el hecho de que a partir de los esquemáticos el ICstation es capaz de trazar los componentes de acuerdo a las propiedades y dimensiones dadas e incluso trazar las conexiones de manera automática (aRoute), semiautomática (trazado hecho por el usuario con la guía del iRoute) o totalmente manual.

La Figura 2.28 muestra la interfaz del ICstation, en la cual se muestran las diferentes herramientas de edición y trazado de componentes, los diferentes materiales que se pueden utilizar y la vista de un circuito finalizado.

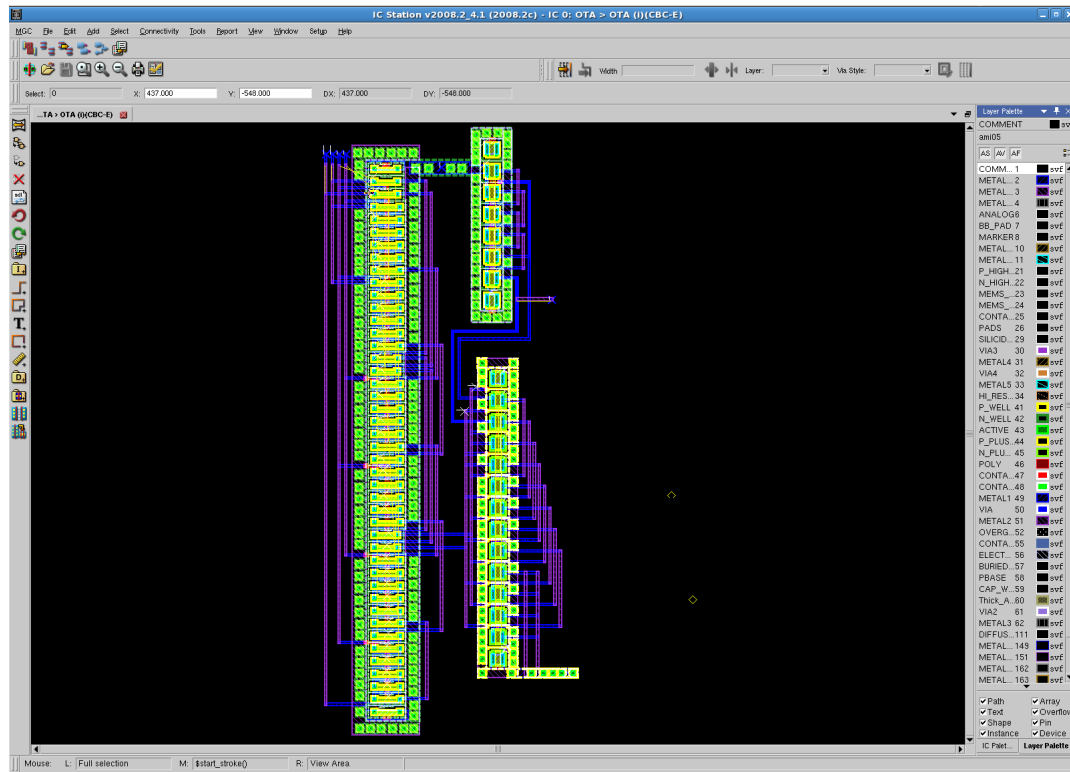


Figura 2.28: Interfaz del ICStation Layout.

2.8.2.4. Calibre IC Verification

Las herramientas que contiene Calibre permiten la verificación de la correcta generación del layout confeccionado. Esta verificación se da en tres niveles, cada uno con su propia aplicación Calibre. El primer paso es el chequeo de las reglas de diseño (Design Rule Checking o DRC, por sus siglas en inglés), el cual se basa en revisar que el layout cumpla con ciertas reglas que se proveen para un proceso de fabricación en particular. Estas reglas establecen el tamaño mínimo de las capas, la separación entre trazados de material idéntico, el traslape de materiales, etc y permiten que el layout generado se adapte a los estándares industriales de manera que el fabricante pueda efectivamente construir el circuito integrado según sus especificaciones. La Figura 2.29 presenta las ventanas que integran el Calibre nmDRC, aplicación que realiza este chequeo y que muestra de manera amigable la presencia de errores o no en el layout y la ubicación exacta de los mismos. En el caso de la Figura 2.29, el circuito cumple con las reglas del proceso y puede continuar al segundo paso de verificación.

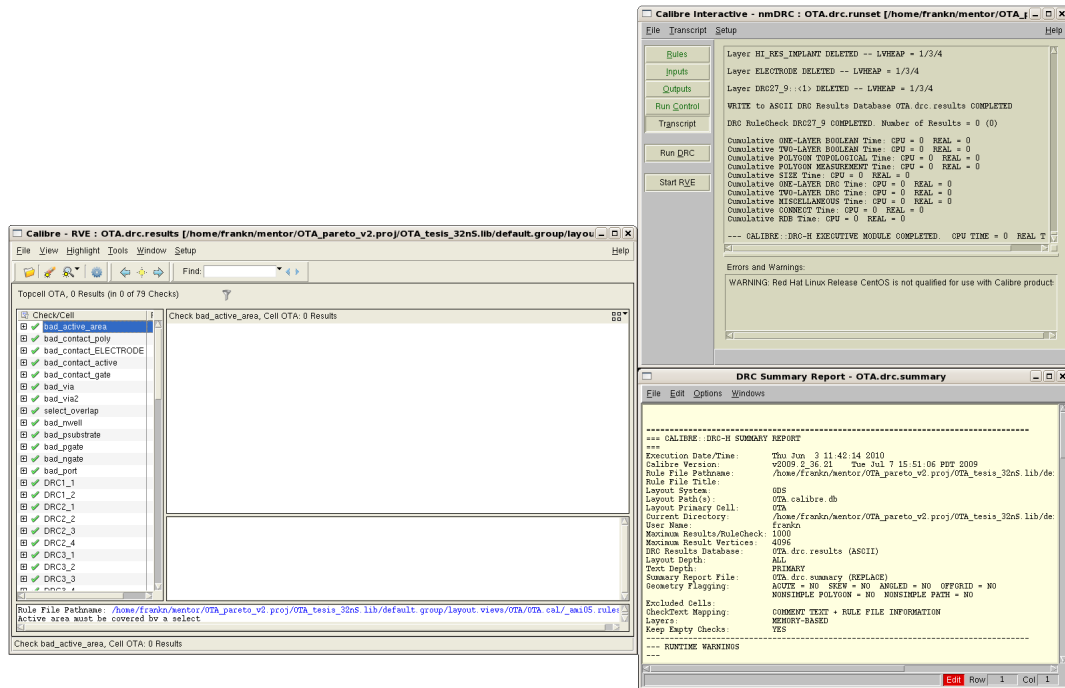


Figura 2.29: Calibre nmDRC.

El siguiente paso es una verificación de la concordancia entre el layout y su esquemático (Layout Versus Schematic verification, o LVS, por sus siglas en inglés). Este paso permite corroborar esencialmente tres cosas: si las conexiones hechas entre los dispositivos son correctas, si se colocaron todos los componentes que contenía el esquemático original y si se colocaron correctamente todos los pines. El programa que efectúa esta operación es el Calibre nmLVS y al igual que sucedía con el chequeo de reglas, este muestra las discrepancias entre ambas vistas para que el usuario pueda proceder a realizar las correcciones pertinentes. La Figura 2.30 muestra las ventanas de este programa, en la cual se obtuvo una revisión positiva, lo que permite su avance hacia una última verificación. Junto al Calibre nmDRC y Calibre nmLVS opera el Calibre RVE (Results Viewing Environment), la cual es la ventana que muestra el estado de los diversos chequeos.

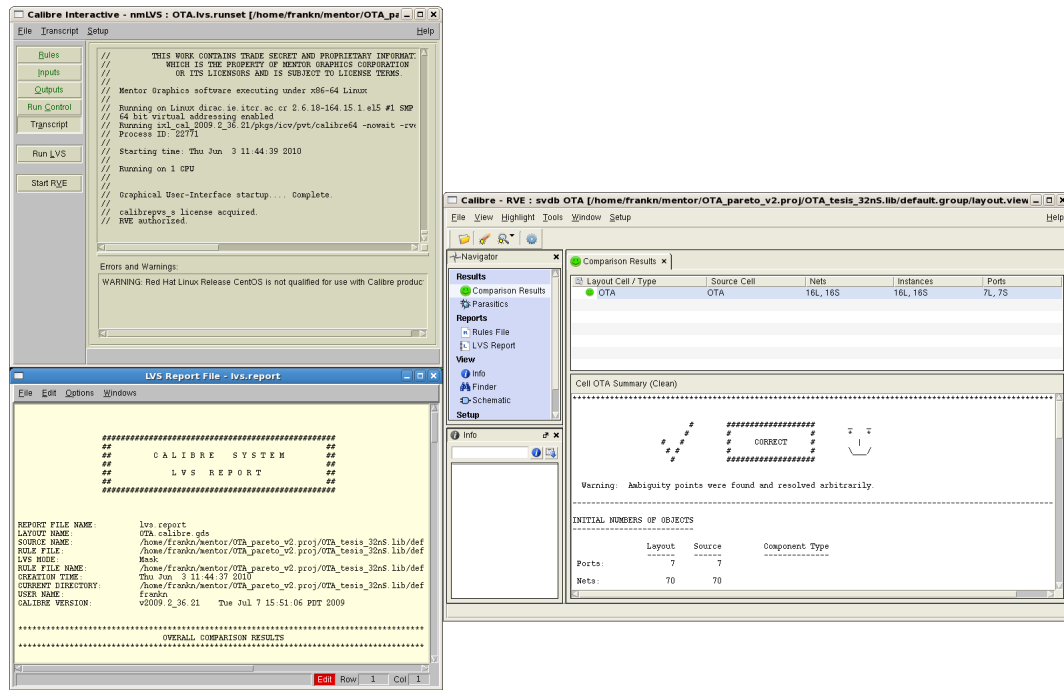


Figura 2.30: Calibre nmLVS.

El paso final consiste en una extracción post-layout (PEX, por sus siglas en inglés), la cual genera un netlist que incluye efectos parásitos (capacitancias y acoples RC no deseados), con el fin de realizar simulaciones post-layout que representarían con mayor fidelidad el comportamiento esperado del circuito integrado una vez este sea fabricado. El netlist generado debe asociarse nuevamente a la vista símbolo para que en los bancos de prueba pueda utilizarse este modelo SPICE del circuito en lugar del esquemático original. La Figura 2.31 muestra la interfaz del Calibre PEX junto con una parte del netlist que genera.

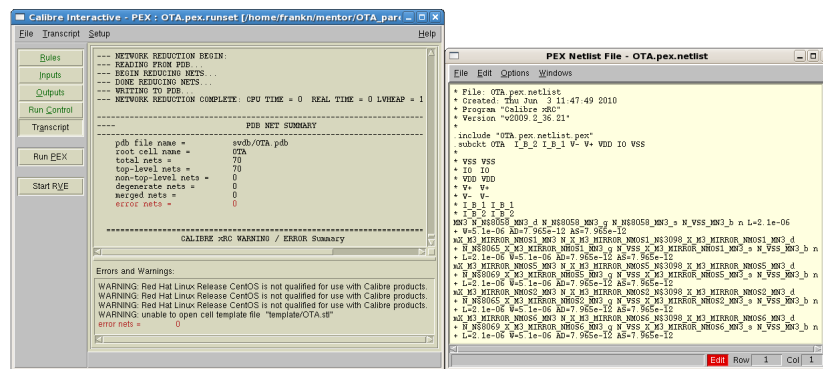


Figura 2.31: Calibre PEX.

2.9. Algoritmos genéticos y el Frente de Pareto

Supóngase que se tiene un circuito A con parametrización u cuya función de aptitud global está definida como (ver [4]):

$$F(A_u) = \Phi(f_1(A_u), f_2(A_u), \dots, f_n(A_u)) \quad (2.61)$$

Con las funciones de aptitud individuales $f_i(A_u)$ definidas de manera que estas sean crecientes con respecto a alguno de los parámetros del comportamiento del circuito. Estas funciones trazan un espacio de aptitud multidimensional, en el cual cada punto representa el desempeño del circuito A de acuerdo a los parámetros u para ese caso particular del espacio de parámetros.

Se asume que la forma general de Φ es usualmente desconocida aunque debe ser también creciente conforme aumente el valor de todas las funciones de aptitud f_i . Esta condición permite asegurar que un punto en el espacio de aptitud tiene mejor ajuste que otro si presenta valores menores en todas sus dimensiones.

En el contexto de la Figura 5.13, se dice que el punto q_1 tiene un mejor ajuste que cualquiera de los puntos dentro del rectángulo gris y por ello se dice que el punto q_1 *domina* al punto q_4 . Los puntos no dominados (q_1 , q_2 y q_3) forman lo que se denomina el Frente de Pareto de ese conjunto de datos. Una parametrización que pertenezca a este frente siempre representará una mejor opción que cualquier otro punto fuera de este ya que posee una mejor aptitud global. Matemáticamente lo anterior se expresa utilizando la siguiente ecuación:

$$\hat{P} = \{ \langle u \in P_A, f(A_u) \rangle \mid \neg \exists v \in P_A : f(A_v) \succ f(A_u) \} \quad (2.62)$$

en donde \hat{P} representa el Frente de Pareto, f el vector de funciones de aptitud $[f_1, \dots, f_n]^T$ y P_A es el espacio de parámetros del circuito A . El operador de orden parcial ' \succ ' muestra la propiedad de dominancia mencionada y se define como:

$$f_i(A_v) \succ f_i(A_u) \Leftrightarrow \forall i : f_i(A_v) \geq f_i(A_u) \wedge \exists i : f_i(A_v) > f_i(A_u) \quad (2.63)$$

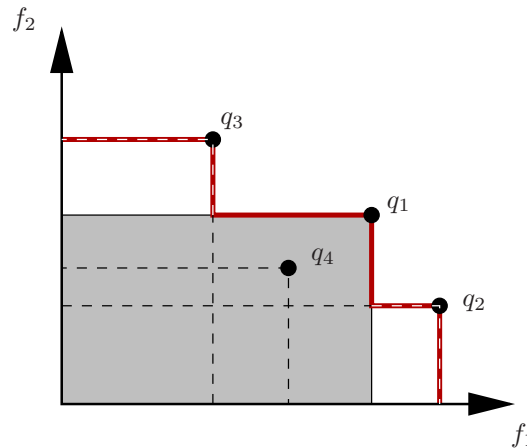


Figura 2.32: Frente de Pareto. Tomado de [4]

Cualquier algoritmo que determine el Frente de Pareto debe implementar las ecuaciones 2.62 y 2.63. Debido a la infinidad de parametrizaciones que el espacio de parámetros P_A pueda contener, es

necesario seleccionar un conjunto de estas parametrizaciones de manera que el Frente de Pareto que se pueda formar con este subconjunto sea una representación confiable del Frente de Pareto original. Una propuesta básica consistiría en muestrear discretamente cada parámetro, aunque la resolución de cada parámetro dispararía exponencialmente el número de evaluaciones si se selecciona una resolución muy alta.

Debido a la anterior, se utilizan algoritmos multiobjetivos como el PESA (Pareto Envelope-based Selection Algorithm, por sus siglas en inglés), que permiten descartar parametrizaciones ineficientes y concentrarse únicamente en aquellas regiones del espacio de parámetros que muestren resultados prometedores. Este algoritmo también aplica la discretización del espacio de parámetros, aunque no se corre el riesgo de aumentar desmedidamente el número de evaluaciones con resoluciones altas debido a la selección mencionada. La sección 2.9.1 muestra una breve reseña sobre el funcionamiento de este algoritmo.

2.9.1. El algoritmo PESA

Este algoritmo, propuesto por Corne, Knowles y Oates [20], utiliza no solo los parámetros comunes de los algoritmos genéticos como las tasas de cruce y mutaciones, sino que incluye dos parámetros adicionales: P_I , que representa el tamaño de la población interna (IP) y P_E , que representa el tamaño de la población externa (EP). A grandes rasgos, el algoritmo ejecuta los siguientes pasos:

1. Generar una población interna inicial de P_I cromosomas e inicializar la población externa con un conjunto vacío.
2. Incorporar los miembros no dominados del conjunto IP dentro del conjunto EP.
3. Utilizando un criterio de finalización, si este es alcanzado, entonces retorna el conjunto EP como resultado del algoritmo. De lo contrario, se vacía el conjunto IP hasta que se generen P_I nuevos candidatos. Estos candidatos se generan tomando dos padres del conjunto EP (con una probabilidad p_c), se cruzan para generar un hijo y este se muta. Con una probabilidad $(1 - p_c)$, se selecciona uno de los padres para mutarlo y producir otro hijo.
4. Se retorna al paso 2.

La incorporación de nuevos miembros al conjunto EP (paso 2) se da uno a la vez. Esto es, si se determina que este elemento no es dominado dentro de IP entonces es trasladado al conjunto EP y todos los elementos que dominaba dentro del conjunto IP son eliminados. Si el conjunto EP se llena (alcanza $1 + P_E$ elementos), entonces alguno de los elementos de dicho conjunto debe eliminarse. El elemento a eliminar es aquel que presente el factor de compresión más alto.

La selección de un padre en el algoritmo PESA (paso 3) se basa en la estrategia de apilamiento que posee este algoritmo. Esta estrategia consiste en la división del espacio de fenotipos en sectores idénticos, los cuales van a contener cierta cantidad de cromosomas dominados y no dominados. A cada cromosoma se le asigna un parámetro especial llamado factor de compresión y este se define como el número de cromosomas que habitan en el mismo sector. Un factor de compresión alto implica que en una misma región hay muchos cromosomas apilados, que eventualmente podrían representar datos ya repetidos. Con esto se orienta el desarrollo del Frente de Pareto hacia aquellas regiones con poca representación en la población.

Capítulo 3

Procedimiento metodológico

3.1. Descripción y síntesis del problema

Luego del envío a fabricación del circuito integrado de la etapa completa de detección se encontró que en el nodo de interconexión de los filtros (ver Figura 1.2) se formaba una capacitancia parásita que modificaba el polo en la frecuencia de interés, haciendo que las bandas pasantes no concordaran con las que se determinaron mediante la CWT y las mediciones de campo realizadas.

Cada OTA presentaba una capacitancia parásita de entrada de 2.2pF. Debido a la conexión en paralelo de los tres filtros, en el nodo de interconexión se presentó una capacitancia total de 11.21pF, lo que produjo un corrimiento de los polos de cada OTA y afectó las bandas pasantes de los tres coeficientes y la amplitud del coeficiente 5. Sumado a esto, cálculos posteriores indicaron que este exceso de capacidad también redujo el *slew rate* del OTA por debajo del requerido, lo que perjudicó también el comportamiento esperado.

Entre las causas de la aparición de esta capacitancia de entrada se determinó principalmente que la gran cantidad de transistores que componen tanto el par diferencial como el difusor añade efectos capacitivos que se acumulan, desembocando en el efecto indeseado mencionado con anterioridad.

A pesar de que el problema se detectó a través de las simulaciones realizadas, los problemas de *matching* entre los condensadores corrieron los polos aún más de lo esperado. Debido a esto, se manejó la posibilidad de ajustar la corriente de polarización del par de transistores de entrada, para así modificar las condiciones de polarización de los mismos y obtener el comportamiento primeramente esperado. Dicho ajuste efectivamente colocó los polos en la ubicación correcta, pero ocasionó que el consumo de potencia del filtro se triplicara según las proyecciones iniciales.

Por otra parte, la relación 7:1 establecida por Krummenacher hace que el circuito aumente considerablemente su tamaño, máxime que el diseño de cada transistor involucra una concatenación de transistores de manera que en conjunto adquieran el comportamiento esperado. Tanto la relación de tamaño como el ajuste de la corriente inciden en la zona de inversión de los transistores, lo que determina el consumo de potencia que va a tener el circuito.

Al buscar reducir el problema de la capacitancia parásita también deberán considerarse simultáneamente otros factores importantes en el circuito, como lo es su consumo de potencia, el rango lineal de la transconductancia obtenida, el *slew rate*, las dimensiones de los transistores, etc. Debido a la interrelación existente entre todos los parámetros, se tiene un problema

de diseño que debe abarcarse desde una óptica de optimización múltiple de parámetros.

El problema a resolver se ha sintetizado en la siguiente pregunta:

¿Cómo mejorar los parámetros de un OTA (G_m , capacitancia parásita de entrada, slew rate, consumo de potencia, *matching* y tamaño) utilizado en la implementación del filtros para la detección de disparos de armas de fuego?

