

Máster Universitario en Ingeniería de Sistemas Elec-
trónicos y Aplicaciones

2018-2019

Trabajo Fin de Máster

“Análisis y diseño de filtros paso banda
basados en VCO para aplicaciones de
extracción de características vocales”

Fernando Hernández Ruiz

Tutor

Luis Hernández Corporales

Convocatoria: Septiembre

ANÁLISIS Y DISEÑO DE FILTROS PASO BANDA BASADOS EN VCO PARA
APLICACIONES DE EXTRACCIÓN DE CARACTERÍSTICAS VOCALES

RESUMEN

El objetivo del presente trabajo es el estudio, diseño e implementación de una novedosa arquitectura de filtros paso banda basados en VCO, con aplicación directa en tareas de extracción de características vocales.

Partiendo del modelo a nivel de sistema del filtro analógico clásico (filtro de variables de estado), se plantean las ventajas de un modelo basado en VCO, conduciéndolo así hasta el modelo final diseñado, mostrando los resultados obtenidos en simulación.

Además, se propone una posible implementación *hardware* del modelo, realizando así una primera aproximación al diseño a nivel de circuito. Se muestran también simulaciones de las distintas partes que forman el circuito, además de una estimación del consumo de cada una de ellas.

Posteriormente, se realiza un análisis de los resultados obtenidos, tanto a nivel de sistema como a nivel de circuito. Además, se exponen los resultados de simulación de las medidas de energía en un banco de filtros ante distintas señales de entrada. De esta manera, se obtienen patrones de la señal de entrada que posteriormente pueden ser introducidas en redes neuronales o árboles de decisión, con el objetivo de detectar la existencia de eventos.

Por último, se incluye el artículo publicado en el *IEEE International Midwest Symposium on Circuits and Systems* (MWSCAS 2019) que sintetiza las ideas principales de este trabajo.

Palabras clave:

Filtro paso banda, VCO, PFM, oscilador en anillo, detección de eventos.

ANÁLISIS Y DISEÑO DE FILTROS PASO BANDA BASADOS EN VCO PARA
APLICACIONES DE EXTRACCIÓN DE CARACTERÍSTICAS VOCALES

ABSTRACT

The goal of this work is the study, design and implementation of a new Band-Pass filter VCO based architecture, with direct application on feature extraction for speech recognition tasks.

Starting with the system-level model of the traditional analog filter (state variable-filter), VCO based model advantages are presented, leading towards the final model designed and showing the simulation results.

Furthermore, a hardware implementation of the model is presented, thus making a first transistor-level design approach. Simulation results and consumption estimation of each part of the circuit are presented.

After that, results are analysed and compared, both system-level and transistor-level. Additionally, simulation results of energy measurements from a filter bank are presented, with several input signals. By this way, patterns of the input signals are obtained, which can be introduced into neural networks or decision trees, in order to detect events.

To conclude, it is included the paper published in the *IEEE International Midwest Symposium on Circuits and Systems* (MWSCAS 2019), which summarizes the main ideas of this work.

Keywords:

Passband filter, VCO, PFM, ring oscillator, voice activity detection (VAD).

ANÁLISIS Y DISEÑO DE FILTROS PASO BANDA BASADOS EN VCO PARA
APLICACIONES DE EXTRACCIÓN DE CARACTERÍSTICAS VOCALES

INDICE DE CONTENIDOS

1. INTRODUCCIÓN.....	1
2. ESTADO DEL ARTE	3
2.1. Soluciones actuales en la detección de eventos	3
2.2. El oído humano como banco de filtros	5
2.3. La escala de Mel y de Bark.....	6
2.4. Extracción de características vocales.....	7
2.5. Procesamiento de señales con VCOs.....	9
3. DISEÑO DE SISTEMA	13
3.1. Filtro bicuadrático analógico	13
Ganancia del filtro	14
3.2. Filtro bicuadrático con VCOs y contadores ideales.....	16
3.3. Filtro bicuadrático diferencial con VCOs y contadores reales	18
3.3.1. Ventajas de la arquitectura diferencial	18
3.3.2. Diseño del contador asíncrono real	20
3.4. Filtro con VCO común y DCOs	22
3.4.1. Diseño del contador asíncrono de múltiples entradas y salidas.....	22
3.5. Filtro programable	24
3.6. Banco de filtros	26
4. DISEÑO DE HARDWARE	27
4.1. Circuito con VCOs ideales	27
4.2. Circuito con VCOs reales	28
Inversor CMOS	29
Oscilador en anillo CMOS	30
Fuentes de corriente del oscilador en anillo	31
4.3. Caracterización del oscilador en anillo	34
4.4. Diseño de los DCO	35
4.4.1. DCO 1.....	35
Diseño del bloque combinacional del DCO 1	39

ANÁLISIS Y DISEÑO DE FILTROS PASO BANDA BASADOS EN VCO PARA APLICACIONES DE EXTRACCIÓN DE CARACTERÍSTICAS VOCALES

4.4.2. DCO 2.....	40
Diseño del bloque combinacional del DCO 2	43
4.5. Estimación de consumo	44
5. RESULTADOS	47
5.1. Resultados de simulación del filtro.....	47
Simulación con entrada multi-senoidal	47
Respuesta en frecuencia del filtro.....	49
5.2. Extracción del patrón de energías	50
5.3. Consumo y área	52
6. PRESUPUESTO Y PLAN DE TRABAJO	53
6.1. Presupuesto	53
6.2. Plan de trabajo (diagrama de Gantt)	54
7. CONCLUSIONES Y LINEAS FUTURAS DE INVESTIGACIÓN	55
7.1. Conclusiones y objetivos cumplidos.....	55
7.2. Líneas futuras de investigación.....	56
BIBLIOGRAFÍA	57
ANEXOS.....	59
I. ESQUEMÁTICOS	59
II. ARTÍCULO PUBLICADO	73

INDICE DE FIGURAS

Figura 1 – Modelo de VAD [3]	3
Figura 2 – Arquitectura de VAD con procesamiento analógico [4].....	4
Figura 3 – Sistema Auditivo [5]	5
Figura 4 – Comparación de los espectros con filtros de banda ancha y estrecha [10]	6
Figura 5 – Gráfica Mel-Hz	7
Figura 6 – Proceso de extracción de coeficientes PLP [13]	8
Figura 7 – Proceso de extracción de coeficientes MFCC [13].....	9
Figura 8 – Modulador por frecuencia de pulsos [15]	10
Figura 9 – Implementación hardware del PFMI.....	11
Figura 10 – Modulador Sigma-Delta de 3 ^{er} orden implementado con VCOs [17]	11
Figura 11 – Estructura del filtro bicuadrático.....	14
Figura 12 – Respuestas en frecuencia del filtro (Q=1, Q=3, Q=5).....	15
Figura 13 – Arquitectura de filtro bicuadrático KHN	16
Figura 14 – Filtro con VCOs, single-ended y contadores ideales	17
Figura 15 – Contador ideal	17
Figura 16 – Simulación del contador ideal.....	18
Figura 17 – Cancelación del ruido en modo común en arquitectura diferencial.....	19
Figura 18 – Contador real de 2 entradas y 2 salidas.....	21
Figura 19 – Filtro con VCOs, diferencial y contadores reales	21
Figura 20 – Filtro con VCO común, DCOs, diferencial, y contadores ideales	22
Figura 21 – Contador real de 4 entradas.....	23
Figura 22 – Circuito divisor de frecuencias.....	24
Figura 23 – Respuesta en frecuencia del filtro con distintos valores de prescalado	25
Figura 24 – Arquitectura del banco de filtros.....	26
Figura 25 – Bloque modulador (LTspice)	27
Figura 26 – Configuración del bloque modulador como VCO	28
Figura 27 – Entrada y salida del modulador configurado como VCO	28
Figura 28 – Oscilador en anillo	28
Figura 29 – Inversor CMOS [19]	29
Figura 30 – Componentes parásitas del inversor CMOS [19].....	29

Figura 31 – Oscilador en anillo implementado en CMOS	30
Figura 32 – Gráfica frecuencia-tensión de un oscilador en anillo [19]	31
Figura 33 – DCO de 5 etapas con 3 posibles frecuencias	32
Figura 34 – Simulación del DCO de 5 etapas	33
Figura 35 – Entrada y salida del Level Shifter	33
Figura 36 – Simulación del comparador de Schmitt conectado a la salida del oscilador33	
Figura 37 – Circuito de caracterización del VCO	34
Figura 38 – Gráfica Frecuencia-Corriente (200kHz – 2MHz)	34
Figura 39 – Sección DCO 1 del filtro digital a nivel sistema.....	35
Figura 40 – Disposición de las fuentes de corriente en el DCO 1.....	37
Figura 41 – Simulación de las fuentes del bloque combinacional del DCO 1	40
Figura 42 – Sección DCO 2 del filtro digital a nivel sistema.....	40
Figura 43 – Disposición de las fuentes de corriente en el DCO 2.....	43
Figura 44 – Simulación temporal del filtro digital (Filtro 7, Simulink).....	47
Figura 45 – Simulación temporal del filtro digital (Filtro 9, LTspice).....	48
Figura 46 – Espectro de la señal de entrada y de salida del filtro (Filtro 7, Simulink)..	48
Figura 47 – Espectro de la señal de entrada y de salida del filtro (Filtro 9, LTspice)....	49
Figura 48 – Respuesta en frecuencia del filtro analógico y digital	49
Figura 49 – Patrón de energías del banco de filtros (analógicos).....	50
Figura 50 – Señal Chirp.....	51
Figura 51 – Señal Senoidal.....	51
Figura 52 – Señal Aleatoria	52

ABREVIATURAS

ADC	–	Analog-to-Digital Converter
BM	–	Basilar Membrane
BPF	–	Band Pass Filter
DCO	–	Digital Controlled Oscillator
DCT	–	Discrete Cosine Transform
FFT	–	Fast Fourier Transform
IAF	–	Integrate And Fire
LNA	–	Low Noise Amplifier
MEMS	–	Micro-Electro-Mechanical System
MFCC	–	Mel Frequency Cepstrum Coefficients
PFM	–	Pulse Frequency Modulation
PFMI	–	Pulse Frequency Modulation Integrators
PSRR	–	Power Supply Rejection Ratio
PWM	–	Pulse Width Modulation
SNR	–	Signal-to-Noise Ratio
STFT	–	Short-Time Fourier Transform
THD	–	Total Harmonic Distortion
VAD	–	Voice Activity Detection
VCO	–	Voltage Controlled Oscillator

ANÁLISIS Y DISEÑO DE FILTROS PASO BANDA BASADOS EN VCO PARA
APLICACIONES DE EXTRACCIÓN DE CARACTERÍSTICAS VOCALES

CAPÍTULO 1

INTRODUCCIÓN

En la actualidad, el reconocimiento automático del habla o reconocimiento automático de voz está presente en numerosos dispositivos que tenemos a nuestro alcance. Los asistentes de voz se apoyan principalmente en esta disciplina (ya que proporciona la interfaz dispositivo-usuario) y, a lo largo de los últimos años, han sido una clara demostración de los importantes avances alcanzados.

Se trata, sin embargo, de una disciplina que comenzó a estudiarse desde principios del siglo XX. Por motivos personales, Alexander Graham Bell fue uno de los pioneros en investigar posibles mecanismos de transcripción, sin llegar a obtener resultados satisfactorios. Sin embargo, los conocimientos que adquirió le llevaron posteriormente a patentar el teléfono.

No fue hasta mediados del siglo XX cuando comenzó a haber avances significativos en el reconocimiento de voz. *AT&T Bell Labs* construyeron el primer sistema de reconocimiento del habla, capaz de reconocer los dígitos del 0 al 9 en inglés [1].

Desde entonces, las tareas de investigación no han cesado para conseguir un sistema de reconocimiento automático de voz cada vez más preciso y robusto. Gracias a los grandes avances en capacidad de computación de las últimas décadas y a la implantación de las redes neuronales, los asistentes de voz se han visto notablemente favorecidos.

Hoy en día, el principal inconveniente de esta herramienta es el alto consumo que supone, especialmente en dispositivos móviles, ya que necesita estar funcionando de forma continuada. Como solución, en los últimos años se están proponiendo distintos sistemas de detección de actividad de voz (*Voice Activity Detection* o VAD). Mientras que el reconocimiento de voz tiene como objetivo transcribir un mensaje transmitido oralmente, en el caso de la detección de actividad de voz es determinar si en una situación hay alguien hablando o si, por el contrario, sólo existe ruido de ambiente. Estos dos sistemas no suelen emplearse independientemente, sino que se complementan. Habitualmente, el sistema VAD es el que permanece encendido, activando el sistema de reconocimiento en caso de detectar actividad de voz. Esto supone una disminución notable del consumo, ya que evita que el procesador principal del dispositivo esté en funcionamiento en un entorno silencioso o de ruido.