3.2. Evaluación de alternativas

3.2.1. Sustitución del difusor simétrico

La solución propuesta preliminarmente se enfoca en sustituir el difusor simétrico por otro componente que permita mantener o mejorar las condiciones de linealidad hasta ahora logradas, pero que a su vez utilice menos recursos que el difusor para así lograr una reducción de la capacitancia parásita.

Un planteamiento preliminar se enfoca en sustituir el difusor por un transistor CMOS con tensión en modo común, tal y como lo muestra la Figura 3.1 y cuyo funcionamiento se mostró en la sección 2.5.3.

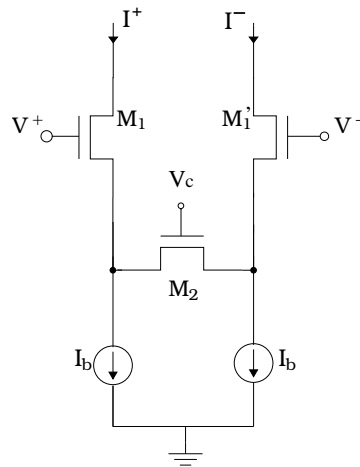


Figura 3.1: OTA con transistor con polarización en modo común.

Este circuito vendría a reducir la cantidad de transistores utilizados, ya que si en esta situación se emplea una relación 1:1 con los transistores del par de diferencial, se reduciría en un 67% la cantidad de transistores necesarios (en el OTA con difusor se necesitan 9 transistores, mientras que en el esquema propuesto se utilizan solo tres). Esta minimización en los componentes impactaría directamente en el tamaño, pero primordialmente reduciría la presencia de componentes con efectos capacitivos, que es lo que en realidad se busca solucionar.

El nuevo OTA deberá tener una capacitancia de entrada no mayor a 0.2pF (10 veces menor que la capacitancia de entrada del OTA actual) para que no genere un efecto considerable en las bandas pasantes de los filtros paralelos. La utilización de menos componentes implica directamente una reducción en el área ocupada por el OTA, la cual actualmente es de 990nm² (2200x450μm). Si bien la cantidad de componentes se reduce en un 67%, esto no implica la

misma reducción en área debido a que debe estudiarse y diseñarse con precisión la ubicación de los componentes para optimizar el funcionamiento del OTA.

Por otra parte, surgen nuevos desafíos como lo es la polarización de los OTAs para el ajuste de los filtros, ya que la tensión en modo común que se aplica al transistor M_2 es en algunos casos difícil de obtener y también estos ajustes tienden a elevar el consumo de potencia.

3.2.2. Software de optimización automática

Pereira, Alvarado y Krautschneider [4] desarrollaron una herramienta de optimización multiobjetivos basada en el uso de algoritmos genéticos (ver sección 2.9) para una librería de compuertas MCML (MOS Current Mode Logic circuits, por sus siglas en inglés). El empleo de algoritmos genéticos permite un nivel de abstracción mucho más alto, por cuanto la herramienta no necesita conocer la topología del circuito, sino un conjunto de valores de aptitud que representen variables dentro de ese circuito (consumo de potencia, respuesta ante algún estímulo, etc.). Dicha herramienta está escrita en C++.

El funcionamiento del programa comprende dos procesos independientes: la fase de representación del circuito y la fase de optimización. El esquema de la Figura 3.2 muestra ambas y los diferentes componentes de cada una. En la fase de representación se incluyen las herramientas de diseño de circuitos, que para el presente caso corresponden al Design Architect, que será en donde se diseñe el circuito y el simulador ELDO, para obtener el comportamiento del mismo. Para modelar los transistores, ELDO requiere los parámetros del proceso y el modelo de simulación, que por defecto utiliza el BSIM3v3, la última generación del modelo BSIM antes de modelar efectos a escalas nanométricas (ver sección 2.8).

La fase de optimización contempla la implementación propia del algoritmo genético (PESA), basado en los valores de aptitud determinados por el usuario, para luego generar un conjunto de parámetros que son realimentados al diseño para una nueva simulación del circuito y así generar un ciclo determinado por el número de evaluaciones que se deseen realizar.

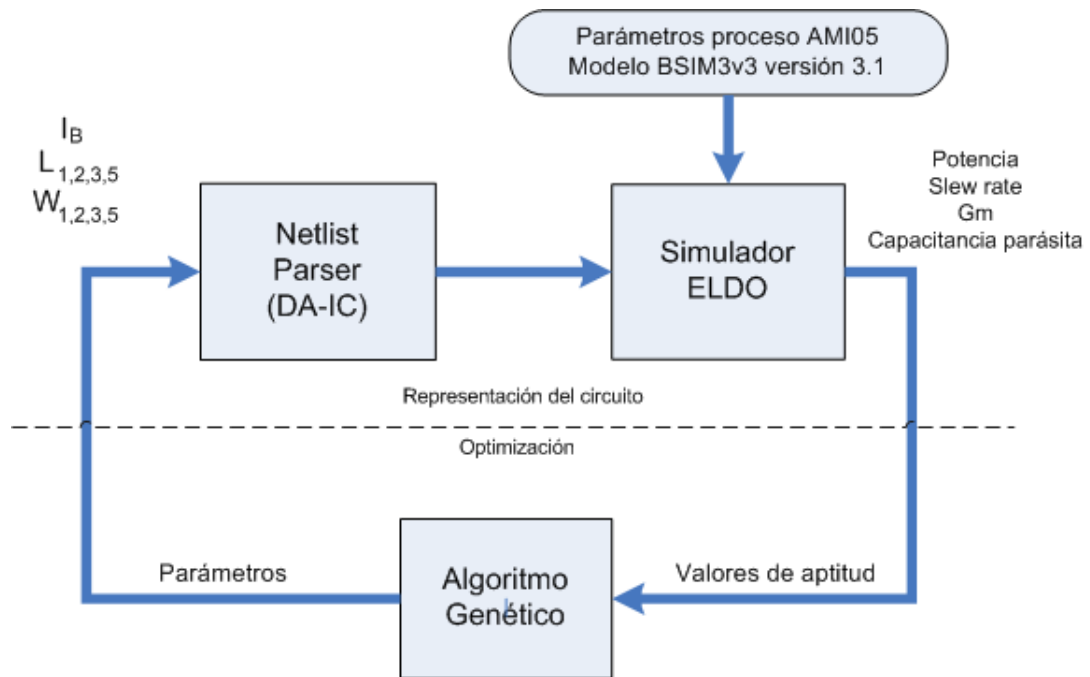


Figura 3.2: Arquitectura del optimizador de circuitos.

La posibilidad de emplear esta herramienta de optimización automática de circuitos surgió durante el desarrollo del proyecto, permitiendo complementar las labores de ajuste y verificación de circuitos. La herramienta no sustituye el proceso de diseño al requerir de un circuito inicial para luego generar un conjunto de valores optimizados según el Frente de Pareto respectivo. La decisión de elegir el mejor conjunto de parámetros sigue siendo tarea del diseñador, según los criterios que requiera satisfacer, aunque para ello ya cuenta con un número finito de posibilidades óptimas de elección.

Capítulo 4

Meta y objetivos

4.1. Meta

Indicador: pruebas de potencia con diversas condiciones de funcionamiento, que demuestren que el consumo de potencia del filtro completo no supere $1\mu W$, con una tolerancia del 10 %.

Reducir el consumo de potencia de la etapa de detección de disparos de armas de fuego de $21.46\mu W$ a $1\mu W$ mediante la optimización de los amplificadores operacionales de transconductancia.

4.2. Objetivos

4.2.1. Objetivo general

Indicador: simulaciones realizadas con Mentor Graphics, las cuales indiquen que los parámetros simulados no superen el 10 % de los valores establecidos dentro del objetivo general (criterio de ingeniería).

Diseñar un amplificador operacional de transconductancia de $137nS$ con una capacitancia de entrada menor o igual a $0.2pF$, un rango lineal de operación de $\pm 500mV$ y con un área menor o igual a $990nm^2$ que permita la implementación de filtros para la detección de disparos de armas de fuego.

4.2.2. Objetivos específicos

1. Diseñar los transistores CMOS necesarios para la construcción del OTA.

Indicador: Simulaciones en Mentor Graphics que muestren un consumo de potencia menor o igual a $28nW$ y un rango lineal de operación de $\pm 500mV$ con un margen de error máximo del 10 %.

2. Establecer la configuración del OTA más eficiente en términos de linealidad y transconductancia.

Indicador: Simulación de modelos en LTSPICE IV que comprueben las características deseadas con respecto a los valores teóricos para el OTA.

3. Caracterizar el comportamiento de la transconductancia en función de la tensión diferencial.
Indicador: gráficas de la función de transferencia corriente-tensión del circuito que reflejen una desviación no mayor al 10 % con respecto a los cálculos teóricos desarrollados.
4. Determinar el consumo de potencia del OTA.
Indicador: Banco de pruebas del Mentor Graphics, cuyos resultados muestren un consumo de potencia total del circuito menor o igual a $1,788\mu W$.
5. Escalar el OTA patrón para obtener dos layouts adicionales de 68.5nS y 34.25nS.
Indicador: gráficas de respuesta de los layouts que reflejen un comportamiento que no se desvíe en más de un 5 % con respecto al OTA patrón de 137nS.

Capítulo 5

Descripción detallada de la solución

5.1. Selección del esquema de linealización

La propuesta inicial de solución se basó en la sustitución del difusor simétrico por un transistor con polarización en modo común. No obstante, era necesario verificar si dicha propuesta realmente representaba una mejora para el OTA y máxime que al realizar el proceso de investigación se encontraron otros esquemas de linealización que podrían resultar útiles, incluyendo el difusor simétrico ya utilizado.

La Figura 2.15 muestra una comparación de tres de los esquemas presentados en la sección 2.5, pero antes de basarse solo en estos resultados, se procedió a realizar simulaciones de cuatro de los esquemas presentados (par diferencial simple, transistor con polarización en modo común, par con difusor simétrico y el par de Nedungadi con fuentes de tensión flotante).

Debido a la menor transconductancia intrínseca de los transistores PMOS (ver [1]), se utilizó este tipo de transistores en las simulaciones realizadas. Idealmente se busca una reducción en la transconductancia porque esto produce una reducción de la capacitancia necesaria para construir el filtro en el cual el OTA a implementar será utilizado. La mayor ventaja de esto es una reducción del área que ocupará el filtro.

Las Figuras 5.1 a 5.4 muestran los circuitos diseñados para simulación, los cuales presentan modificaciones a los presentados en la sección 2.5 debido a que utilizan transistores PMOS por la razón mencionada anteriormente. El espejo de corriente para la generación de la corriente de salida utiliza un factor de $m=1$ para todos los casos.

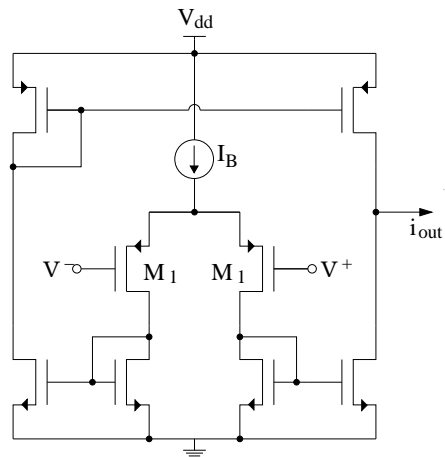


Figura 5.1: OTA simple con transistores PMOS

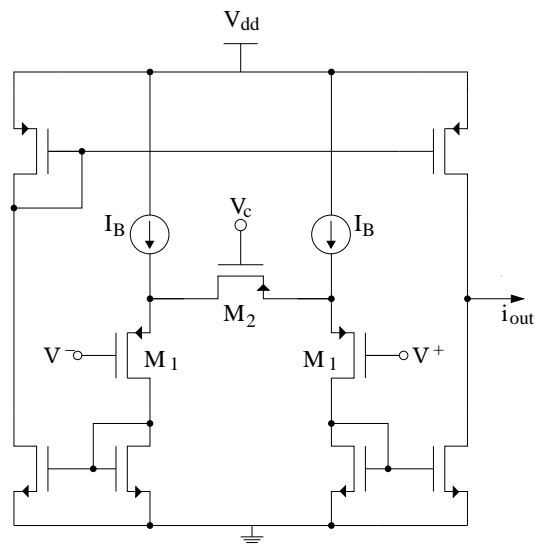


Figura 5.2: OTA con transistor en modo común, con transistores PMOS

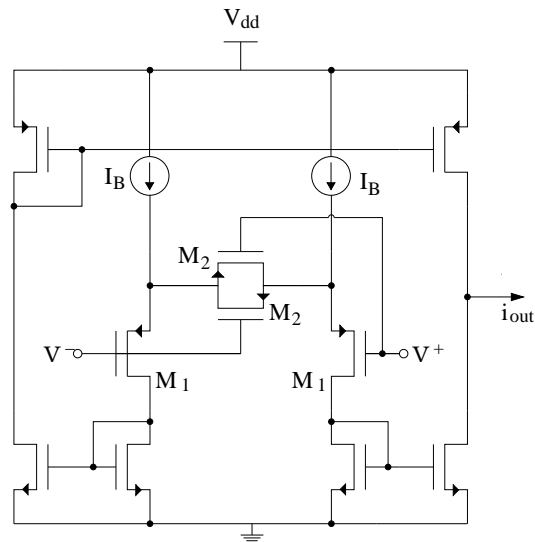


Figura 5.3: OTA con difusor simétrico, implementado con transistores PMOS

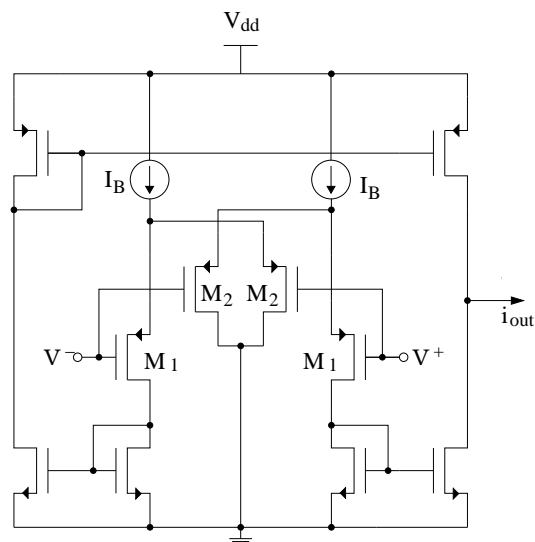


Figura 5.4: OTA de Nedungadi utilizando transistores PMOS

Implementando los circuitos anteriores en el simulador LTSPICE y utilizando los modelos SPICE para el proceso AMI05, se utilizó el circuito de prueba que muestra la Figura 5.5, en el cual la fuente de entrada aplica un barrido de tensión de $\pm 2,5V$ como tensión de entrada diferencial.

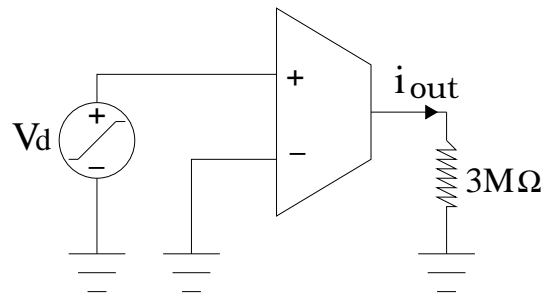


Figura 5.5: Circuito de prueba para OTAs para la obtención de la transconductancia

Mediante dicho circuito de prueba se obtuvo la gráfica de la Figura 5.6 en la que se obtiene la transconductancia de los cuatro circuitos en función de la entrada aplicada. Estas curvas se obtuvieron utilizando la definición de transconductancia que muestra la ecuación 2.43 y se colocaron sobre el mismo eje para apreciar las características que presentan entre sí.

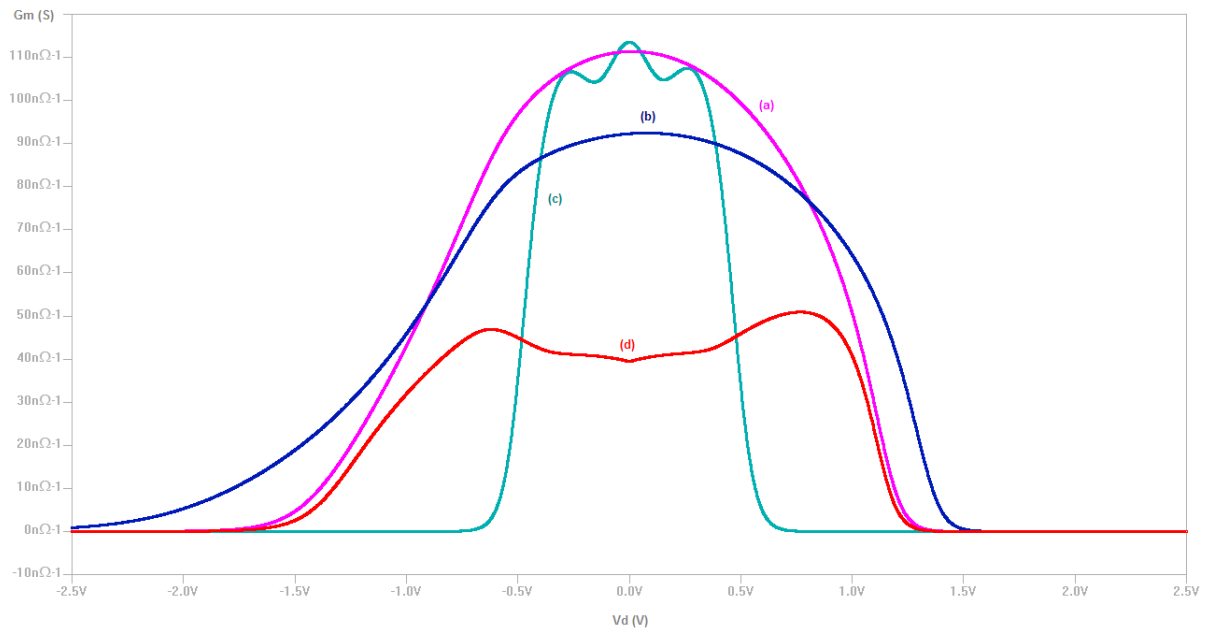


Figura 5.6: Gráfica de G_m versus V_d para los cuatro esquemas de OTA a comparar: a) OTA simple; b) OTA con transistor en modo común; c) OTA con difusores simétricos y d) OTA de Nedungadi

El diseño de los circuitos de simulación se basó en las ecuaciones respectivas presentadas en la sección 2.5. La Tabla 5.1 muestra los valores teóricos calculados de transconductancia para cada caso y los valores prácticos (valor en $V_d = 0$) que se muestran en la Figura 5.6.

Tabla 5.1: Valores teóricos y prácticos de G_m para cada configuración de OTA.

Configuración	G_m teórico (nS)	G_m práctico (nS)	Porcentaje de error (%)
OTA simple	136.01	111.34	18.14
OTA modo común	314.69	92	70.76
OTA con difusor simétrico	50.72	39.48	22.16
OTA de Nedungadi	403.84	113.5	71.89

El parámetro en común entre todos los circuitos fue la corriente de polarización I_B ya que se busca un circuito que presente poco consumo y que permita la obtención de los requerimientos presentados. Uno de los principales problemas presentados (ver sección 3.1) fue que el ajuste del circuito mediante la corriente de polarización elevó grandemente el consumo de potencia, por lo que es esencial que este parámetro sea lo más invariable posible, por lo que el ajuste de los circuitos de compaación se produjo a nivel de las dimensiones de los transistores.

La Tabla 5.2 provee la información que resume las dimensiones de los circuitos presentados y el comportamiento de los mismos arrojado por el simulador y la Figura 5.6. En todos los casos, los transistores que forman los espejos de corriente son todos idénticos y la razón de su dimensiones es $W/L = 10\mu\text{m}/20\mu\text{m}$.

Tabla 5.2: Resumen de las características de los OTAs simulados

Esquema	$M_1(\mu\text{m})$	$M_2(\mu\text{m})$	$I_B(\text{nA})$	$V_c(\text{V})$	Rango lineal (mV)	Potencia (nW)
OTA simple	1/200	—	100	—	± 231	990
OTA modo común	3/600	2/570	50	-3.3	± 370	990
OTA con difusor	3/80	2/570	50	—	± 404	990
OTA de Nedungadi	0,8/60	8/60	50	—	± 315	660

Para el caso del OTA simple, la corriente de polarización es del doble de las restantes debido a que esta fuente debe polarizar ambos transistores del par diferencial. En el caso del OTA de Nedungadi, en donde en la sección 2.5.5 se mencionaba que este es un circuito de alto consumo y los datos anteriores muestran un comportamiento opuesto, esto se debe a que la corriente que realmente polariza a cada transistor es diez veces menor. Si se utilizara una corriente de 550nA para poder mantener la misma corriente de polarización en los transistores del par diferencial con respecto a los otros casos, el consumo de potencia de este circuito pasaría de 660nW a 7.26mW. Este cambio en la corriente de polarización no se introdujo para mantener el valor de la transconductancia cercano al de los otros casos.

Los resultados preliminares muestran que el OTA con difusor simétrico es la mejor opción de las presentadas y por ello se eligió esta configuración para el diseño del OTA en cuestión. La sección siguiente muestra un análisis extensivo de la operación del dicho OTA, ya que no se encontró algún análisis como el que se efectuó y las ecuaciones de la sección 2.5 son válidas solamente para inversión fuerte y no son utilizables en el presente caso donde se desea utilizar la zona de inversión moderada.

5.2. Análisis matemático del OTA con difusor simétrico

El análisis realizado comienza con el establecimiento de las condiciones de operación debidas a la polarización en DC y a partir de estas se derivó el modelo en AC de pequeña señal para establecer la transconductancia equivalente del circuito. Para lograr esto, se utilizaron las ecuaciones del modelo general tanto para DC como para AC (ver secciones 2.2 y 2.3), siempre y cuando no se exprese lo contrario. Debido al uso de transistores PMOS en el circuito de la Figura 5.3, se reescribieron las ecuaciones necesarias adaptadas a estos transistores.

5.2.1. Análisis de polarización en DC

La Figura 5.7 muestra la etapa de entrada del OTA bajo la cual se efectuará el análisis de DC. En este caso se diferencian los transistores simétricos entre sí para diferenciar las corrientes y las tensiones en sus terminales.

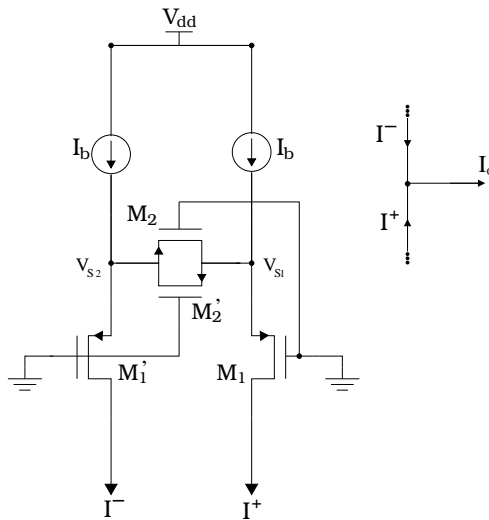


Figura 5.7: Polarización DC de la etapa de entrada diferencial del OTA

Aplicando LCK en el nodo V_{S1} se obtiene:

$$I_B + I_{SM2} = I'_{SM2} + I_{SM1} \quad (5.1)$$

Aplicando LCK en el nodo V_{S2} se obtiene:

$$I_B + I'_{SM2} = I'_{SM1} + I_{SM2} \quad (5.2)$$

Sumando 5.1 y 5.2 se obtiene:

$$I_B = \frac{I'_{SM1} + I_{SM1}}{2} \quad (5.3)$$

Restando 5.1 y 5.2 se obtiene:

$$I'_{SM1} - I_{SM1} = 2(I'_{SM2} - I_{SM2}) \quad (5.4)$$

Debido a la simetría del circuito y a las ecuaciones 5.3 y 5.4, se hacen las siguientes suposiciones:

$$V_{S1} = V_{S2}$$

$$I'_{SM1} = I_{SM1} = I_B$$

En las siguientes secciones se tratarán los transistores M_1 y M_2 por separado.

5.2.1.1. Polarización de M_1 .

La ecuación de la corriente según el modelo general para un transistor PMOS es la que se muestra en la ecuación 5.5. Se debe recordar que esta ecuación es aplicable para el transistor en la región de saturación.

$$I_S = I_Z \ln^2 \left[1 + \exp \left(\frac{V_{SG} - |V_{TP}|}{2n\phi_t} \right) \right] \quad (5.5)$$

Donde V_{TP} representa la tensión de umbral para un transistor PMOS. Este valor es naturalmente negativo para un transistor PMOS en modo incremental, por lo que se utiliza el valor absoluto de este para que se mantenga la analogía con la ecuación 2.25 correspondiente a un transistor NMOS.

La ecuación 5.6 se dedujo al despejar el término V_S a partir de la ecuación 5.5 y dejarlo en función de la corriente de polarización I_B , las dimensiones del transistor y las características propias de la tecnología de fabricación de los mismos. Esto se basa al considerar que $I_{SM1} = I_B$, $V_{SG} = V_S - V_G = V_S$ (las tensiones de entrada son puestas a tierra por estar en el análisis de DC) y que además M_1 se encuentra en saturación.

$$V_S = |V_{TP}| + 2n\phi_t \ln \left[\exp \left(\sqrt{\frac{I_S}{I_Z}} \right) - 1 \right] = |V_{TP}| + 2n\phi_t \ln \left[\exp \left(\sqrt{\frac{I_B}{I_Z}} \right) - 1 \right] \quad (5.6)$$

El cálculo de I_Z se conserva como se presentó en la ecuación 2.36. La ecuación 5.7 demuestra que la función en DC de los transistores M_1 radica en el establecimiento de la tensión del terminal de fuente (source) para determinar las condiciones de operación de los transistores M_2 .

Reescribiendo la ecuación 5.7 como la diferencia de tensiones que en ella se encuentran permiten obtener una expresión que resultará de utilidad posteriormente.

$$V_S - |V_{TP}| = 2n\phi_t \ln \left[\exp \left(\sqrt{\frac{I_B}{I_Z}} \right) - 1 \right] \quad (5.7)$$

5.2.1.2. Polarización de M_2 .

Para este transistor se desea que funcione dentro de la región de no saturación, para lo cual debe garantizarse que (ver sección 2.2):

$$V_{SD} \leq V'_{SD} \quad (5.8)$$

en donde V'_{SD} se calcula como:

$$V'_{SD} = \frac{V_{SG} - |V_{TP}|}{\alpha} \quad (5.9)$$

Utilizando las ecuaciones 5.7¹ y 5.9, puede reescribirse la condición de operación dentro de la zona de no saturación (ecuación 5.8) como:

$$V_{DS} \leq \frac{2n\phi_t}{\alpha} \ln \left[\exp \left(\sqrt{\frac{I_B}{I_Z}} \right) - 1 \right]$$

La cual siempre se cumple dado por simetría es posible verificar que $V_{S1} = V_{S2}$, lo que implica que para los transistores M_2 se tiene una tensión $V_{SD} = 0V$ y la función logarítmica siempre será real y positiva para valores de $I_B > 0$. Al verificarse la condición 5.8 se concluye que los transistores del difusor simétrico siempre operan dentro de la región de no saturación. El hecho de que $V_{SD} = 0V$ ocasiona que las corrientes por estos transistores sean cero², es decir, $I_{SM2} = I'_{SM2} = 0$, lo que mantiene la validez de las ecuaciones 5.3 y 5.4. Asimismo la corriente de salida en DC es cero también puesto que las corrientes I^+ e I^- son iguales y estas se restan a través del espejo de corriente.

5.2.2. Análisis de pequeña señal

Las obtenciones anteriores permiten encontrar que el transistor M_1 trabaja en la zona de saturación y M_2 en la zona de no saturación. Con ello, un modelo simplificado de operación en pequeña señal implicaría representar a M_1 solamente a través de su transconductancia g_m , la cual resulta más significativa por su relación con la tensión de entrada aplicada. M_2 por su parte, será representada solo por la transconductancia entre fuente y drenaje g_{ds} .

La Figura 5.8 contiene el circuito equivalente de pequeña señal y las ecuaciones posteriores definen las corrientes del circuito:

¹La ecuación 5.7 puede utilizarse para redefinir en este caso la condición mencionada pues la tensión V'_{SD} representa el límite entre la zona de saturación y la de no saturación.

²En la ecuación 2.24 del modelo general para todas las regiones esto ocasiona que los argumentos de ambos logaritmos sean iguales y por ende la corriente es cero en el dispositivo.

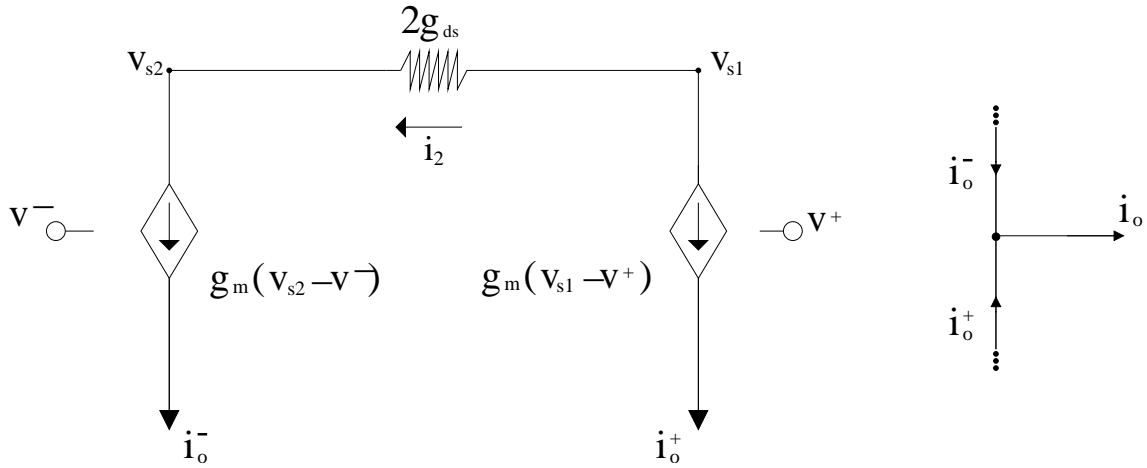


Figura 5.8: Circuito equivalente de la etapa de entrada (análisis de pequeña señal)

$$i_o^- = g_m (v_{s2} - v^-) \quad (5.10)$$

$$i_o^+ = g_m (v_{s1} - v^+) \quad (5.11)$$

$$i_2 = 2g_{ds} (v_{s1} - v_{s2}) \quad (5.12)$$

$$i_o^- = i_2 \quad (5.13)$$

$$i_o^+ = -i_2 \quad (5.14)$$

$$i_o = i_o^- - i_o^+ \quad (5.15)$$

$$i_o = 2i_2 \quad (5.16)$$

De 5.12 se obtiene:

$$\frac{i_2}{2g_{ds}} = v_{s1} - v_{s2} \quad (5.17)$$

Combinando 5.16 con 5.17 se obtiene:

$$\frac{i_o}{4g_{ds}} = v_{s1} - v_{s2} \quad (5.18)$$

Sustituyendo 5.10 y 5.11 en 5.15 se obtiene que:

$$i_o = g_m (v^+ - v^-) - g_m (v_{s1} - v_{s2}) \quad (5.19)$$

Sustituyendo 5.18 dentro de 5.19 y despejando para i_o , se obtiene la transconductancia equivalente del circuito:

$$i_o = g_m (v^+ - v^-) - g_m \frac{i_o}{4g_{ds}}$$

$$G_m = \frac{i_o}{(v^+ - v^-)}$$

$$G_m = \frac{g_m}{\left(1 + \frac{g_m}{4g_{ds}}\right)} \quad (5.20)$$

en donde g_m puede obtenerse utilizando la ecuación 2.26 y reescrita en la ecuación 5.21 con las variables del circuito propuesto.

$$g_m = \frac{I_B}{n\phi_t} \frac{1}{f(x)} \quad (5.21)$$

La transconductancia entre fuente y drenaje requiere un tratamiento especial, ya que las ecuaciones del modelo general de la sección 2.3.1 aplican solo para la zona de saturación. Además de esto, el cálculo de g_{ds} para este modelo general (ecuación 2.28) se basa en la ecuación 2.23, la cual en la sección 2.2 se comentó que la misma representa un aproximado de poca precisión para incluir el efecto de V_{DS} sobre I_B .

De los modelos presentados por Tsvividis [8] para la región débil y fuerte, el más adecuado es el de no saturación en inversión fuerte (ver Tabla 2.1), por cuanto los demás modelos (general e inversión débil) predicen una transconductancia $g_{ds} = 0$, lo que ocasionaría que la transconductancia del OTA según la ecuación 5.20 sea cero también. Esto contradiría las simulaciones de la sección 5.1, donde se obtuvieron OTAs perfectamente funcionales. La ecuación de g_{ds} modificada por su uso con transistores PMOS y para las condiciones del circuito de polarización (Figura 5.7) es:

$$g_{ds} = \left(\frac{W}{L}\right) \mu C_{ox} (V_S - |V_{TP}|) \quad (5.22)$$

Sustituyendo la ecuación 5.7 dentro de la ecuación anterior se obtiene la ecuación para g_{ds} en función de las dimensiones de los transistores, las características del proceso y la corriente de polarización.

$$g_{ds} = \left(\frac{W}{L}\right) \mu C_{ox} 2n\phi_t \ln \left[\exp \left(\sqrt{\frac{I_B}{I_Z}} \right) - 1 \right] \quad (5.23)$$

El análisis anterior no proporciona información sobre el rango lineal de la transconductancia equivalente. Arnaud [21] presenta una aproximación para el rango lineal (ΔV) del par diferencial básico (ecuación 5.24), que puede tomarse como base y ejemplificar la dependencia de este rango en términos de las características del circuito.

$$\Delta V \approx 3n\phi_t \sqrt{\alpha (1 + i_f)} \quad (5.24)$$

Donde i_f es la corriente normalizada y es un indicativo de la región de inversión de acuerdo al modelo ACM [22]. Para valores menores a 1, el dispositivo trabaja en la región de inversión

débil y para valores mayores a 100 trabaja en inversión fuerte. Valores intermedios corresponden a la zona de inversión moderada. El cálculo de i_f se presenta en la ecuación 5.25.

$$i_f = \frac{2 \cdot I_B}{\left(\frac{W_1}{L_1}\right) \mu C_{ox} n \phi_t^2} \quad (5.25)$$

Finalmente, la potencia máxima que debe suministrar la fuente depende tanto de las fuentes de polarización como de las corrientes escaladas que atraviesan los espejos de corriente, como se aprecia en la ecuación:

$$P_{max} = 2I_B V_{dd} \left(1 + \frac{2}{m}\right) \quad (5.26)$$

5.3. Diseño inicial del OTA

Antes de implementar la herramienta de optimización, resultó necesario diseñar un OTA inicial para la obtención de archivos con la información que requeriría el programa. Este diseño utilizó como base los transistores usados por Chacón [1], en donde para obtener cada transistor se utilizaron combinaciones de transistores unitarios en serie. Esto se hace con dos fines: reducir los efectos de difusión lateral (ver [23]), que ocasionan una disminución del área efectiva del canal y principalmente para mitigar los problemas relacionados al *mismatch* si se utilizaran transistores más grandes (ver [14]). En la sección 2.7 se mostró que esta configuración de transistores equivaldría a un transistor con un largo de tres veces al de los transistores unitarios.

Cada transistor del difusor simétrico es seis veces mayor al transistor del par diferencial. No se empleó la relación de siete veces de Krummenacher por dos razones: mediante esto se buscó minimizar un poco el área del circuito, ya que por las razones ya mencionadas, su tamaño con respecto a otros esquemas es mayor y además porque a través de las dimensiones de los transistores unitarios sería posible ajustar dicha relación. El material de Dualibe [2] muestra que la relación óptima de Krummenacher es que M_2 sea 6.7 veces más largo que M_1 .

Este diseño inicial también utilizó un factor de escalamiento $m = 4$, ya que de los OTAs requeridos, este es el que requiere más componentes, por lo que resultaba importante que el optimizador buscara mejorar estos también. Los transistores de los espejos de corriente también están conformados por tres transistores unitarios en serie.

La Figura 5.9 muestra el OTA inicial completo. Los transistores que tienen el mismo nombre son idénticos entre sí y representan: M_1 , para los transistores de entrada; M_2 , para los transistores del difusor simétrico; M_3 , para los transistores del espejo de corriente inferior y M_5 para los transistores del espejo de corriente superior (el que forma la corriente de salida).

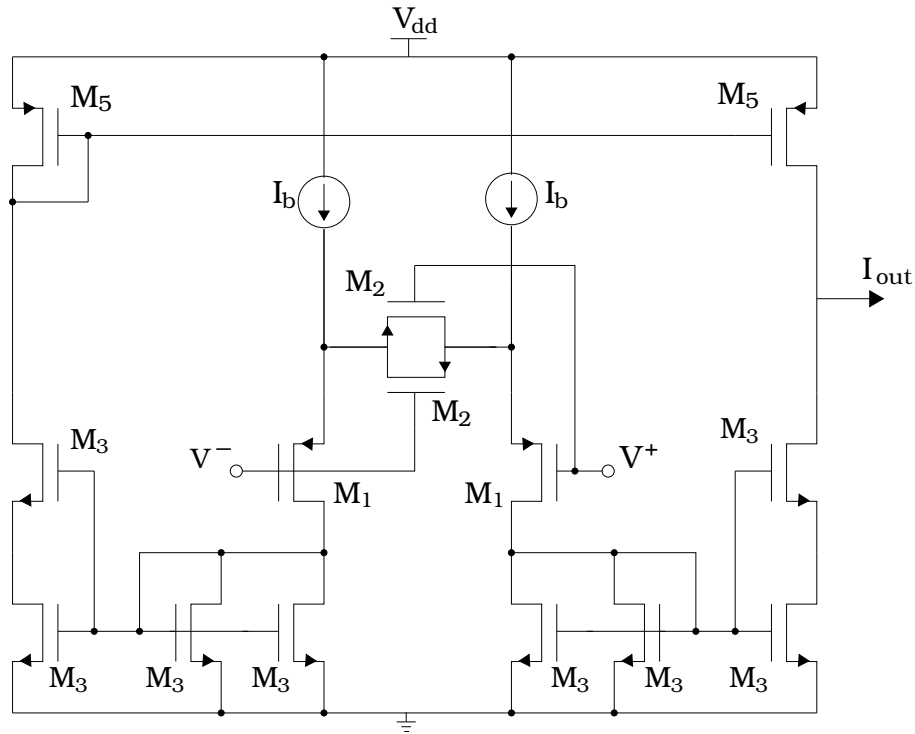


Figura 5.9: OTA inicial implementado.

La Tabla 5.3³ contiene los datos del diseño inicial, calculados empleando las ecuaciones obtenidas en la sección 5.2 y complementadas por las ecuaciones de las secciones 2.2 y 2.3. Al ser un punto de partida para la etapa posterior de optimización del mismo, se utilizaron todos los transistores unitarios idénticos ($W/L = 3\mu\text{m}/8\mu\text{m}$) con el fin de visualizar los cambios que efectuaba el optimizador a estos componentes.

Tabla 5.3: Parámetros de diseño del OTA inicial.

G_m (nS)	M_1 (μm)	M_2 (μm)	M_3 (μm)	M_5 (μm)	V_{dd} (V)	I_B (nA)
36	3/24	3/144	3/24	3/24	3.3	56

La simulación del circuito con las características anteriores deparó los datos que se muestran en la Tabla 5.4, la cual resume las especificaciones obtenidas mediante mediciones hechas a las gráficas obtenidas con el visor EZwave y a datos que que suministra ELDO en el registro de la simulación finalizada.

Tabla 5.4: Especificaciones del OTA inicial.

$G_{m_{max}}$ (nS)	Rango lineal (mV)	I_B (nA)	Potencia (nW)	Capacitancia parásita (fF)	SR (mV/ μs)
36.57	± 260	26	174.93	267.79	0.489

³Se utilizan las dimensiones del transistor equivalente para cada caso, aunque debe recordarse que todos estos transistores emplean combinaciones de transistores unitarios en serie y M_2 es seis veces más largo que M_1 por razones ya mencionadas.

5.4. Implementación de la herramienta de optimización

Las ecuaciones 2.37, 5.20, 5.21, 5.23, 5.24 y A.4 muestran la relación presente entre los parámetros y las especificaciones que se buscan para el circuito. La gran interrelación existente entre ellos muestra que para la obtención de un circuito óptimo se requiere de un proceso de optimización múltiple de parámetros, el cual debe darse simultáneamente para que se de esta manera se obtengan parámetros aceptables, es decir, que al mejorar alguno no se sacrifique otro.