El sistema VAD requiere de una etapa de pre-procesamiento del audio y una etapa de toma de decisión (red neuronal o similar). El pre-procesamiento lleva ligado, habitualmente, el uso de un banco de filtros paso banda, como se verá en el estado del arte.

La motivación del presente trabajo es diseñar y analizar una arquitectura de filtros dedicados a dicho pre-procesamiento, que sea más eficiente en área y consumo que las soluciones existentes hoy en día.

En el capítulo 2 se presenta el estado del arte, donde se muestran algunas de las soluciones más actuales, aún en desarrollo. Posteriormente se describe brevemente la anatomía del oído humano, modelable como banco de filtros. Se muestran también algunas de las principales técnicas de extracción de características vocales. Por último, se desarrolla el concepto de procesado de señales mediante VCOs y las implicaciones directas que tiene sobre el presente trabajo.

En el capítulo 3 se desglosa el diseño propuesto a nivel de sistema, partiendo del modelo analógico convencional de un filtro biquadrático y llegando finalmente al modelo propuesto basado en el uso de VCOs, programable por octavas.

En el capítulo 4 se expone el diseño hardware de la arquitectura propuesta, empleando para ello contadores reales, circuitos combinacionales, fuentes de corriente y osciladores en anillo. En este capítulo se muestran resultados de simulaciones funcionales de algunas de sus partes, así como una primera estimación del consumo del circuito.

En el capítulo 5 se recopilan, comparan, y analizan los resultados obtenidos tanto a nivel de sistema como a nivel de circuito. También se extraen los resultados procedentes de un banco de filtros al que se introducen distintas señales de entrada.

En el capítulo 6 se muestra el presupuesto y el plan de trabajo llevado a cabo, haciendo uso de un diagrama de Gantt.

En el capítulo 7 se presentan las conclusiones extraídas del trabajo, los objetivos alcanzados y las posibles líneas futuras de investigación.

Por último, se incluye un capítulo de anexos, donde se registran todos los modelos diseñados, analizados y simulados en el presente trabajo, tanto en Simulink como en LTspice. Adicionalmente, se incluye el artículo publicado en el *IEEE International Midwest Symposium on Circuits and Systems* (MWSCAS 2019).

CAPÍTULO 2

ESTADO DEL ARTE

2.1. Soluciones actuales en la detección de eventos

Existen innumerables enfoques en el campo del reconocimiento del habla y la detección de eventos. No es el objeto de este proyecto presentarlas todas y compararlas, sino encontrar una solución a algunos de los problemas que existen en las arquitecturas propuestas a día de hoy.

Tradicionalmente, las tareas de procesado de señales se han realizado mayoritariamente con técnicas y circuitos digitales, debido a su versatilidad, escalabilidad y capacidad de computación. Por ejemplo, cuando se desea comprimir una muestra de audio, obtener su transformada de Fourier, eliminar su ruido de fondo o aplicar cualquier otro efecto, la inmensa mayoría de soluciones se apoyan en el uso de tecnología digital. Como consecuencia, muchos sistemas de reconocimiento de audio se apoyan principalmente en dichas técnicas [2]. Sin embargo, estos procedimientos presentan una desventaja: habitualmente presentan un alto consumo. Algunas soluciones digitales recientes han conseguido consumos muy reducidos mediante la técnica de *wake-up* que proporciona el sistema VAD (Figura 1).

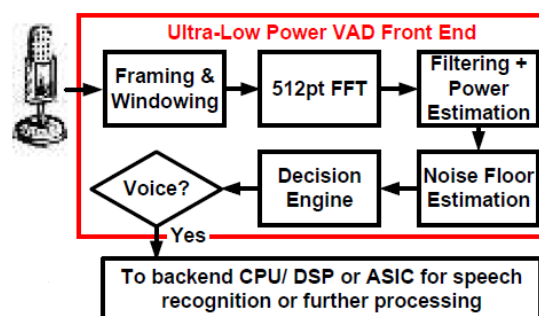


Figura 1 – Modelo de VAD [3]

La muestra de entrada digitalizada se divide en *frames* y se aplica una ventana a cada uno de ellos para evitar el fenómeno de fuga espectral. Posteriormente se aplica la transformada rápida de Fourier (FFT). Tras realizar un proceso de filtrado, se estima la potencia de la señal y el nivel de ruido y, en función de los resultados obtenidos, el motor de decisión (previamente entrenado) determina si existe voz o ruido en la muestra. Tal y como indica el artículo, mediante el uso de frecuencia y alimentación duales, se

reduce considerablemente el consumo, debido principalmente a factores de escalabilidad.

Sin embargo, la tendencia de las últimas investigaciones apunta hacia un diseño híbrido digital-analógico que, mediante un proceso similar a la transformada de Fourier, consiga resultados similares con un consumo muy inferior. Este planteamiento se ve plasmado en artículos muy recientes [4].

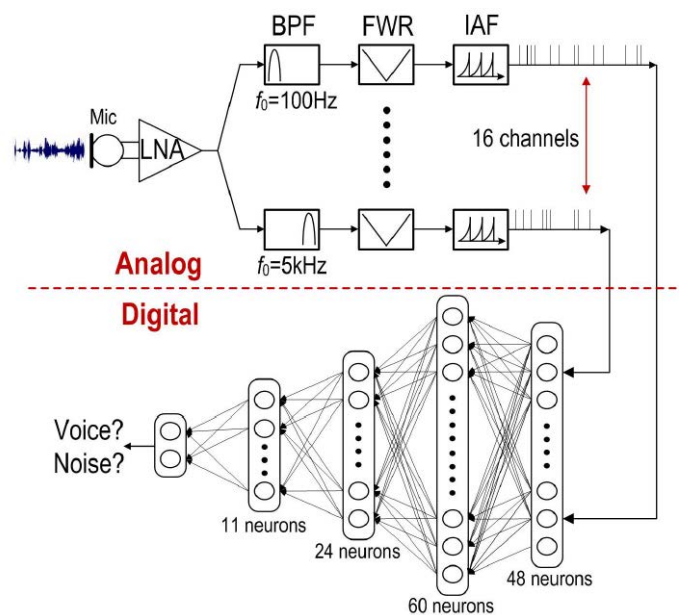


Figura 2 – Arquitectura de VAD con procesamiento analógico [4]

Como se puede observar en la Figura 2, la señal procedente del micrófono pasa por un amplificador de bajo ruido (LNA). Posteriormente la señal se bifurca en 16 canales con sus correspondientes filtros paso banda analógicos (BPF), rectificadores (FWR) y bloques codificadores IAF (*Integrate and Fire*). De esta manera, se realiza un proceso analógico similar a la transformada de Fourier. Por último, los pulsos procedentes de los 16 canales se introducen en una red neuronal entrenada, que determina si en la muestra procesada hay voz o solo ruido.

La transformada de Fourier analógica es un proceso que mimetiza el comportamiento del oído humano como banco de filtros. Cada filtro extrae información en forma de energía en su banda de paso correspondiente, de modo que el cómputo total de las energías forma una imagen similar a la de un espectrograma. En el siguiente apartado se desarrolla esta característica del oído humano que desde hace décadas se viene estudiando. El modelado de este sistema tan fundamental y a la vez tan complejo es una tarea necesaria en el campo del reconocimiento del habla y la detección de eventos.

2.2. El oído humano como banco de filtros

La percepción tonal o de altura es una actividad llevada a cabo por el sistema auditivo. Exceptuando los elementos más periféricos, el conocimiento sobre este sistema aún es muy limitado y fragmentado. Por esta razón, el modelado de este sistema depende en gran medida de los estudios psicoacústicos en los que se apoye.

El sistema auditivo se divide en dos partes principales: el sistema auditivo periférico y el sistema auditivo central. A su vez, el periférico, también denominado oído, se divide en oído externo, oído medio y oído interno. El primero canaliza la energía acústica, el segundo la amplifica, y el tercero la transforma en impulsos eléctricos. Por tanto, es el oído interno el que actúa como transductor. Esta tarea la realiza la cóclea, una estructura con forma de caracol que contiene una membrana denominada basilar. La membrana basilar (BM) es delgada y rígida en la base (zona más externa), y más gruesa y maleable en el apéndice (zona más interna). Los estereocilios son las células que conforman la membrana, capaces de transformar las vibraciones en impulsos. Debido a las diferencias de rigidez de la membrana, las vibraciones de baja frecuencia (tonos graves) son capaces de atravesar toda la cóclea y excitar los estereocilios de la zona interna. Por el contrario, las altas frecuencias excitan tan solo la membrana en la base y sus vibraciones se disipan en la zona interna. De esta manera, la cóclea puede interpretarse como un conjunto de muchos filtros que se excitan en función de la frecuencia de la señal de entrada. En la Figura 3 se puede apreciar la estructura de la cóclea tal y como se ha descrito:

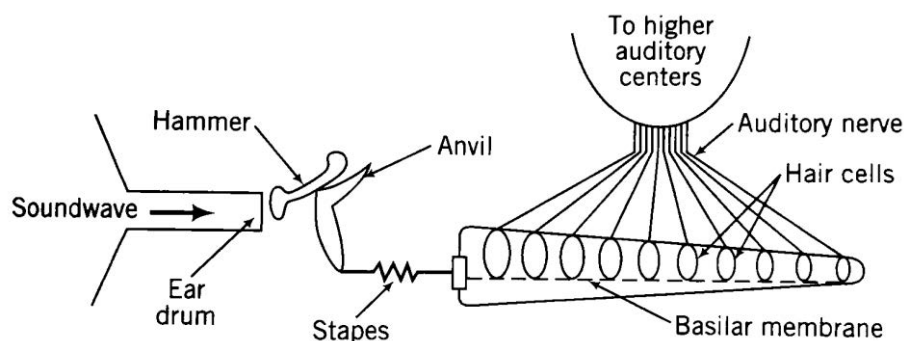


Figura 3 – Sistema Auditivo [5]

Hermann von Helmholtz fue el primero en concebir el sistema auditivo como un banco de filtros paso banda solapados. Este conjunto de filtros le permite al oído realizar un proceso de descomposición de la señal equivalente a la transformada de Fourier. Desde entonces, en numerosas investigaciones se ha modelado el oído de esta manera, llevando a cabo una tarea de mimesis muy compleja, que requiere de experimentos psicoacústicos [6]-[9].

El número de filtros que debe contener el banco no tiene una única respuesta. Históricamente se ha resuelto mediante ensayo y error. Por lo general, la cifra oscila entre 10 y 20 filtros. Tampoco existe consenso en el ancho de banda de los filtros. Dependiendo de

si se usan filtros más restrictivos o menos, se obtienen a la salida espectros como los que muestra la Figura 4.

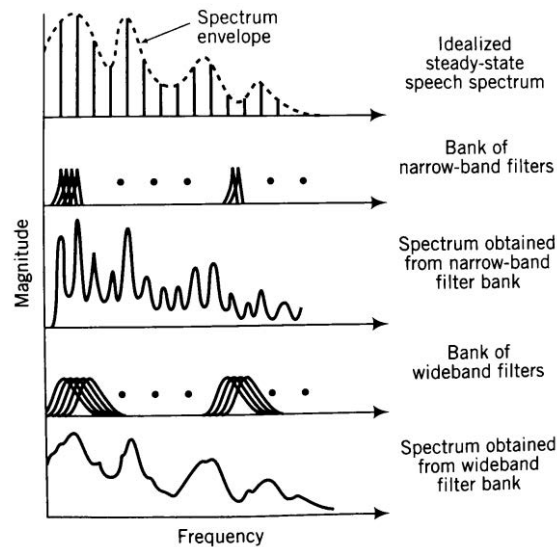


Figura 4 – Comparación de los espectros con filtros de banda ancha y estrecha [10]

2.3. La escala de Mel y de Bark

Para comprender algunas de las principales técnicas de extracción de características vocales, es necesario introducir algunas escalas definidas específicamente para estas tareas.

Se conoce como escala de Mel a la escala musical perceptual cuyos intervalos se sitúan equiespaciados para sus oyentes. Fue definida en 1937 por Stevens, Volkman y Newman, y su nombre procede de la palabra *melody*, ya que es una escala basada en la comparación de tonos.

La fórmula más popular para obtener los mels a partir de los hercios es la siguiente:

$$mel(f) = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right) \quad (1)$$

Sin embargo, al tratarse de una escala perceptual, es decir, completamente subjetiva y que varía dependiendo de cada oyente, no existe una única expresión. Diferentes fórmulas han sido propuestas desde entonces para definir la escala de Mel.