En la sección 3.2.2 se introdujo una herramienta de optimización automática, cuya funcionalidad se adaptó eficientemente a las necesidades del proyecto. La tarea realizada con respecto a este programa consistió en la modificación de la misma para adaptarla al tipo de archivos que emplea Mentor Graphics, ya que fue implementada inicialmente para circuitos hechos en Cadence (ver [4]). Las secciones siguientes muestran las labores de adaptación realizadas y cómo permitieron la operación del optimizador basado en algoritmos genéticos (ver sección 2.9).

5.4.1. Parametrización y simulación del circuito

El objetivo del diseño inicial del OTA era el de obtener un circuito cuyas características se encontraran parametrizadas, es decir, representadas por medio de variables, las cuales puedan asignársele valores por medio del netlist correspondiente a este circuito. En el DA-IC se utilizaron los nombres L_1 , L_2 , L_3 , L_5 y W_1 , W_2 , W_3 , W_5 para identificar los cuatro tipos de transistores. Aunque en teoría los transistores del espejo de corriente solo afectan las especificaciones por medio del factor de escalamiento únicamente, se incluyeron dentro de los parámetros que genera el optimizador en caso de que las simulaciones mostraran una mejora en el circuito debido a esto. Dentro de los parámetros se incluyó también I_B al ser un valor fundamental en la determinación de todas las características del OTA. Esta parametrización puede apreciarse en la Figura 3.2, en donde se presentó la arquitectura del optimizador.

El DA-IC permite la inclusión de parámetros directamente por medio del comando *Add Parameters*, añadiendo una línea del comando `.PARAM` (lenguaje SPICE) al netlist correspondiente. No obstante, se realizó la modificación manual del netlist para que todos los parámetros utilizaran la misma del comando, por cuanto los métodos posteriores requerirán de este cambio para que el programa sea capaz de leer constantemente este archivo de manera correcta.

Para la verificación del funcionamiento de OTAs, se emplearon dos pruebas. La primera de ellas consiste en un barrido de la tensión de entrada diferencial que utiliza un circuito como el mostrado en la Figura 5.5. La Figura 5.10 muestra el esquemático generado en el DA-IC, el cual es utilizado para obtener las gráficas de i_{out} versus V_d y G_m versus V_d para todos los OTAs que se diseñaron como parte de la solución.

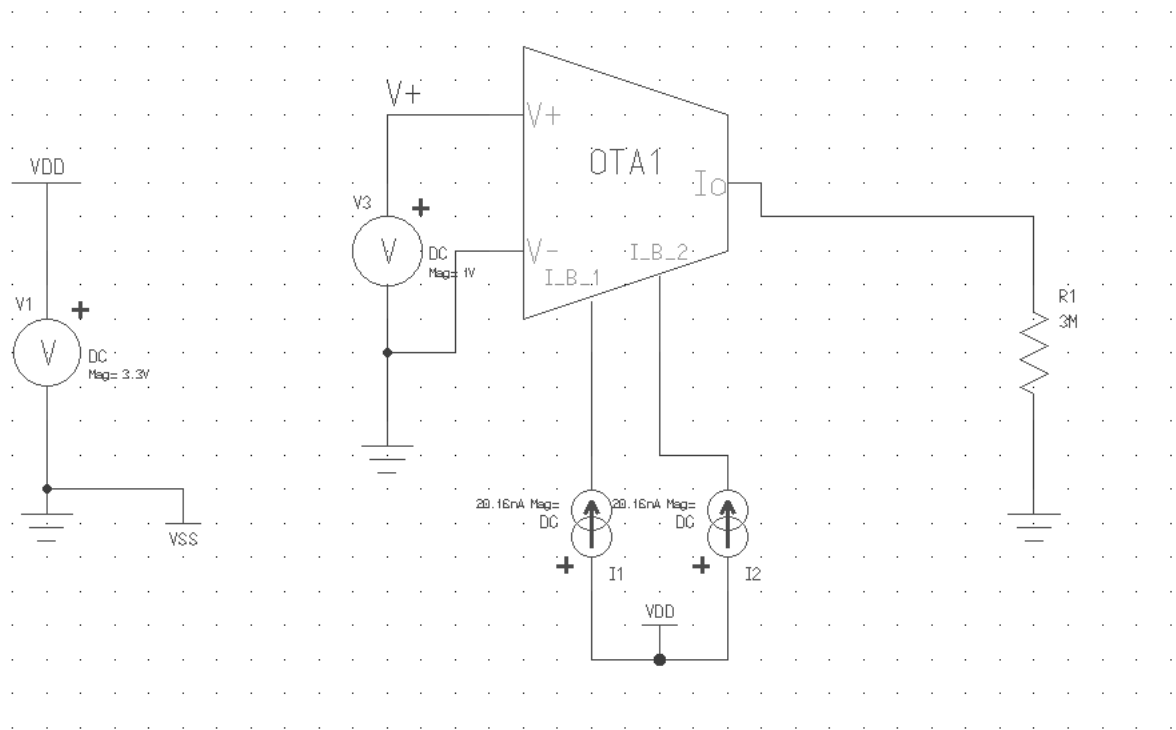


Figura 5.10: Prueba de DC para los OTAs.

La segunda prueba consiste en un análisis transitorio, con el fin de visualizar la forma de la corriente de salida ante una señal de entrada senoidal de 1V de amplitud. Esta prueba permite visualizar la deformación de la señal debido a las características no lineales del OTA, especialmente cuando la amplitud de la señal de entrada se acerca a los límites del rango lineal. La Figura 5.11 muestra el esquemático para esta prueba, cuya diferencia con la prueba de DC es el tipo de entrada que se aplica. Además, la prueba en transitorio permite a ELDO calcular las capacidades en cada nodo del circuito. Así se cuenta con datos para la optimización de la capacidad de entrada.

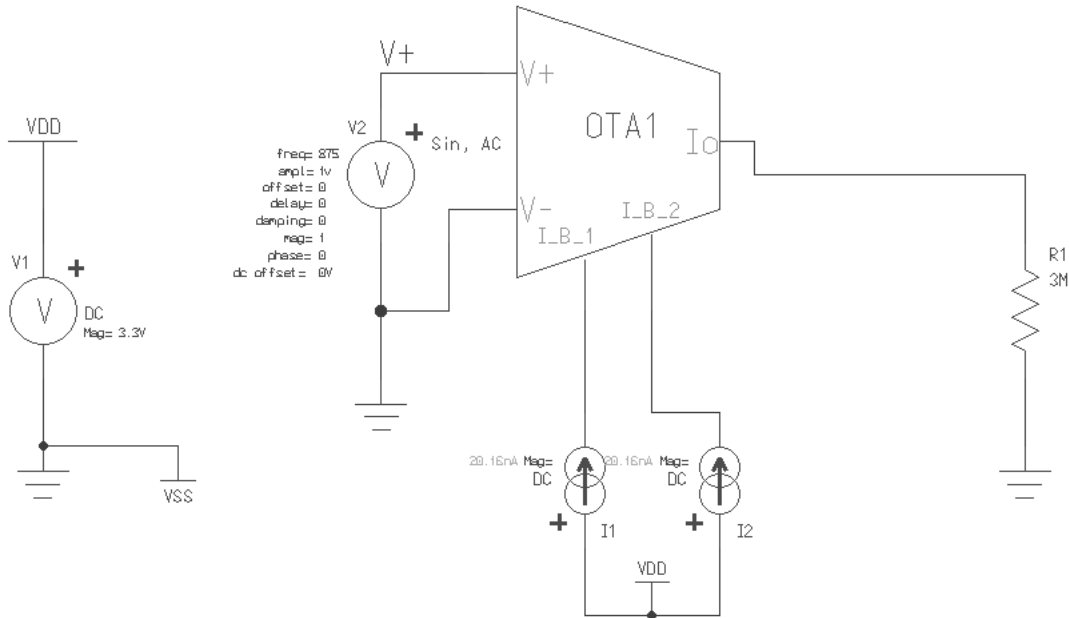


Figura 5.11: Prueba de transitorio para los OTAs.

5.4.2. Modificaciones a la herramienta de optimización

La primera etapa de las modificaciones se realizaron en el programa que actuaba como interfaz (simInterface) entre los archivos de simulación y el programa que ejecutaba el algoritmo. Esta etapa representó el verdadero enlace entre la herramienta y el ICstudio. Las tareas realizadas a este programa fueron:

1. *Definición de variables acorde a los parámetros:* esta modificación consistió en la creación de un arreglo en el cual se almacenarían los valores para cada parámetro que asigne el optimizador.
2. *Funciones de lectura y escritura de parámetros:* una de las tareas recurrentes del programa era la lectura de resultados de simulación y la escritura de los parámetros de manera que esto pudieran ser interpretados por ELDO. Los métodos de lectura se encargaron de leer la base de datos generada por ELDO para obtener la información de corriente y tensión y así poder procesar dichos datos. En el registro de simulación debía hacerse la lectura del valor de capacitancia de entrada parásita, del cual se tomaba el condensador equivalente que se formaba entre el nodo V^+ y tierra.
La función de escritura operaba cada vez que el optimizador enviaba un conjunto nuevo de parámetros insertándolos en el netlist para una siguiente simulación.
3. *Cálculo de las especificaciones del OTA:* con la información extraída, se calcularon tres funciones que a la postre representarían valores de aptitud con los que operaría el optimizador. Utilizando el arreglo de valores de corriente, se obtuvo un arreglo con los valores de

transconductancia, al aplicar la definición de derivada discreta central o numérica que muestra la ecuación 5.27 (ver [24]):

$$G_{m(i)} = di_{out(i)} = \frac{i_{out(i+1)} - i_{out(i-1)}}{2h} \quad (5.27)$$

en donde el subíndice i indica la posición en el arreglo de la corriente donde se calcula el valor de su pendiente, la cual se almacena en un arreglo asignado para la transconductancia. El parámetro h identifica la separación entre las muestras, que en el DA-IC se especificó una resolución del barrido de DC de $1mV$, por lo que $h = 1mV$.

Para el rango lineal, determinado a partir de la transconductancia máxima como se explica en la sección 1.3, se aplicó el diagrama de flujo de la Figura 5.12. En esta labor se analiza que se determine solo un valor máximo y que los límites sean definidos conforme a este valor, así como que dentro de este rango se cumpla la limitación del 5% por debajo del valor máximo. El último cálculo consistió en la obtención del slew rate, para lo cual se implementó la ecuación A.4 del Apéndice A.

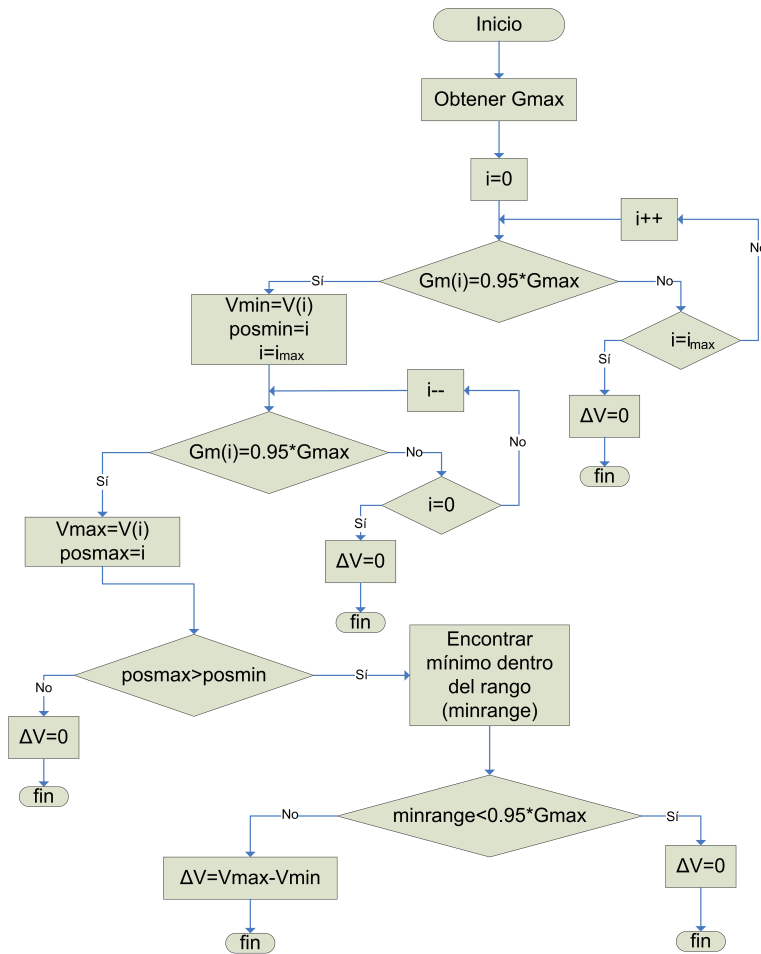


Figura 5.12: Diagrama de flujo para el método del cálculo del rango lineal

4. *Generación de los valores de aptitud (fitness values)*: en la sección 2.9 se explicó que las funciones de aptitud individuales, independiente de su forma, tenían que ser crecientes con respecto a alguno de los parámetros del circuito. En el caso presentado, estas funciones representarían cuatro especificaciones del OTA: su transconductancia, la capacitancia parásita de entrada, el slew rate y el rango lineal. En el caso del slew rate y el rango lineal, se deseaba que estos fueran crecientes conforme su parametrización u creciera; mientras que para la transconductancia y la capacitancia parásita debían ser inversamente proporcionales al crecimiento de su función asociada (se buscaba minimizar estos valores). De esta manera los cuatro valores de aptitud que son enviados al algoritmo genético se muestran en la Tabla 5.5. C_{in} denota la magnitud de la capacitancia parásita de entrada.

Tabla 5.5: Valores de aptitud asociados al OTA.

Especificación	Valor de aptitud
Rango lineal	ΔV
Slew rate	SR
G_m	$\frac{1}{G_m}$
Capacitancia parásita	$\frac{1}{C_{in}}$

La segunda etapa consistió en la modificación de clases y constructores propios del algoritmo genético, para así adaptarlos a los parámetros específicos para este circuito. Mediante el algoritmo PESA, implementado a partir de la librería LTI-LIB (ver [4]), se obtenía el Frente de Pareto y las parametrizaciones no dominadas era almacenadas en un archivo texto con aquellas combinaciones de parámetros que representaban las mejores opciones del conjunto total de datos.

5.5. Pruebas de preselección de resultados

Luego de la adaptación completa de la herramienta, se procedió a la ejecución de simulaciones con el fin de obtener los mejores casos posibles, para así configurar el OTA con estos datos y generar una simulación individual para verificar mediante las gráficas correspondientes el comportamiento del mismo. El Apéndice B contiene un historial con las simulaciones realizadas hasta llegar al OTA finalmente seleccionado. El siguiente código ejemplifica el contenido de cada archivo de simulación y cómo los datos eran leídos para su posterior análisis.

```
Internal evaluation 7/10
Valores de L1, L2: 5.82258 6.54839
Valores de L3, L5: 7.27419 1.95161
Valores de W1, W2: 0.5 0.5
Valores de W3, W5: 2.19355 3.40323
Valor de Ib: 23.3548
We received this response from the server: "Rango_lineal: 1.131 1/Gm: 5.97015e+07
1/C: 1.82298e+13 Slew_rate: 17.0678"
```

En la información anterior, una vez que se iniciaba la simulación, la etapa del algoritmo genético enviaba por medio de un socket todos los parámetros al simInterface para que este ejecutara la

simulación con dichos parámetros. Una vez finalizada, enviaba la respuesta de nuevo al algoritmo para que desplegara la información como la mostrada anteriormente. Para cada iteración el algoritmo ejecutaba diez evaluaciones y cada una de estas solicitaba una simulación a ELDO.

Para la preselección se utilizó un criterio de discriminación, basado en el cumplimiento o no de los requerimientos del OTA. Primeramente se seleccionaban los casos cuyo rango lineal superara 1V y de estos se tomaban los casos con los valores más bajos de I_B posibles, dado que estos casos presentarían el menor consumo. Si aún así existieran casos similares, se comparaban los valores recíprocos de la capacitancia parásita y la transconductancia en procura del valor más alto, aunque este criterio rara vez se utilizó pues por lo general cada simulación arrojaba un dato con condiciones más favorables que el resto. La Tabla 5.6 muestra una de las preselecciones hechas con datos válidos, una vez que la herramienta se depuró en su totalidad.

Tabla 5.6: Preselección de los mejores resultados dados por la herramienta de optimización.

G_m (nS)	ΔV (mV)	C_{in} (fF)	SR (mV/ μ s)	I_B (nA)
33.55	± 949	235.58	1.789	87.226
40.90	± 581	86.35	1.091	64.871
44.80	± 624	93.40	0.800	52.097
28.90	± 549	94.63	0.936	39.323
18.0	± 504	85.74	1.015	26.548
16.55	± 506	88.64	0.838	20.161
16.35	± 502	89.63	0.848	20.161

La Figura 5.13 muestra el Frente de Pareto correspondiente a la penúltima simulación efectuada (ver Apéndice B). En esta figura se emplean tres de los cuatro valores de aptitud asociados al OTA para poder obtener una gráfica en tres dimensiones.

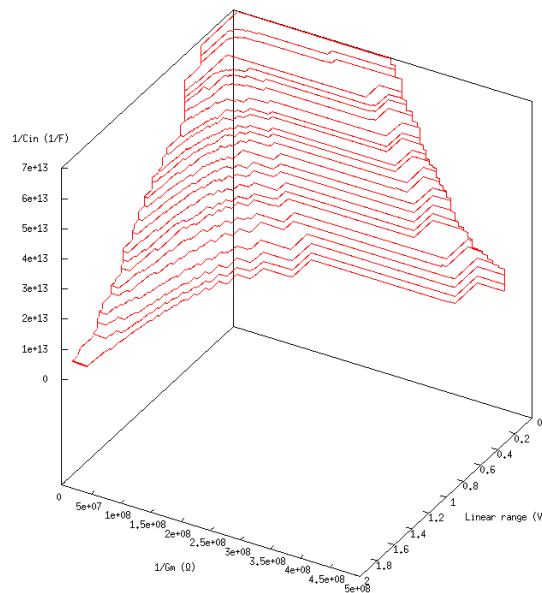


Figura 5.13: Frente de Pareto obtenido para una de las simulaciones realizadas.

5.6. Resultados seleccionados y escalamiento del OTA

Una vez seleccionado el OTA de 16.35nS como el OTA que cumplía las especificaciones dadas de rango lineal y slew rate, empleando un menor consumo y teniendo una capacitancia parásita y una transconductancia más baja que los otros casos, se procedió a la simulación individual de este resultado para obtener las gráficas de corriente y transconductancia. También se reescaló este OTA al generar dos circuitos adicionales con factores de $m=2$ y $m=1$ para obtener transconductancias de 32nS y 64nS, respectivamente. La Tabla 5.7 muestra las dimensiones de los transistores unitarios empleados. Se asignó $I_B = 20,16nA$ para todos los casos, tal y como se aprecia en la Tabla 5.6.

Tabla 5.7: Dimensiones de los transistores unitarios empleados en los OTAs

Transistor	L (μm)	W (μm)
M_1	7.2	0.8
M_2	8	0.8
M_3	2	5
M_5	2	5

Las Figuras 5.14-5.18 muestran los transistores construidos en el DA-IC, correspondientes a M_1 , M_2 , M_3 y M_5 . Las imágenes fueron tomadas a impresiones de pantalla de la interfaz del Design Architecture.

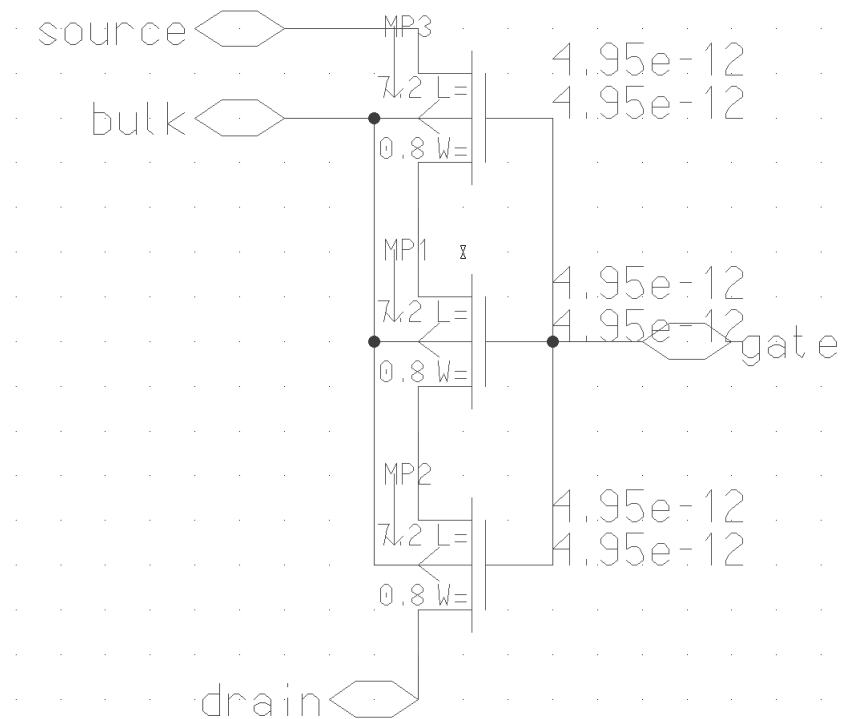


Figura 5.14: Módulo del transistor M_1 .