El comportamiento logarítmico de esta escala es el reflejo de una característica importante del oído humano: a bajas frecuencias se percibe con mayor facilidad las variaciones en el tono. A medida que la frecuencia aumenta, dichas variaciones deben ser mayores para que el oyente las perciba. La escala de Mel, por tanto, refleja el comportamiento no-lineal del oído, y es un buen punto de partida para los sistemas de reconocimiento del habla y de reconocimiento de eventos.

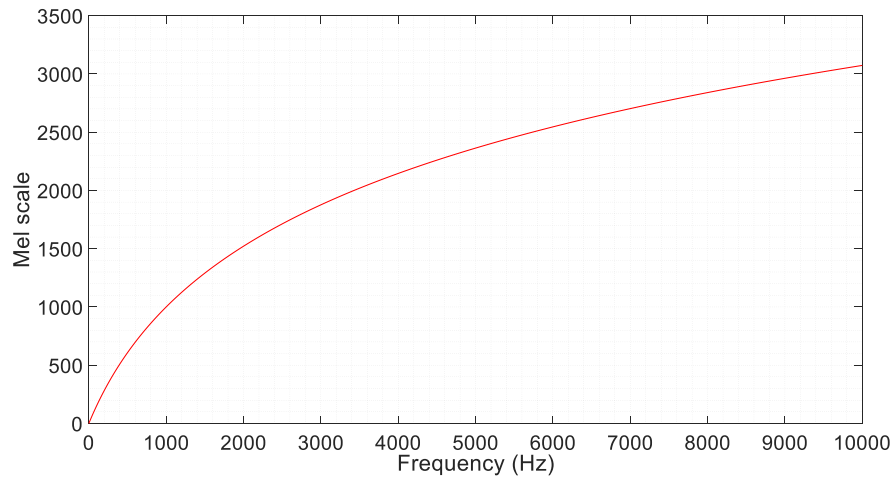


Figura 5 – Gráfica Mel-Hz

La escala de Bark ofrece una alternativa perceptual a la escala de Mel. A diferencia de esta, la escala de Bark se centra más en el ancho de las bandas críticas que en la frecuencia central de estas. La escala de Bark se rige por la expresión:

$$bark = 6 \cdot \log_e \left(\left(\frac{f}{600} \right) + \sqrt{\left(\frac{f}{600} \right)^2 + 1} \right) \quad (2)$$

Existe cierta controversia en el uso de este tipo de escalas, dado que se basan en resultados subjetivos. El uso de una u otra no es determinante para un buen o mal funcionamiento de un sistema, ya que esto dependerá de otros factores de diseño (por ejemplo, de cómo se entrene la red neuronal que se sitúa en etapas posteriores) [11]. No obstante, estas escalas juegan un papel importante en el cálculo de los coeficientes para la extracción de características vocales, como se verá a continuación.

2.4. Extracción de características vocales

La extracción de características es una tarea conocida en el ámbito del *machine learning*, el reconocimiento de patrones y el procesamiento de audio e imagen. Se trata de un proceso mediante el cual, a partir de una muestra inicial, se extraen ciertas características relevantes que permiten analizarla y clasificarla de forma efectiva y no redundante con una finalidad determinada.

En el caso de la extracción características vocales, la finalidad es obtener un conjunto de parámetros que definan la muestra inicial, facilitando así tareas como la detección de eventos (si hay alguien hablando), la clasificación del orador (quién está hablando) o el reconocimiento del habla (qué está diciendo).

Históricamente, los métodos más utilizados para la extracción de características vocales en el ámbito del reconocimiento del habla son:

- **Codificación Predictiva Lineal (LPC)**

Esta técnica se apoya sobre la suposición de que la voz es producida por un zumbador al final de un tubo, con la adición de sonidos sibilantes y sonidos plosivos. Esto, que a priori puede parecer una simplificación torpe, se aproxima notablemente a la realidad. La glotis produce un zumbido de una intensidad y frecuencia específica. Acto seguido, el tracto vocal, que es, en esencia, un tubo con resonancias características, se encarga de producir los formantes. Los formantes son picos de intensidad presentes en el espectro de un sonido.

La técnica LPC realiza el proceso inverso cada cierto periodo de tiempo o *frame*, obteniendo así los coeficientes LPC. Mediante estos coeficientes (a_k), que describen los formantes, se genera un modelo lineal a partir de la combinación de las N muestras anteriores y una señal de error:

$$s[n] = \left(\sum_{k=1}^N a_k \cdot s[n - k] \right) + e[n] \quad (3)$$

Esta técnica es ampliamente usada tanto en reconocimiento del habla, como en sintetizadores de voz y como método de compresión de voz (en estándares como GSM) [12].

- **Predicción Lineal Perceptiva (PLP)**

Esta técnica guarda gran similitud con la técnica LPC. La principal diferencia radica en que la predicción lineal perceptiva modifica las características espectrales atendiendo a las que posee el sistema auditivo, descartando así información irrelevante. Esto puede entenderse de una manera más simple: el oído humano se ha desarrollado de tal forma que percibe con gran claridad el espectro de la voz humana. Por tanto, si se aplican ciertas modificaciones al espectro de una señal, se pueden obtener unos coeficientes que determinarán posteriormente si dicha señal pertenece a una voz humana o no. Para obtener esos coeficientes, denominados coeficientes cepstrales PLP, se lleva a cabo el proceso de la Figura 6.

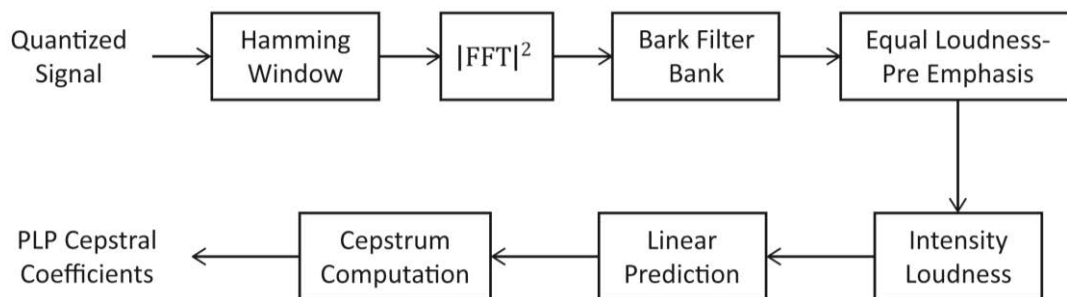


Figura 6 – Proceso de extracción de coeficientes PLP [13]

Este proceso incluye el uso de un banco de filtro Bark, que consiste en un conjunto de filtros espaciados haciendo uso de la escala Bark, no en la escala de hercios tradicional.