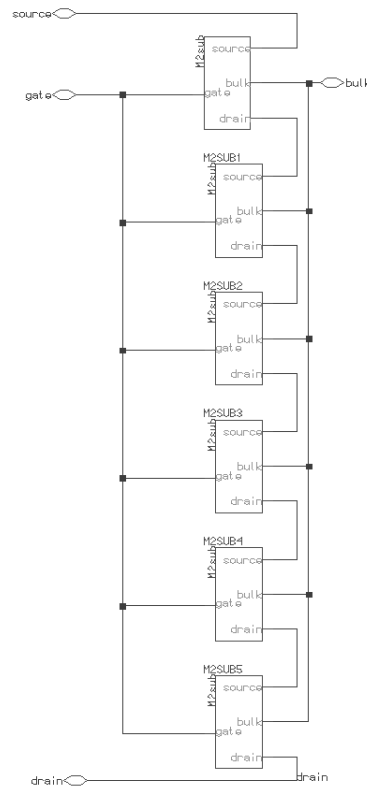


Figura 5.15: Módulo del transistor M_2 .

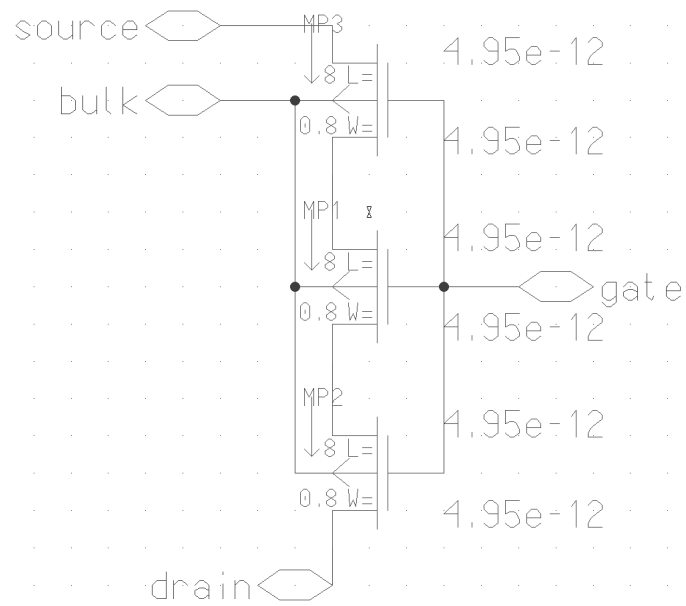


Figura 5.16: Sub-módulo del transistor M_2 .

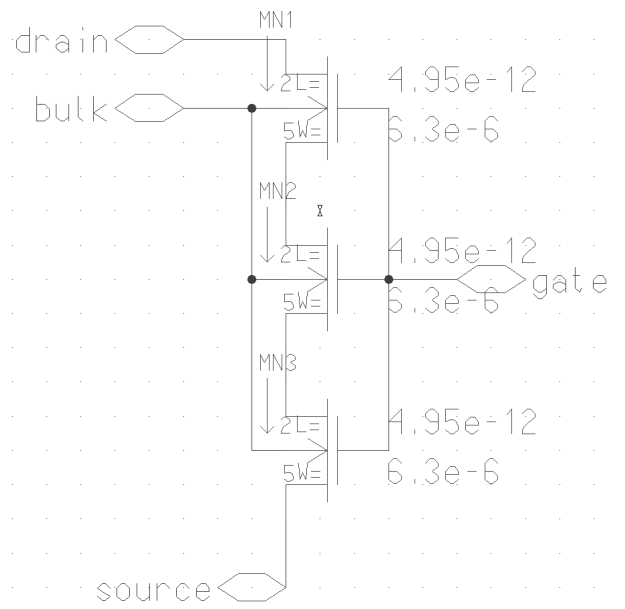


Figura 5.17: Módulo del transistor M_3 .

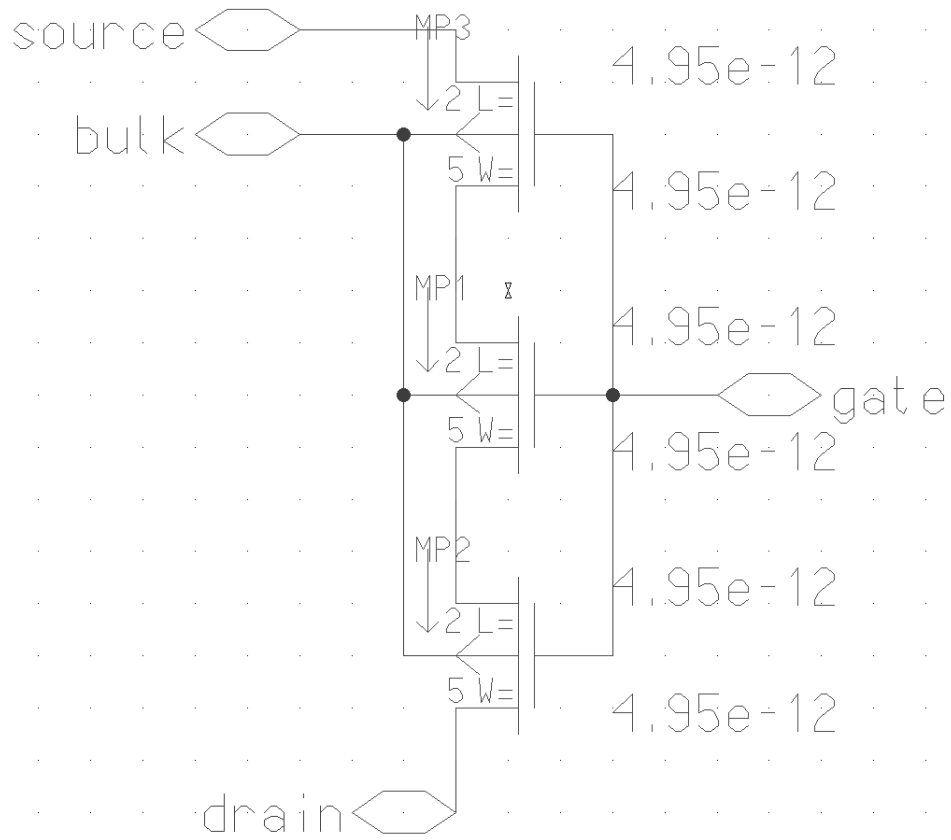


Figura 5.18: Módulo del transistor M_5 .

Las Figuras 5.19, 5.20 y 5.21 representan los esquemáticos completos para los tres OTAs requeridos.

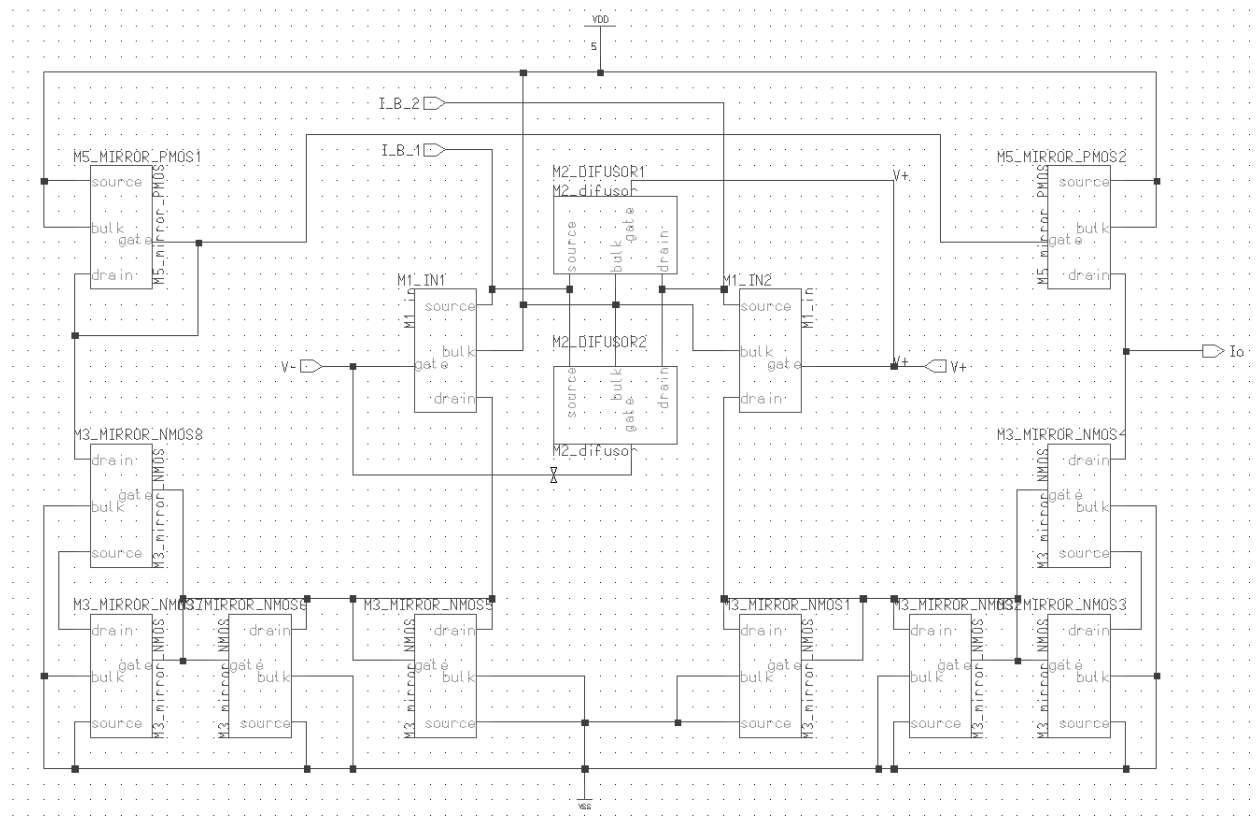


Figura 5.19: Esquemático del OTA de 16nS.

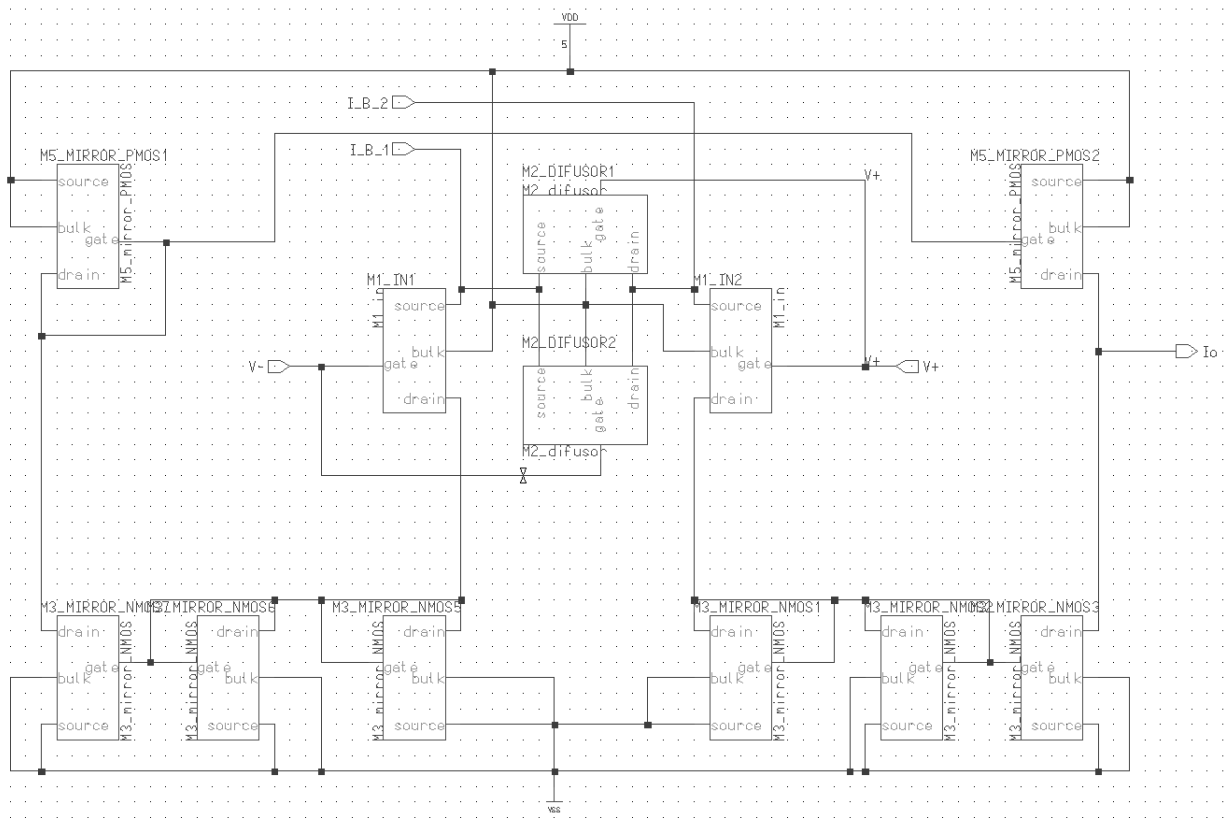


Figura 5.20: Esquemático del OTA de 32nS.

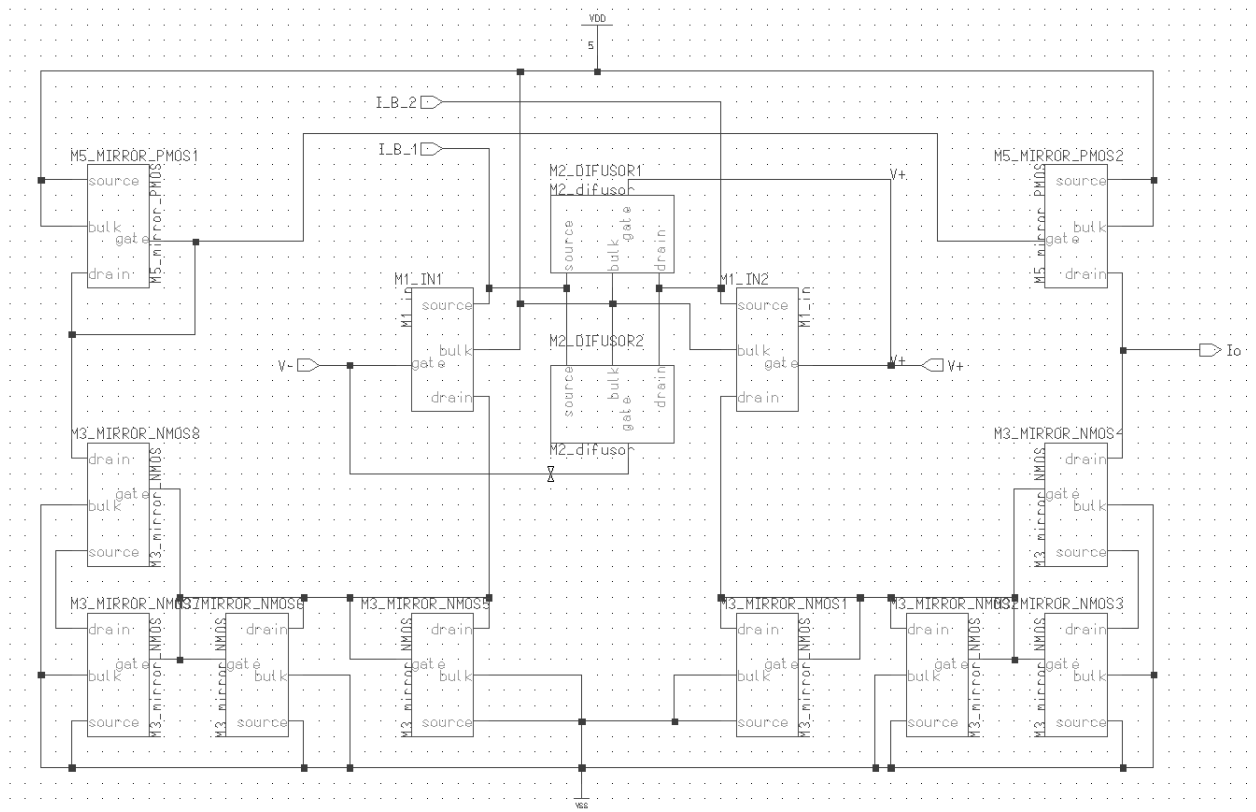
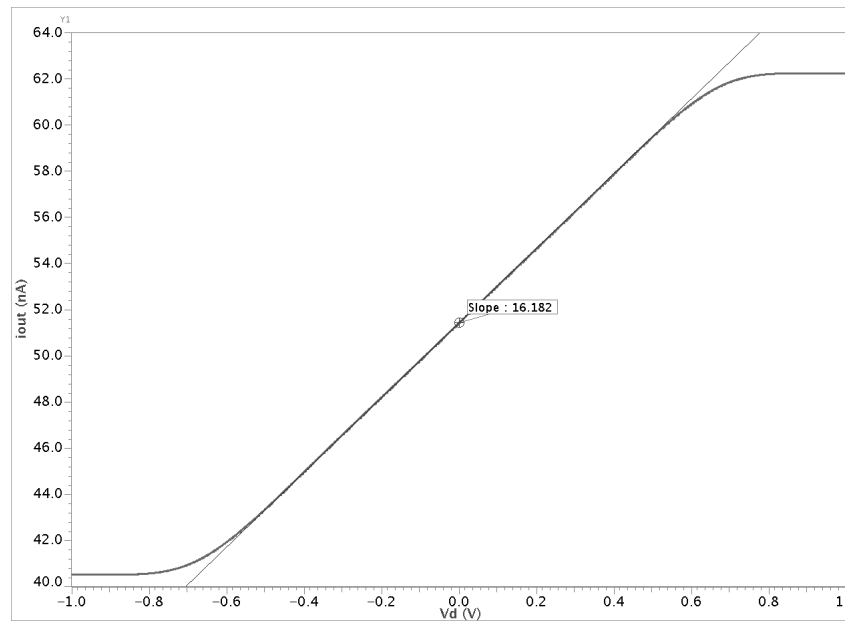
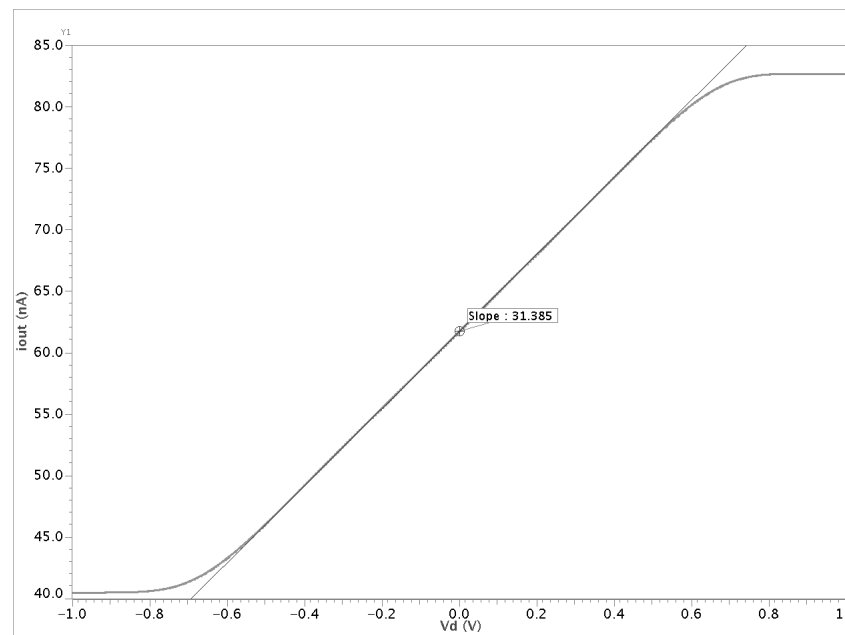


Figura 5.21: Esquemático del OTA de 64nS.

La diferencia primordial entre los OTAs radica en la cantidad de transistores del espejo de corriente inferior para modificar el valor del factor de escalamiento. En las secciones siguientes se muestran las gráficas de corriente, transconductancia y transitorio en AC para estos casos.

5.7. Gráficas de corriente, transconductancia y transitorios

Empleando el banco de prueba de la Figura 5.10, se obtuvieron las Figuras 5.22, 5.23 y 5.24, que representan la respuesta de corriente de salida en función de la tensión de entrada diferencial, para un rango de $\pm 1V$. Estas gráficas se obtuvieron mediante impresiones de pantalla a la interfaz del EZwave.

Figura 5.22: i_{out} vrs V_d para el OTA de 16nS.Figura 5.23: i_{out} vrs V_d para el OTA de 32nS.

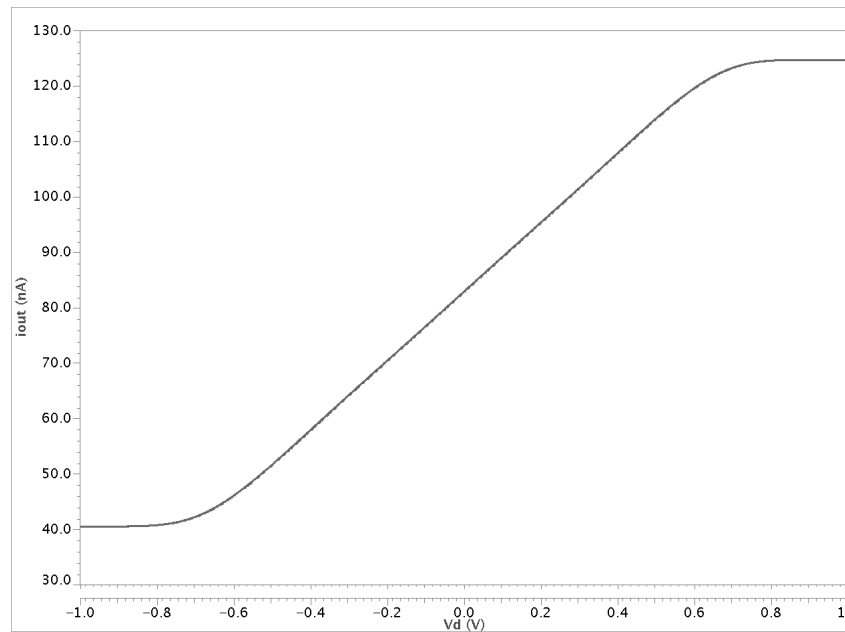


Figura 5.24: i_{out} vrs V_d para el OTA de 64nS.

Empleando las herramientas matemáticas del EZwave, se calcularon las derivadas a la curvas de corriente, con el fin de obtener la transconductancia de los OTAs en función de la tensión de entrada diferencial. Los resultados obtenidos se muestran en las Figuras 5.25, 5.26 y 5.27.

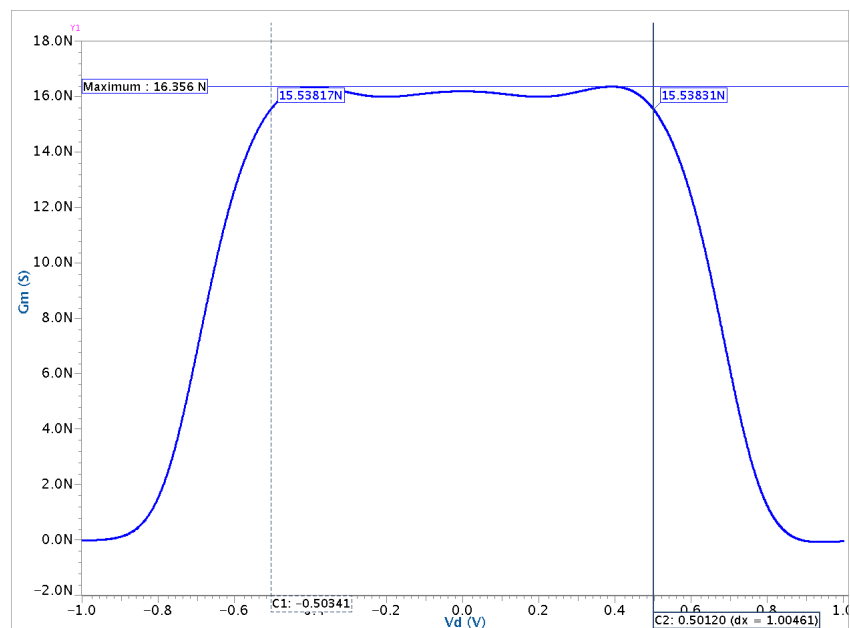


Figura 5.25: G_m vrs V_d para el OTA de 16nS.

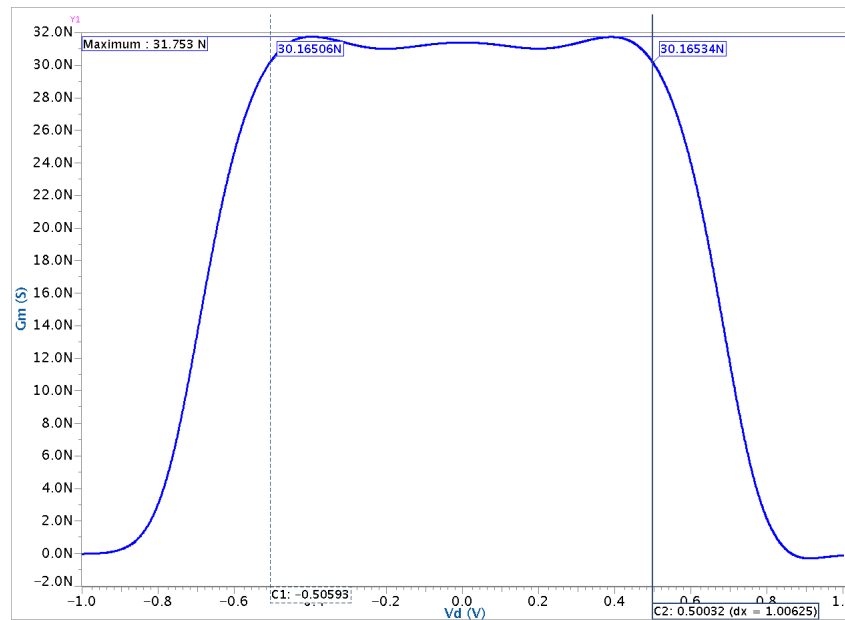


Figura 5.26: G_m vrs V_d para el OTA de 32nS.

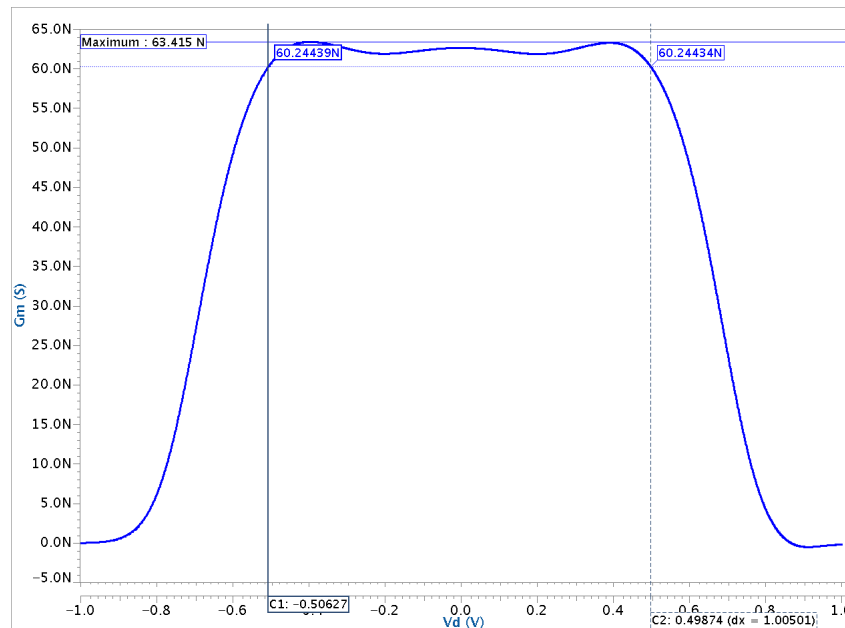
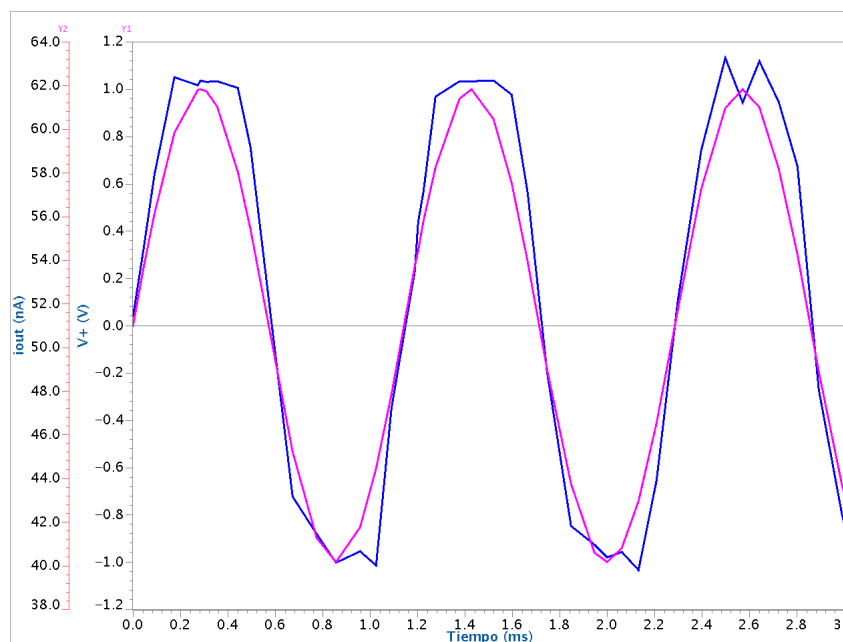
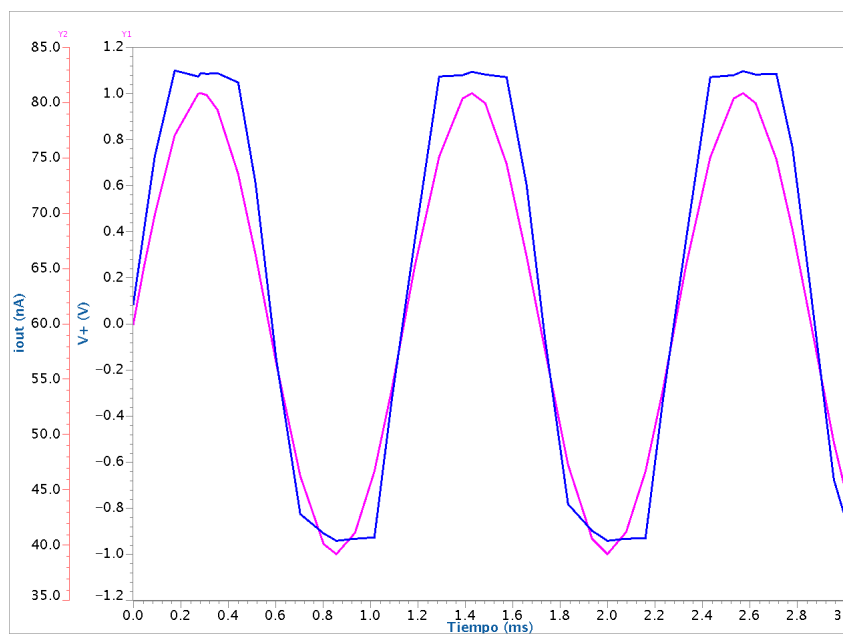


Figura 5.27: G_m vrs V_d para el OTA de 64nS.

Utilizando el banco de pruebas en AC de la Figura 5.11, se obtuvieron las gráficas de las Figuras 5.28, 5.29 y 5.30, en las que el trazo en color azul representan la forma de onda de la corriente de salida, mientras que el trazo en color magenta representa la tensión de entrada aplicada.

Figura 5.28: Gráficas de i_{out} y V_d en función del tiempo para el OTA de 16nS.Figura 5.29: Gráficas de i_{out} y V_d en función del tiempo para el OTA de 32nS.

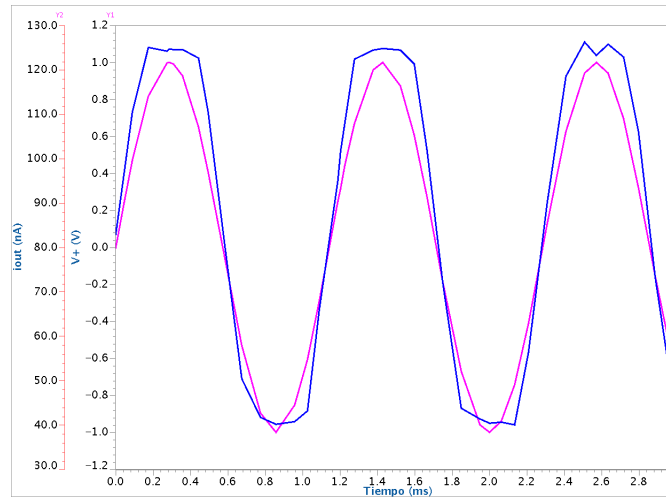


Figura 5.30: Gráficas de i_{out} y V_d en función del tiempo para el OTA de 64nS.

La Tabla 5.8 resume las principales características de los OTAs generados, mientras que en la Tabla se comparan los circuitos obtenidos con las especificaciones dadas inicialmente (ver sección 1.3) y el porcentaje de mejoramiento con respecto a estas.

Tabla 5.8: Resumen de las características de los OTAs diseñados

OTA	$G_{m_{max}}$ (nS)	ΔV (mV)	C_{in} (fF)	Potencia (nW)	SR (mV/ μ s)
16nS	16.356	± 502	89.625	150.08	0.848
32ns	31.753	± 502	89.625	217.49	1.743
64nS	63.415	± 503	89.625	356.11	3.496

Tabla 5.9: Comparación entre las especificaciones de los OTAs obtenidos y las especificaciones iniciales.

Especificación		Inicial	Lograda	Porcentaje de mejoramiento (%)
Slew rate (mV/ μ s)	@219Hz	0.69	0.848	22.9
	@437Hz	1.37	1.743	27.2
	@875Hz	2.75	3.496	27.1
Rango lineal (mV)		± 500	± 502	0.4
Capacitancia parásita de entrada (fF)		200	89.625	55.2
Transconductancia (nS)		137	63.415	53.7
Potencia (nW)		1788	356.11	80.0

5.8. Layouts trazados

Las Figuras 5.31, 5.32 y 5.33 muestran el layout generado para cada OTA. En estos layouts, la columna de transistores de la izquierda contienen los transistores del par diferencial y los del difusor simétricos debidamente pareados (ver [23]).

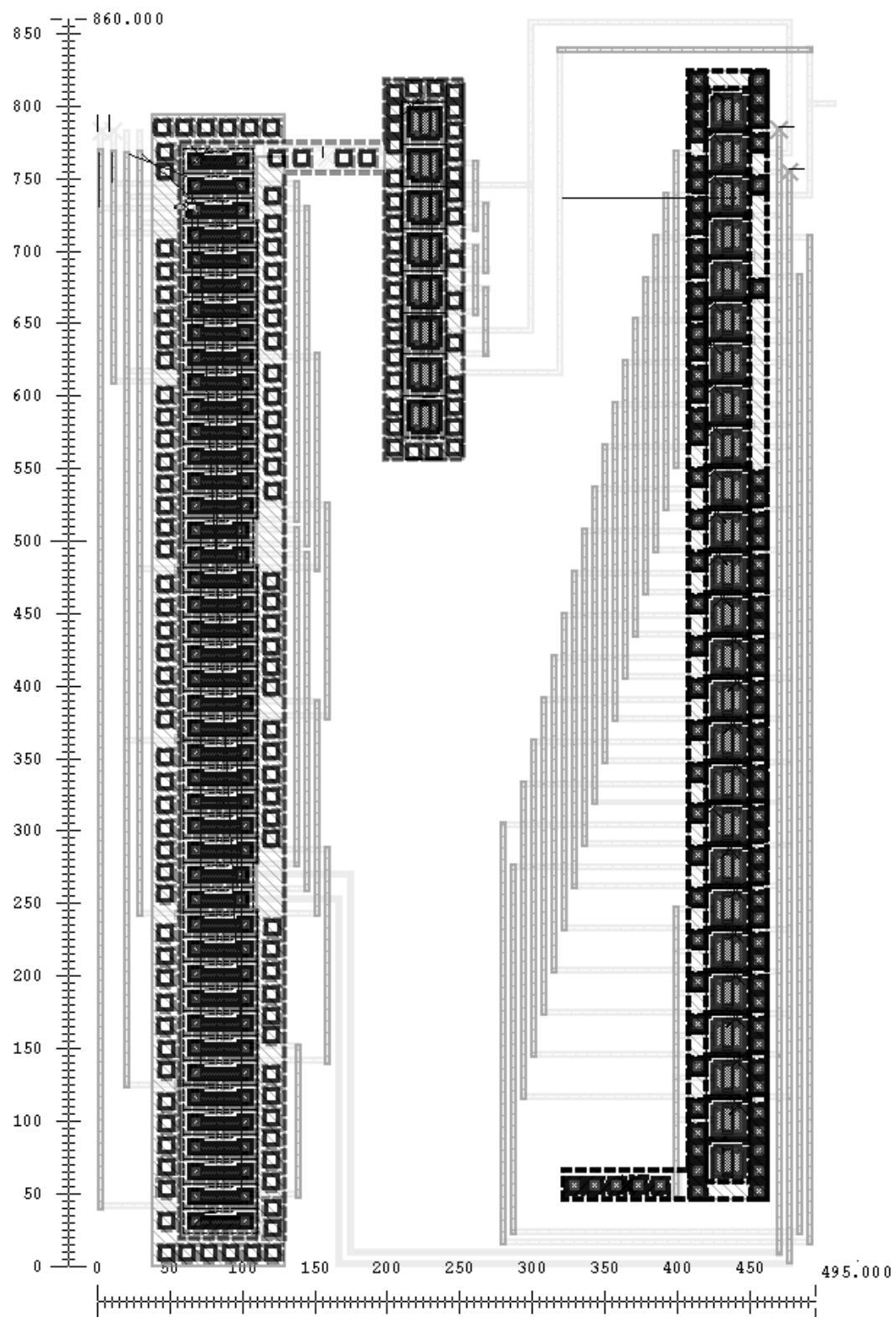


Figura 5.31: Layout del OTA de 16nS.