Como consecuencia de esto, el método PLP se adapta mejor que el LPC a las particularidades del sistema auditivo (que es altamente no lineal).

- Coeficientes Cepstrales en la Frecuencia Mel (MFCC)

El método dominante en la actualidad es el uso de los Coeficientes Cepstrales en la Frecuencia Mel (MFCC). La palabra *cepstrum* (que también aparece en el método PLP) procede de *spectrum*, dándole la vuelta a la mitad de la palabra. El cepstrum de una señal es el resultado de aplicarle la Transformada de Fourier inversa de su espectro en escala logarítmica.

Para calcular los MFCC se procede de la siguiente manera: se aplica un filtro de pre-énfasis a la señal. Se divide la señal en trozos o *frames* solapados y se aplica una ventana (generalmente de Hamming) para evitar el fenómeno de fuga espectral. Tras ello, se realiza la transformada de Fourier de tiempo reducido (STFT) en cada uno de los *frames*, obteniendo así su espectro. Cada *frame* se introduce en un banco de filtros equiespaciados en la escala Mel. Posteriormente se aplica el logaritmo y la transformada de coseno discreta (DCT). La DCT es un proceso muy utilizado en compresión de imágenes, ya que tiene una buena capacidad de compactación de energía. Como resultado de esta operación, se obtienen los coeficientes, generalmente normalizados entre 0 y 1, cuyo valor depende de la importancia que tiene en el vector acústico. La Figura 7 muestra el procedimiento seguido para obtener los MFCCs.

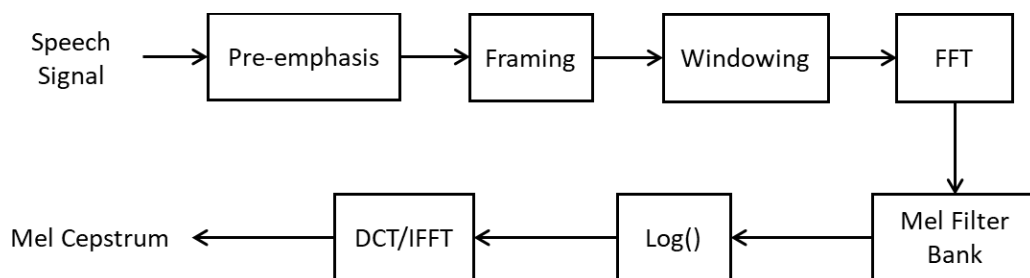


Figura 7 – Proceso de extracción de coeficientes MFCC [13]

Como se ha podido observar, el uso de bancos de filtros en las tareas de extracción de características vocales es esencial, por lo que el estudio de estos es una labor importante de cara a reducir el consumo y tamaño en futuras implementaciones.

Como ya se indicó inicialmente, la principal motivación de este trabajo es mostrar una alternativa novedosa para implementar el banco de filtros, con un consumo y área menor, en la medida de lo posible, a las soluciones actuales.

2.5. Procesamiento de señales con VCOs

El uso de VCOs para el procesamiento de señales es el pilar fundamental de este trabajo. Tal y como se expone en [14], los integradores, generalmente implementados con operacionales y RC (Opamp-RC) o con amplificadores de transconductancia y capacidad (Gm-C), también pueden realizarse con VCOs. Los Opamp-RC presentan una gran

linealidad con respecto a los Gm-C. Sin embargo, su ancho de banda disponible es menor. Estos integradores presentan problemas de ganancia en DC, por lo que se recurre al uso de cascodos o de etapas conectadas en cascada. No existen soluciones a dicho problema que no conlleven una penalización en potencia o ancho de banda.

En el caso de un integrador mediante osciladores en anillo, su ganancia en DC es infinita e independiente de las dimensiones de los transistores y de la alimentación, lo cual es una ventaja notable en dispositivos móviles. Además, el uso de osciladores en anillo permite que el procesado de señales sea en el dominio del tiempo y no en el dominio de la tensión. Consecuentemente, no es necesario el uso de grandes tensiones para conseguir valores altos de SNR (*Signal-to-Noise-Ratio*).

En [15] se demuestra que, al situar un detector de pulsos tras un VCO, se obtiene un modulador de frecuencia de pulsos (PFM), que actúa como un integrador (Figura 8).

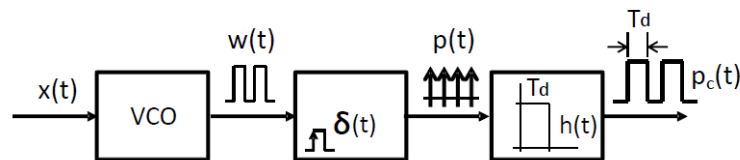


Figura 8 – Modulador por frecuencia de pulsos [15]

La PFM es una técnica de modulación mediante la cual se representa una señal analógica utilizando sólo 2 niveles de tensión. A diferencia de la modulación de ancho de pulso (PWM), en la cual se codifica la magnitud de la entrada mediante una señal de frecuencia fija y anchura de pulso variable, en la PFM el ancho de pulso es constante y es la frecuencia de pulsos la que varía. A mayor amplitud de la entrada, mayor frecuencia con la que se suceden los pulsos.

Se puede apreciar cómo, en la Figura 8, el VCO genera una señal de frecuencia variable $w(t)$ que depende de la amplitud de $x(t)$. Posteriormente, la señal atraviesa un bloque generador de impulsos de Dirac, produciendo así impulsos en los flancos de subida de la señal $w(t)$, generando así $p(t)$. Por último, un filtro con respuesta cuadrada ante impulso proporciona la señal $p_c(t)$, de anchura fija y frecuencia variable. De esta manera, el valor medio de la señal varía acorde a la señal de entrada. Si la amplitud de la señal de entrada aumenta, también lo hace la señal a la salida del VCO. Habiendo fijado el ancho de pulso, al aumentar la frecuencia, aumenta el valor medio de la señal. Posteriormente, se puede demodular la señal original mediante un filtro paso bajo suficientemente restrictivo.

En realidad, el bloque $\delta(t)$ es conceptual y no es necesario su uso. La Figura 9 muestra la implementación *hardware* del bloque PFMI. El VCO seguido de un contador actúa como un bloque integrador de modulación por frecuencia de pulsos (PFMI). También es necesario el oscilador de referencia para evitar que el oscilador en estado de reposo haga desbordar al contador.