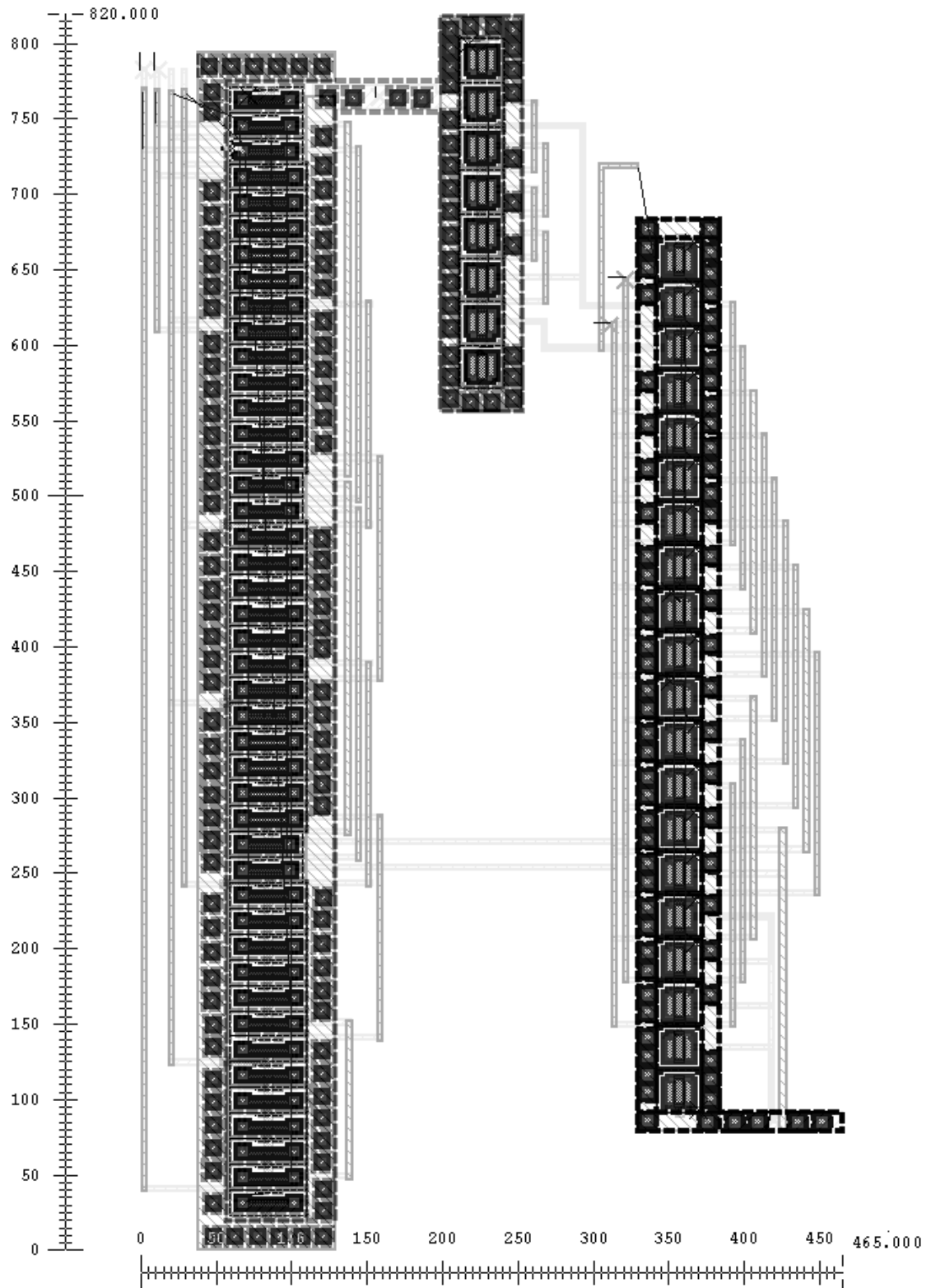


Figura 5.32: Layout del OTA de 32nS.

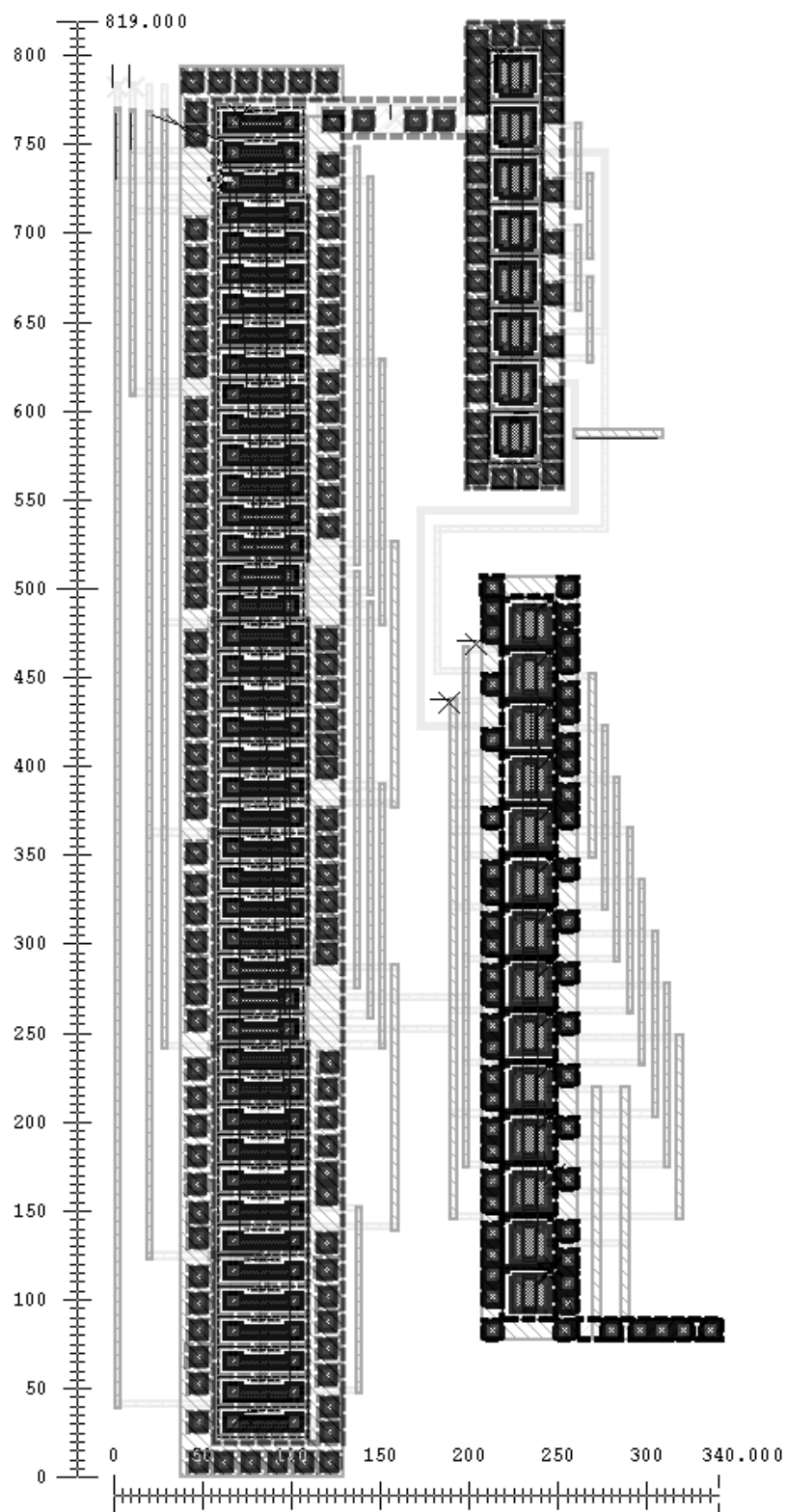


Figura 5.33: Layout del OTA de 64nS.

Las medidas que muestran los ejes de estas figuras corresponden a las dimensiones escalables del proceso de fabricación, denotadas por λ . Para la tecnología empleada, $\lambda = 0,3\mu m$. La Tabla 5.10 contiene las dimensiones de los layouts en micrómetros y el área que requieren.

Tabla 5.10: Dimensiones físicas de los layouts generados

OTA	Dimensiones (μm)	Área requerida (nm^2)
16nS	148.5x258	38.31
32nS	139.5x246	34.32
64nS	102x245.7	25.06

Empleando el circuito de pruebas en AC (Figura 5.11), se obtuvieron las Figuras 5.34, 5.35 y 5.36, que muestran la simulación post-layout incluyendo capacitancias parásitas.

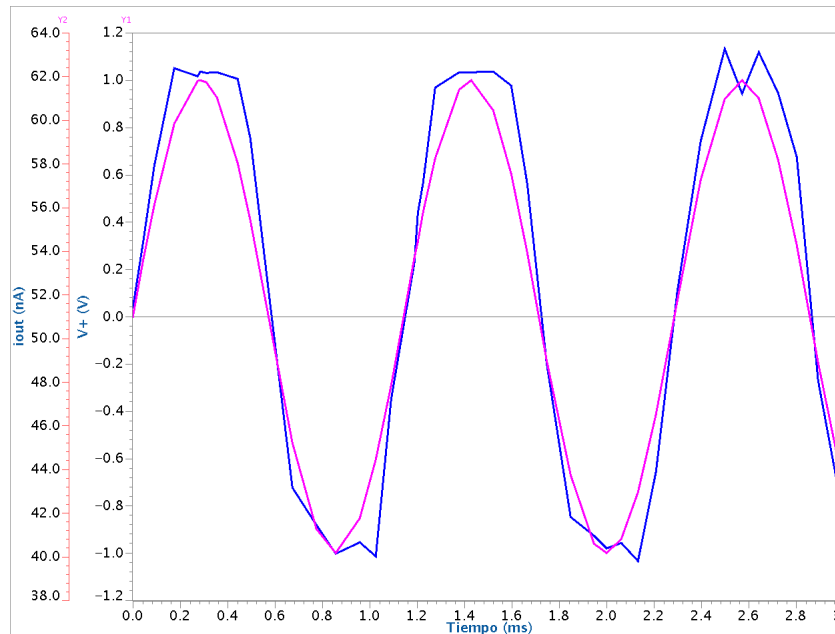


Figura 5.34: Gráficas de i_{out} y V_d en función del tiempo para el layout del OTA de 16nS.

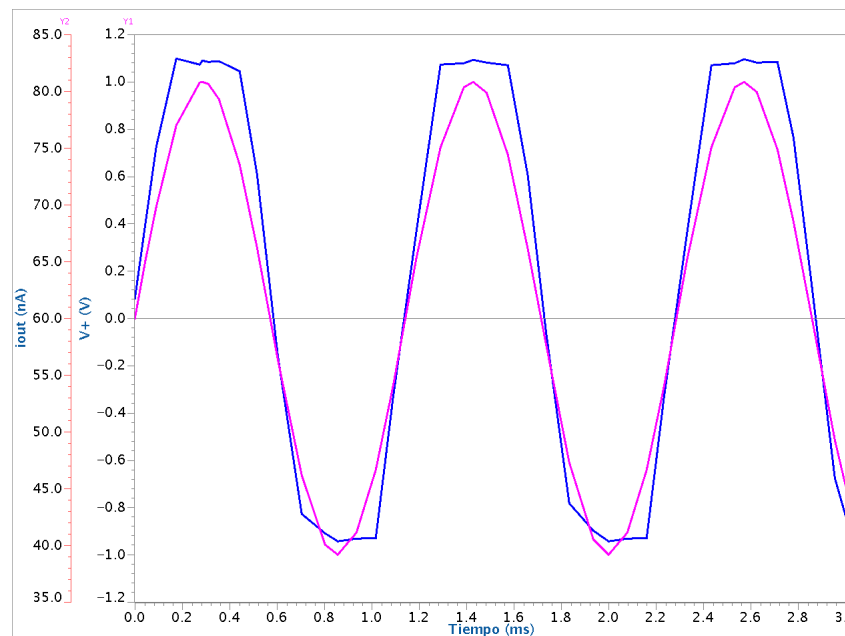


Figura 5.35: Gráficas de i_{out} y V_d en función del tiempo para el layout del OTA de 32nS.

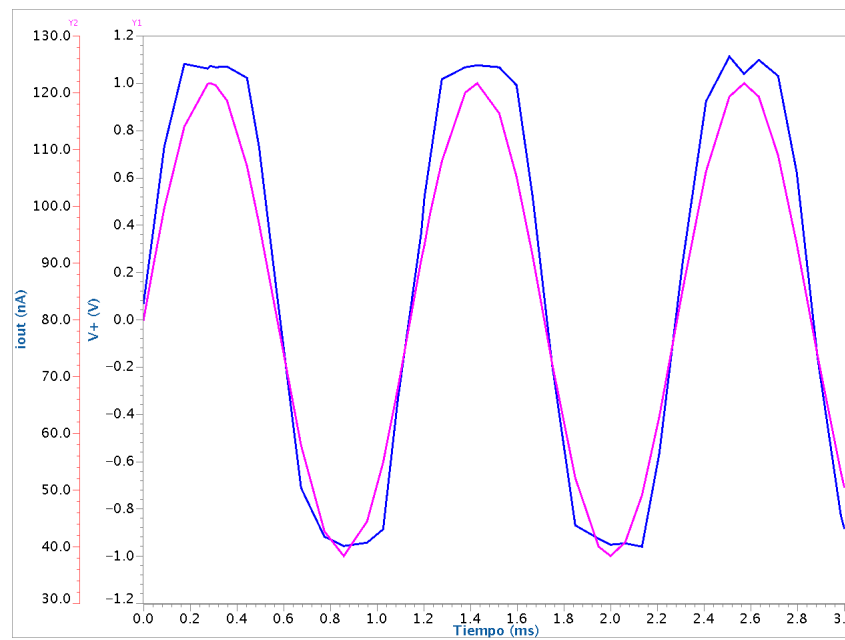


Figura 5.36: Gráficas de i_{out} y V_d en función del tiempo para el layout del OTA de 64nS.

Capítulo 6

Análisis de resultados

En la primera etapa de la solución se realizó un estudio de cuatro de las estructuras de OTAs planteadas en la sección 2.5, con el fin de determinar cuál de ellas representaba la mejor opción. Para el diseño de las mismas se utilizaron las ecuaciones de esa misma sección, aunque los cálculos teóricos no concordaron con la respuesta obtenida en el simulador, lo cual se aprecia en la Tabla 5.1, en donde los casos del transistor en modo común y el OTA de Nedungadi presentaron un error muy grande, por encima del 70 %, lo que resta confiabilidad a las ecuaciones de estos dos esquemas. Para el caso del OTA simple y el que emplea el difusor simétrico, se obtuvieron porcentajes de error menores pero siempre rondando cerca del 20 %, lo que implica que los circuitos requieran de ciertos ajustes luego de los cálculos para obtener el comportamiento deseado. El problema radica en que debido a la interrelación existente entre todos los parámetros, estos ajustes impactan en las demás características y no se puede garantizar que esto suceda de una manera positiva.

Estos errores pueden atribuirse a diversas causas. La primera compete a la zona de operación efectiva en la que trabaja cada transistor. Para poder emplear las ecuaciones específicas para cada esquema (ver sección 2.5), debe garantizarse que estos trabajen en inversión fuerte. Esto genera una limitante, dado que no hay forma de evaluar el comportamiento preciso en todas las zonas, sino solamente mediante las aproximaciones de cada región dadas en la sección 2.2. El uso de la corriente de 50nA se presentó primeramente para buscar un bajo consumo y también debido a que en el rango de corrientes de nanoAmperios fue donde se presentaron las características más similares a las estipuladas por las ecuaciones, como lo fue la respuesta cuadrática del par diferencial simple (ecuación 2.44).

Otra razón se deriva de las diferencias entre los modelos que se utilizan para las ecuaciones y los que utiliza el simulador. Por una parte, el simulador LTSPICE IV utiliza el modelo BSIM3v3 cuando se emplean los parámetros de la tecnología de fabricación empleada (AMI05). Estos modelos computacionales tienden a evaluar más efectos que los que se podrían con el cálculo a mano, por lo que parte del error obtenido ha de atribuirse a este factor. Las ecuaciones empleadas utilizan modelos como el EKV o el Shichman-Hodges (ver sección 2.8) que permiten una aproximación al diseñador en cuanto a los resultados a obtener, incluso con mejor precisión que el BSIM en ciertas circunstancias. Sin embargo, al no ser derivadas para todas las zonas de operación, es posible la obtención de resultados de baja precisión (especialmente si se aleja de la zona de operación para la cual fueron estipuladas), con porcentajes de error como los que se obtuvieron.

Luego de algunos ajustes a estos esquemas, se obtuvo la simulación de la Figura 5.6. Los datos de la Tabla 5.2 muestran que el OTA con difusor simétrico presentó el mejor rango lineal de los cuatro casos y la menor transconductancia, mientras que el circuito de Nedugadi presentó un consumo 33 % menor al de los otros casos. Esto se debe a que en el caso de Nedungadi no se cumple la ecuación de potencia máxima dada en la ecuación 5.26 puesto que los transistores de tensión flotante proveen un ruta a tierra para la corriente de polarización en caso de que uno de los transistores tenga polarización inversa y no conduzca. En los otros tres esquemas, la ruta alterna se provee a través del transistor opuesto del par diferencial, por lo que los espejos copiarían en el caso extremo el doble de corriente de polarización. Si bien bajo este supuesto el esquema de Nedungadi pareció resultar una buena opción en términos de potencia, el factor de $(n + 1)I_B$ para lograr polarizar a todos los transistores en la misma región eventualmente podría requerir de un incremento inesperado en el consumo de potencia para lograr mantener un rango lineal alto, lo que sugirió ser una opción riesgosa en caso de no resultar favorable. Por otra parte, el esquema del difusor simétrico ha probado ser una configuración con un rango lineal muy amplio (ver [1, 14]) y cuyo consumo pueda controlarse de manera que para obtener las especificaciones deseadas puedan manipularse mejor las dimensiones de su componentes.

El siguiente paso consistió en diseñar un primer OTA con dimensiones definidas mediante parámetros para lograr acoplar el diseño con la herramienta. Previo a este diseño, se realizó un análisis del OTA con difusor simétrico, empleando las ecuaciones del modelo general y los circuitos equivalentes del modelo en pequeña señal (secciones 2.2 y 2.3). La ecuación 2.47 mostró ser un cálculo con un error cercano al 20 %, pero a partir del modelo general se esperó obtener una ecuación para la transconductancia con mayor precisión que esta. El modelo obtenido se presenta en la ecuación 5.20 y si se compara con el modelo de inversión fuerte de la ecuación 2.47, puede apreciarse la similitud entre ambos mediante los valores de k_1 con g_m y k_2 con g_{ds} , respectivamente.

Basado en lo anterior se obtuvo un OTA cuyas características se resumen en la Tabla 5.4. Para este OTA se logró reducir el consumo de potencia a 174nW, pero a su vez sacrificando el rango lineal del mismo, como un efecto que muestra la ecuación 5.24. El slew rate tampoco llega al requerimiento mínimo de 0.69mV/ μ s. Asimismo, debe notarse el ajuste hecho en la corriente para lograr una transconductancia de 36nS, al pasar de 56nA calculados teóricamente a 26nA requeridos en el esquemático para lograr dicha transconductancia, lo que muestra a su vez que el modelo general obtenido genera una mejor aproximación pero que también carece de la precisión de los modelos computacionales por ya razones ya expuestas. El caso de este OTA inicial buscó establecer un marco de referencia para medir cuanto se pueden llegar a mejorar estos parámetros.

En este sentido, la herramienta de optimización representó un mecanismo efectivo en la búsqueda de mejores parámetros al otorgar resultados con la garantía de que estos son los más óptimos dentro de las posibilidades dadas por el usuario. Estos límites impuestos consistieron en que para las dimensiones de los transistores, la herramienta evaluara opciones entre el tamaño mínimo para la tecnología y un máximo de 10 veces ese valor mínimo. Para la corriente de polarización, se dio un rango desde 1nA hasta 100nA y como se mencionó anteriormente, la herramienta buscó maximizar el rango lineal y el slew rate, mientras minimizaba la transconductancia y la capacitancia parásita. Estos valores de aptitud fueron dados de acuerdo a las especificaciones de diseño mencionadas en la sección 1.3 e incluidas en los objetivos del proyecto. La potencia no fue incluida dentro de estos valores porque implícitamente esta sería optimizada mediante el compromiso de estos cuatro valores entre sí mediante la corriente de polarización.

Dada la enorme cantidad de datos que arrojaba cada simulación, la Tabla 5.6 contiene una muestra de datos que en ciertos momentos de las simulaciones se consideraron como las mejores opciones. Entre estos datos hay casos en los que se alcanzaban rangos lineales de prácticamente 2V, pero con un alto consumo de potencia y una alta capacitancia de entrada, así como casos con bajas transconductancias con rangos lineales dentro de las especificaciones buscadas. El Frente de Pareto mostrado en la Figura 5.13 ejemplifica el tipo de compromisos que existe entre los valores de aptitud empleados: para lograr maximizar uno de ellos, los demás deberán sufrir algún costo.

El último caso de la Tabla 5.6 (16.35nS) fue seleccionado como el OTA a implementar pues se ajustaba a todos los requerimientos mínimos establecidos y presentó el menor consumo de potencia que se podía lograr mediante esta herramienta y los límites dados anteriormente. La diferencia leve con el caso de 16.55nS se da por el redondeo que debe hacerse a las dimensiones previo a elaboración del layout. El entorno del ICStation hace estos ajustes de manera automática, pero este redondeo se aplicó de manera manual en el esquemático para visualizar el resultado preliminar.

Las dimensiones empleadas en los transistores unitarios se listan en la Tabla 5.7. Tanto en este como en todos los casos se notó el hecho de que los transistores tienden a ser más largos que anchos con el fin de reducir su transconductancia intrínseca. Esto se deriva a partir de que en la mayoría de ecuaciones de diseño se repite constantemente el factor $(W/L)\mu C_{ox}$, en donde claramente se aprecia que un aumento en el largo de los transistores ocasiona una reducción en dicho factor y por ende, en la transconductancia obtenida. Los tres transistores M_3 y M_5 fueron manualmente modificados con las dimensiones de la Tabla 5.7 por recomendación del profesor Chacón de utilizar transistores más cuadrados. Los esquemáticos de los transistores implementados se muestran en las Figuras 5.14-5.18. El transistor M_2 (Figura 5.15) se compuso de seis submódulos (Figura 5.16), por la relación de tamaño con M_1 . Al evaluar la relación de tamaño entre los transistores completos del par diferencial y el difusor simétrico se obtiene que $\frac{w_1/L_1}{w_2/L_2} = 6,7$, en donde se alcanza la relación óptima de Krummenacher sin que el optimizador conociera la existencia de la misma, demostrando así la capacidad del algoritmo de encontrar los mejores conjuntos de parámetros.