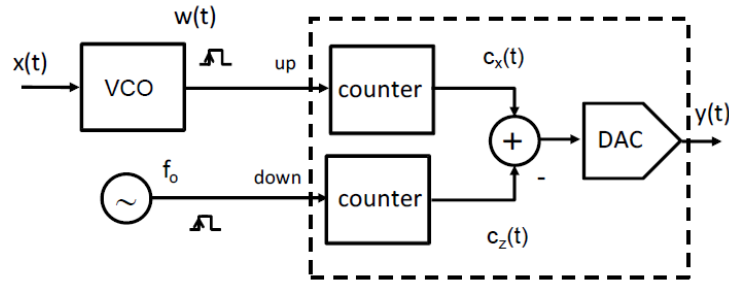


Figura 9 – Implementación hardware del PFMI

El uso de un VCO junto con un contador para formar un integrador es un recurso reciente pero cada vez más empleado en aplicaciones de audio. En [16], se presenta un convertidor analógico-digital (ADC) Sigma Delta ($\Sigma\Delta$) de segundo orden, basado en VCOs. Mediante el uso de osciladores, se consigue reducir drásticamente el área ocupada de silicio y, consecuentemente, su precio. Se evita de esta manera el uso de amplificadores operacionales u otros circuitos lineales.

En [17] se encuentra otro ejemplo reciente en el uso de VCOs para procesar información y utilizarlo como bloque integrador en un ADC sigma delta de tiempo continuo (Figura 10).

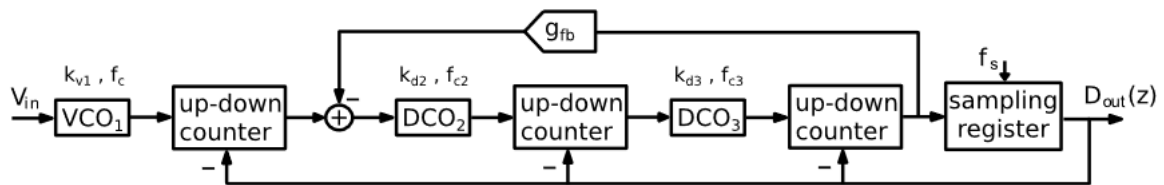


Figura 10 – Modulador Sigma-Delta de 3^{er} orden implementado con VCOs [17]

En definitiva, se trata de una técnica muy potente, de reciente aparición, que puede dar lugar a nuevas arquitecturas. De esta manera, todas aquellos circuitos que contienen bloques integradores pueden verse favorecidos al introducir bloques PFMI. Es el principal objetivo de este proyecto implementar una arquitectura de filtros apoyada en esta técnica de procesamiento de señales y analizar los resultados obtenidos.

CAPÍTULO 3

DISEÑO DE SISTEMA

3.1. Filtro bicuadrático analógico

El filtro bicuadrático o filtro de variables de estado es una topología de segundo orden, utilizada de forma extendida en diseños analógicos debido a su versatilidad. Su arquitectura permite obtener la salida de filtro paso alto, paso banda y paso bajo simultáneamente en distintos puntos del sistema.

La función de transferencia general de este filtro, como su propio nombre indica, se compone de dos funciones cuadráticas:

$$H(s) = \frac{a_1 s^2 + b_1 s + c_1}{a_2 s^2 + b_2 s + c_2} \quad (4)$$

Dependiendo del valor de los coeficientes del numerador, la función de transferencia obtenida es la del filtro paso alto, banda o bajo. Para obtener la función de transferencia del filtro paso banda, los coeficientes a_1 y c_1 han de ser nulos. Situando varios bicuadráticos en cascada se da lugar a un filtro de mayor orden.

La función de transferencia del paso banda resulta:

$$H_{BP}(s) = \frac{k \cdot s}{s^2 + \frac{\omega_0}{Q} s + \omega_0^2} \quad (5)$$

Donde ω_0 es la frecuencia natural o frecuencia propia del filtro (donde la ganancia es mayor) y Q es el factor de calidad o de selectividad del filtro. Este factor se define como la frecuencia natural dividida por el ancho de banda:

$$Q = \frac{f_0}{f_2 - f_1} = \frac{f_0}{BW} \quad (6)$$

Un factor Q alto indica una alta restricción de frecuencias de paso y un pico de resonancia mayor, mientras que un filtro con factor Q bajo es menos restrictivo y posee un pico de resonancia menos pronunciado [18].

Ganancia del filtro

Dada la expresión (5), la amplitud en el pico de resonancia, que se encuentra en la frecuencia natural del filtro, tiene un valor de:

$$|H_{MAX}| = |H(\omega_0)| = \frac{k \cdot \omega_0}{\omega_0^2 + \frac{\omega_0}{Q} \omega_0 - \omega_0^2} = \frac{a_1}{\frac{\omega_0}{Q}} = \frac{k \cdot Q}{\omega_0} \quad (7)$$

En el caso de la arquitectura propuesta, la amplitud máxima a la entrada sin que el filtro sature es aquella que multiplicada por el valor de la ganancia de un valor de 1. Es decir:

$$A_{MAX_INPUT} = \frac{1}{|H_{MAX}|} \quad (8)$$

El diagrama de bloques del filtro biquadrático se muestra en la Figura 11.

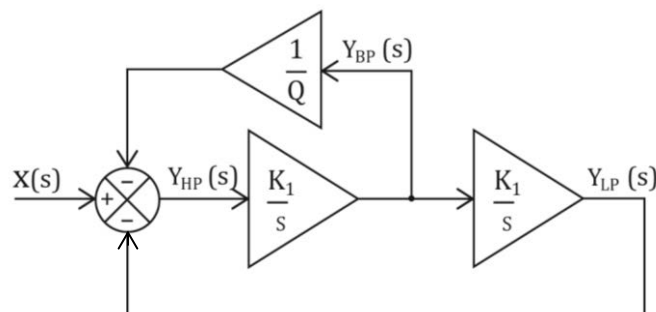


Figura 11 – Estructura del filtro biquadrático

Donde K_1 es la ganancia del integrador. Como se indicó inicialmente, las tres posibles salidas están presentes en el modelo, de forma simultánea. En el presente trabajo, solamente Y_{BP} (*Band Pass*) es la señal de salida de interés, por lo que, de ahora en adelante, al mencionar la salida del filtro se estará refiriendo en todo momento a la del paso banda.