El circuito seleccionado se reescaló para contar con los tres OTAs solicitados, cuyos esquemáticos se muestran en las Figuras 5.19, 5.20 y 5.21, para luego generar las curvas de comportamiento de corriente y transconductancia (Figuras 5.22-5.27). Estas gráficas permiten visualizar primordialmente el rango lineal del dispositivo basado en la transconductancia máxima de cada curva. De manera representativa cada OTA fue nombrado con una transconductancia de 16nS, 32nS y 64nS, respectivamente.

La prueba de transitorios produjo los resultados de las Figuras 5.28, 5.29 y 5.30. En estas se aprecia la deformación de la respuesta de corriente en los valores máximos de la tensión de entrada, lo que se debe a que a esos valores de tensión ya se presenta una caída en la transconductancia cercana al 5%, por la pérdida del comportamiento lineal en esa región de tensiones. Esto ocasiona una distorsión en la señal de salida reflejado en el aplanamiento o la aparición de cambios drásticos en la amplitud máxima de la corriente. En los tres OTAs la distorsión varía ligeramente, aunque siempre se da en la misma zona cercana al máximo de amplitud.

Los valores exactos para cada transconductancia junto con otras características extraídas del registro de simulación se presentan en la Tabla 5.8. Con estos datos puede verificarse que el escalamiento por medio de espejos de corriente no afecta ni el rango lineal ni la capacitancia

parásita de entrada, los cuales dependen del esquema de par diferencial empleado, sus dimensiones y su polarización. La potencia tiende a aumentar conforme se disminuya el factor de escalamiento (aumento en la transconductancia), debido a que los espejos reflejan menos corriente por las ramas que fluyen directamente de la fuente de alimentación.

En la Tabla 5.9 se efectuó una comparación entre las especificaciones que los OTAs requerían cumplir. En el caso del slew rate, se logró un valor entre 22 % y 27 % al que necesitaba el filtro, lo que implica que para las frecuencias de interés, el OTA será capaz de responder con la suficiente velocidad que las mismas lo exijan. En el caso del rango lineal se logró satisfacer el rango lineal de manera ajustada, mientras que en la capacitancia parásita de entrada se logró una reducción del 55 % con respecto al valor que se buscaba. Puesto que la transconductancia también se redujo un 53.7 % con respecto al valor inicial, el condensador necesario para el filtro va a experimentar esa misma reducción en el valor de su capacidad, pero aún sigue siendo lo suficientemente grande como para que la capacitancia parásita no afecte el corrimiento de los polos. En la potencia se logró una reducción de 80 % primordialmente por la disminución de la corriente de polarización de 45nA (usada por Chacón [1]) a una de 20.16nA. Con el consumo de potencia de los OTAs obtenidos (Tabla 5.8), es posible reducir el consumo del filtro completo de $21.46\mu\text{W}$ a $3.28\mu\text{W}$, para una reducción del 85 %. A pesar de que no se cumple la meta establecida originalmente, queda establecido un antecedente que permitiría una próxima mejora al OTA.

El último requerimiento consistió en la elaboración de los layouts de los tres OTAs. Estos se muestran en las Figuras 5.31, 5.32 y 5.33. En estos layouts, el conjunto de transistores de la izquierda corresponden al par diferencial junto con el difusor simétrico. Los transistores de este bloque se intercalaban entre sí para mejorar el *matching* entre ellos, al lograr que todos los transistores tuvieran los mismos elementos alrededor (ver las recomendaciones de [1, 23, 8]). El bloque superior derecho representa los transistores del espejo superior. Ambos bloques se conservaban idénticos en los tres OTAs, ya que el bloque inferior derecho es el que corresponde al espejo de corriente que implementa el escalamiento de la corriente. El tamaño de este bloque es el que determinaba finalmente el espacio a ocupar por el OTA, como lo muestran las dimensiones de la Tabla 5.10. En ellas se aprecia, que conforme se aumentaba el factor de escalamiento para disminuir la transconductancia, aumentaba el área del layout. En el caso del layout de mayor transconductancia obtenida (63.41 nS), su área total es de 38.31nm^2 , el cual sustituirá al OTA de 137nS cuya área original era de 990nm^2 . Este cambio representa un ahorro del 96 % en el espacio que ocupará el nuevo OTA.

Para verificar el funcionamiento de estos layouts, se realizó una simulación de transitorios con el propósito de identificar variaciones causadas por la inclusión de efectos parásitos en el netlist. No obstante las simulaciones de las Figuras 5.34, 5.35 y 5.36 no muestran diferencias notables con las simulaciones en AC de los esquemáticos (Figuras 5.28, 5.29 y 5.30). Esto muestra que la disposición de componentes se realizó de manera que no se generan efectos parásitos adicionales. La verificación real de esto se presentaría cuando los OTAs sean posteriormente agregados a la etapa de filtros del sistema de detección de disparos y el circuito integrado sea construido.

Capítulo 7

Conclusiones

1. El OTA con difusor simétrico representó la mejor opción de los esquemas evaluados en términos de linealidad y menor transconductancia.
2. Los modelos de diseño a mano poseen grandes índices de error (entre 18 % y 71 % dependiendo del modelo) con respecto a los modelos computacionales de simulación.
3. El factor de escalamiento obtenido mediante espejos de corriente no perjudica el rango lineal ni la capacitancia parásita de entrada.
4. Los transistores completos para cada sección del OTA presentan las siguientes dimensiones: $0,8\mu m/21,6\mu m$ para los transistores del par diferencial, $0,8\mu m/144\mu m$ para los transistores del difusor simétrico y $2\mu m/5\mu m$ para los transistores de los espejos de corriente.
5. Se obtuvo un modelo matemático de transconductancia para el OTA con difusor simétrico que permite el diseño en todas las regiones de operación en saturación.
6. La herramienta basada en algoritmos genéticos permitió optimizar el diseño de un OTA, otorgando las mejores parametrizaciones a partir de un conjunto de posibilidades definidas por el usuario. Los datos seleccionados cumplieron la relación óptima de Krummenacher.
7. Se diseñaron tres OTAs de 63.41nS, 31.75nS y 16.35nS mediante la técnica de escalamiento de corriente y cumpliendo todos los requerimientos de diseño otorgados.
8. Las transconductancias logradas representan una reducción en magnitud del 53 % con respecto a las transconductancias inicialmente planteadas, permitiendo una reducción de las capacitancias que necesitan los filtros en los que se emplearán estos OTAs.
9. El comportamiento de los OTAs presenta poca distorsión dentro de los límites cercanos de su rango lineal.
10. Los OTAs obtenidos presentan una capacitancia parásita de entrada de 68.625fF, lo que representa una reducción del 55 % por debajo de la especificación inicial.
11. Los OTAs obtenidos poseen un rango lineal de $\pm 502\text{mV}$, satisfaciendo el requerimiento de linealidad solicitado.
12. Se obtuvieron slew rates de 22 % a 27 % superiores a los requeridos inicialmente.

13. El consumo de potencia es inversamente proporcional al factor de escalamiento utilizado en un OTA.
14. Se logró una reducción del 80 % en el consumo de potencia, pasando de 1788nW a 356nW para el OTA que trabajará en el filtro cuya frecuencia de corte se localiza en 875Hz.
15. Se proyecta que el filtro completo consumirá $3.28\mu\text{W}$, lo que representa una disminución del 85 % del consumo original ($21\mu\text{W}$). No se alcanza la meta de lograr un consumo neto de $1\mu\text{W}$.
16. Los layouts de los OTAs poseen las siguientes dimensiones: $148.5 \times 258\mu\text{m}$, $139.5 \times 246\mu\text{m}$ y $102 \times 245.7\mu\text{m}$, respectivamente.
17. Se logró una reducción del 96 % en el área del OTA que trabajará en el filtro con frecuencia de corte de 875Hz, pasando de 990nm^2 a 38.21nm^2 .

Capítulo 8

Recomendaciones

A continuación se enumeran una serie de recomendaciones que eventualmente permitirían expandir o mejorar los logros alcanzados con este proyecto:

1. Difundir, en la medida de lo posible, el uso de la herramienta de optimización de circuitos por cuanto se ha comprobado su efectividad como complemento en las labores de ajuste y verificación de circuitos. El tiempo que se puede ahorrar permitiría el desarrollo de sistemas más complejos y diversos.
2. Flexibilizar los límites de la herramienta para que esta tenga más posibilidades de elección de resultados. Esto no ocasionaría un aumento en las evaluaciones, ya que el algoritmo busca concentrarse en las zonas más prometedoras del espacio de parametrizaciones.
3. Aumentar la resolución de la herramienta de optimización para que pueda explorar opciones más cercanas entre sí y con ello proveer una mejor parametrización.
4. Utilizar factores de escalamiento aún más grandes para reducir el consumo de potencia de los OTAs. En la realización de esto deberá velarse por el condensador necesario para el filtro (por la reducción de la transconductancia), ya que de ser muy pequeño podría equipararse con la capacitancia parásita y el funcionamiento del filtro se vería afectado también.
5. Investigar sobre la existencia de esquemas de OTAs más efectivos que los que se compararon en este proyecto. El área de investigación de los OTAs resultó ser bastante amplio debido a que no es un campo relativamente nuevo y existen muchas más fuentes relacionadas a esto. Eventualmente podría proponerse una nueva estructura basada en las experiencias mostradas en este documento y las de otros autores.

Bibliografía

- [1] A. Chacón, *Circuitos integrados de bajo consumo para detección y localización de disparos de armas de fuego*. PhD thesis, Facultad de Ingeniería, Departamento de Ingeniería Electrónica, Universidad Mar del Plata, Mar del Plata, Argentina, May 2009.
- [2] C. Dualibe, “Notas de clase: Diseño de filtros gmc (presentación power point),” *2da Escuela Argentina de Microelectrónica*, Julio 2007.
- [3] Mentor Graphics Corporation, “Mentor graphics homepage.” <http://www.mentor.com>.
- [4] R. Pereira-Arroyo, P. Alvarado-Moya, and W. H. Krautschneider, “Design of a mcm gate library applying multiobjective optimization,” in *Proc. IEEE Computer Society Annual Symp. VLSI ISVLSI '07*, pp. 81–85, 2007.
- [5] M. Baru, “Analog design methodologies when every nanoampere counts,” in *CMOS ET Workshop Vancouver 2009*, September 2009.
- [6] F. Krummenacher and N. Joehl, “A 4-mhz cmos continuous-time filter with on-chip automatic tuning,” vol. 23, no. 3, pp. 750–758, 1988.
- [7] R. Boylestad and L. Nashelsky, *Electrónica: teoría de circuitos y dispositivos electrónicos*. Pearson Educación, 8 ed., 2003.
- [8] Y. Tsividis, *Mixed Analog-Digital VLSI Devices and Technology*. Colombia University, USA: World Scientific Publishing, 2002.
- [9] J. M. Rabaey, A. Chandrakasan, and B. Nikolicž, *Digital Integrated Circuits*. Prentice Hall, 2 ed., January 2003.
- [10] R. C. Jaeger and T. N. Blalock, *Microelectronic circuit design*. McGraw-Hill series in electrical and computer engineering, Dubuque, Iowa: McGraw-Hill, 2 ed., 2003.
- [11] S. Franco, *Diseño con amplificadores operacionales y circuitos integrados analógicos*. McGraw-Hill, 3 ed., 2005.
- [12] R. Amezcua-Castrejón, F. Sandoval-Ibarra, and J. Santana, “Diseño de un amplificador operacional de transconductancia completamente diferencial sin cmfb,” *SOMI XV, CINVESTAV-Unidad Guadalajara*, 2000.
- [13] A. Nedungadi and T. Viswanathan, “Design of linear cmos transconductance elements,” vol. 31, no. 10, pp. 891–894, 1984.

- [14] A. Arnaud, *Very Large Time Constant Gm-C Filters*. PhD thesis, Instituto de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay, Abril 2004.
- [15] H. Shichman and D. A. Hodges, “Modeling and simulation of insulated-gate field-effect transistor switching circuits,” vol. 3, no. 3, pp. 285–289, 1968.
- [16] B. R. Group, “Bsim3 homepage.” <http://www-device.eecs.berkeley.edu/~bsim3/>.
- [17] C. Enz, F. Krummenacher, and E. Vittoz, “An analytical mos transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications,” *Journal on Analog Integrated Circuits and Signal Processing*, Kluwer Academic Publishers, pp. 83–114, Julio 1995.
- [18] T. Abate, “Mentor graphics rejects cadence design bid,” *San Francisco Chronicle*, Junio 2008.
- [19] K. S. Desai, “Eda innovation through merger and acquisitions,” *EDACafé*, Setiembre 2006.
- [20] D. Corne and J. Knowles, “The pareto envelope-based selection algorithm for multiobjective optimization,” in *PPSN VI: Proceedings of the International Conference on Parallel Problem Solving from Nature*, pp. 839–848, 2000.
- [21] A. Arnaud, R. Fiorelli, and C. Galup-Montoro, “On the design of very small transconductance otas with reduced input offset,” in *Proc. th Symp. Integrated Circuits and Systems Design*, pp. 15–20, 2005.
- [22] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro, “An mos transistor model for analog circuit design,” vol. 33, no. 10, pp. 1510–1519, 1998.
- [23] B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw-Hill Science/Engineering/Math, 1 ed., August 2000.
- [24] S. C. Chapra and R. P. Canale, *Métodos numéricos para ingenieros*. McGraw-Hill, 5 ed., Enero 2007.

Apéndices

Apéndice A: Consideraciones sobre el slew rate

El OTA a diseñar debe cumplir también con las condiciones de slew rate, es decir, que sea capaz de producir una respuesta tan rápida como lo exija la señal de entrada, cuando este sea finalmente implementado en el filtro GmC. Razavi [23] define matemáticamente el slew rate (SR) de forma general tal y como lo muestra la ecuación A.1:

$$SR = \left. \frac{dv_{out}}{dt} \right|_{v_{out}=v_{outmax}} \quad (A.1)$$

Para un filtro GmC, la derivada anterior forma parte de la ley general del capacitor, por lo que el slew rate para este caso puede reescribirse como:

$$SR = \left. \frac{i_{out}}{C} \right|_{i_{out}=i_{outmax}} \quad (A.2)$$

La máxima corriente de salida se presenta en el caso que ambas corrientes de polarización I_B sean direccionadas hacia la salida. No obstante, estas corrientes son escaladas por la etapa de los espejos de corriente, por lo que esta corriente máxima también es escalada por el factor m . El slew rate entonces podría entonces redefinirse nuevamente como:

$$SR = \frac{2 \cdot I_B}{m \cdot C} \quad (A.3)$$

Utilizando la definición de la frecuencia de corte de un filtro GmC (ecuación 2.56), es posible expresar el slew rate en términos de variables del OTA y de la frecuencia de corte deseada:

$$C = \frac{G_m}{2\pi f_c}$$

$$SR = \frac{2 \cdot I_B \cdot (2\pi f_c)}{m \cdot G_m} \quad (A.4)$$

La ecuación A.4 permite independizar el cálculo del slew rate necesario del condensador del filtro, cuya construcción excede los alcances de este informe. Si se representa la corriente máxima de salida a través de la función de transferencia del OTA ideal (ecuación 2.42), este valor puede verse a partir del valor máximo de la señal de entrada. Si se considera una señal senoidal como entrada de tensión diferencial, el valor máximo ocurre en el valor pico de dicha función (V_{pico}).

$$i_{out_{max}} = G_m \cdot V_{d_{max}}$$

$$i_{out_{max}} = G_m \cdot V_{pico} \quad (A.5)$$

La ecuación anterior permite redefinir el slew rate en términos de la tensión de entrada y la frecuencia de corte del circuito:

$$SR_{min} = V_{pico} \cdot 2\pi f_c \quad (A.6)$$

Si se considera que el slew rate necesario que requiere el filtro debe ocurrir a la frecuencia de corte, debido a que esta será la señal más rápida que debe atravesar de manera correcta el filtro, entonces la ecuación A.6 en realidad identifica al slew rate mínimo que debe poseer el filtro para su correcto funcionamiento. En otras palabras, debe buscarse que $SR \geq SR_{min}$, por lo que retomando las ecuaciones A.4 y A.6 dicha condición implica que:

$$\frac{2 \cdot I_B \cdot (2\pi f_c)}{m \cdot G_m} \geq V_{pico} \cdot 2\pi f_c$$

$$\frac{2 \cdot I_B}{m} \geq V_{pico} \cdot G_m \quad (A.7)$$

En conclusión, si el OTA cumple con la condición de la ecuación A.7, entonces cumple con la condición del slew rate necesario para que el filtro opere correctamente.

Apéndice B: Historial de simulaciones de la herramienta de optimización

A continuación se listan las fechas de las simulaciones realizadas junto con los principales cambios hechos a la herramienta y resultados preliminares obtenidos:

- Jueves 15/4/2010: primera corrida del simulador. Presentaba una definición errónea del rango lineal, mala lectura de parámetros y solo un caso aceptable (rango lineal superior a 1V).
- Viernes 16/4/2010: se corrige el rango lineal pero no la lectura de parámetros y se ejecutan dos corridas, de 500 y 2500 iteraciones, respectivamente. Aparecen dos casos “buenos”.
- Lunes 19/4/2010: se detectan y corrigen los errores mencionados. Comienza a aparecer una mayor cantidad de casos posibles y uno de ellos (caso A) se presenta al profesor Chacón para visualizar el avance logrado gracias a la adaptación de esta herramienta.
- Martes 20/4/2010: se buscan opciones que consuman menor potencia que el caso A. Se encuentran, pero debe sacrificarse un poco el rango lineal.

- Miércoles 21/4/2010: se disminuye la tensión de alimentación de 5V a 3.3V por recomendación del profesor Chacón para disminuir el consumo de potencia y además porque esa es la tensión máxima de alimentación para la tecnología utilizada. Esto ocasiona un cambio en el netlist empleado anteriormente. Se incluye también el cálculo del slew rate, sólo para visualizarlo. Se generan dos simulaciones adicionales, que se diferencian por el rango de corriente establecido para el algoritmo (1-100nA y 1-70nA, respectivamente) y la inclusión del slew rate en una de ellas.
- Viernes 23/4/2010: se evalúan las simulaciones del miércoles. Sólo se genera un caso prometedor. El cálculo del slew rate era la mitad del verdadero, lo que se corrigió y se desarrollaron dos simulaciones (con y sin slew rate, respectivamente).
- Lunes 26/4/2010: se evalúan las simulaciones del viernes 23, sin resultados novedosos. Se detecta un problema: el ancho mínimo para los transistores establecidos en el optimizador estaban por debajo de los límites de la tecnología, lo que inhabilitó todas las simulaciones realizadas pues el ancho mínimo era utilizado en prácticamente todos los resultados prometedores. Se corrige el problema y se ejecuta una nueva simulación.
- Martes 27/4/2010: revisión de la última simulación, la cual indica que el consumo de potencia inevitablemente va a aumentar y los requerimientos de rango lineal apenas se van a cumplir (en los casos anteriores existían varias opciones de grandes rangos lineales y bajo consumo). Con la disminución de la transconductancia debe considerarse la disminución del condensador necesario para el filtro, lo que lo hace más comparable con la capacitancia parásita. Debe procurarse elegir un OTA que mantenga una relación de que el condensador del filtro sea 10 veces mayor a la capacitancia parásita de entrada.
- Martes 3/5/2010: se ejecuta otra simulación, en la que se incluye el slew rate como un cuarto valor de aptitud. Aparecen nuevamente casos aceptables, que luego de preseleccionarlos, deberán simularse en ELDO y visualizar su comportamiento.
- Martes 10/5/2010: se ejecuta una última simulación, en procura de obtener los últimos resultados. De aquí se obtiene el mejor caso y es que se presenta tanto en el paper como en este informe.
- Jueves 27/5/2010: simulación no evaluada por falta de tiempo. Se aumentó la resolución de 5 a 8 bits para cada valor de aptitud y la cantidad de iteraciones a 2500.