Se han realizado simulaciones en Matlab del modelo analógico mostrado en la Figura 11 (Filtro 1, en el apartado Anexos) para poder conocer la respuesta en frecuencia del filtro. Para ello, se ha introducido un seno cardinal, o función sinc, en el dominio del tiempo. La función sinc se define como:

$$sinc(x) = \frac{\sin(x)}{x} \quad (9)$$

Una de las características principales de esta función es que su transformada de Fourier corresponde a una función rectangular. Por tanto, la función sinc en el dominio del tiempo equivale a un rectángulo en el dominio de la frecuencia, y viceversa. Este es un fenómeno favorable debido a que, escogiendo adecuadamente su periodo, al introducir la función sinc al filtro, en el dominio de la frecuencia se estaría introduciendo un rec-

tángulo de ancho de banda finito. Es decir, si la frecuencia del seno cardinal es de 20kHz, el rectángulo introducido al filtro en el dominio de la frecuencia tendría un ancho de banda de 20kHz, que es el rango de interés para aplicaciones de audio.

La Figura 12 muestra la señal de entrada en el dominio de la frecuencia y la respuesta del filtro para distintos valores de factor Q. El filtro se comporta según lo esperado, mostrando una respuesta paso banda hasta llegar al ancho de banda de la señal de entrada. El esquemático simulado corresponde al Filtro 1 de los anexos.

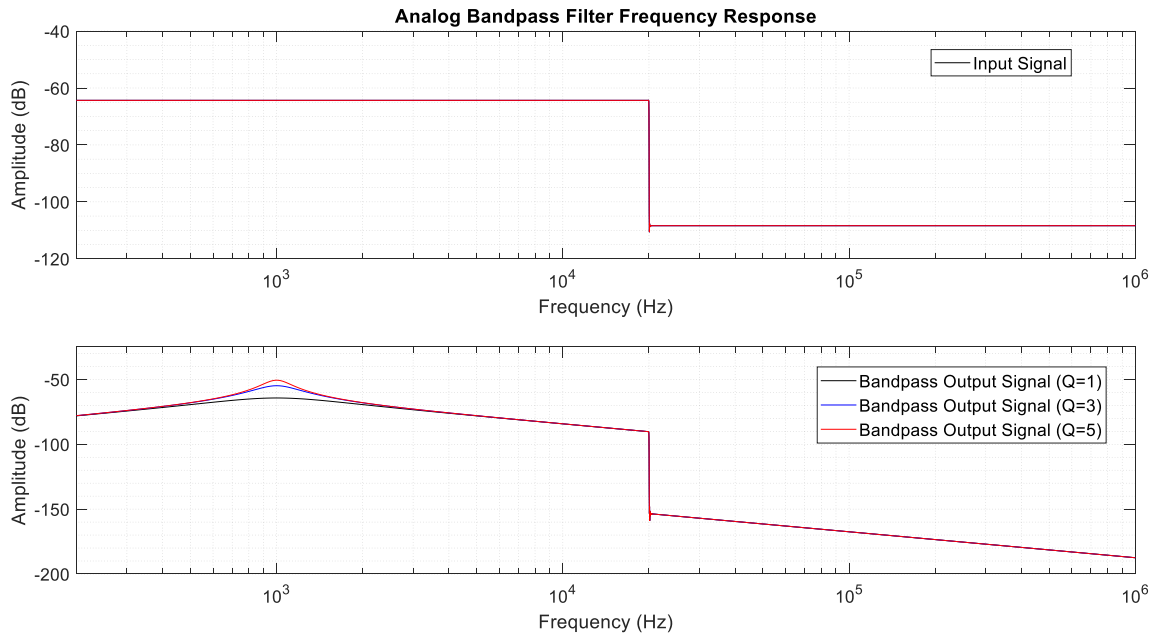


Figura 12 – Respuestas en frecuencia del filtro (Q=1, Q=3, Q=5)

Existen diversos métodos para identificar y caracterizar sistemas, además de usar el seno cardinal, como por ejemplo utilizar un impulso o delta de Dirac, una señal de tipo multi-seno o una señal *chirp*.

Los filtros bicuadráticos analógicos se implementan comúnmente con amplificadores operacionales, resistencias y condensadores, llegando a obtenerse niveles de consumo muy reducidos. A modo de ejemplo, se muestra en la Figura 13 una de las múltiples arquitecturas de filtros de variables de estado que existen: el filtro Kerwin-Huelsman-Newcomb (KHN). Su función de transferencia se expresa:

$$\frac{V_{out}}{V_{in}}(s) = \frac{c_1}{a_2 s^2 + b_2 s + c_2} \quad (10)$$

Cuyos coeficientes se obtienen a partir de los valores de sus componentes:

$$c_1 = \frac{R_2(R_4 + R_5)}{R_4(R_1 + R_2)} \quad (11)$$

$$a_2 = R_1 R_3 C_1 C_2 \quad (12)$$

$$b_2 = \frac{R_1 R_3 C_2 (R_4 + R_5)}{R_4 (R_1 + R_2)} \quad (13)$$

$$C_2 = \frac{R_5}{R_4} \quad (14)$$

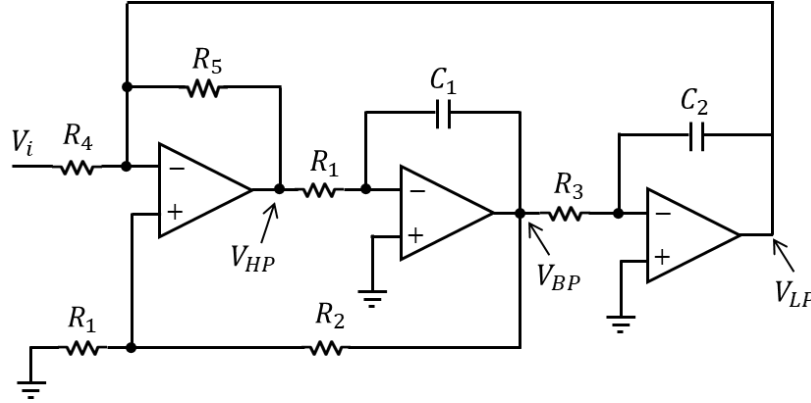


Figura 13 – Arquitectura de filtro bicuadrático KHN

Por último, la frecuencia natural del filtro y su factor Q se rigen por las expresiones:

$$\omega_n = \sqrt{\frac{R_5}{R_1 R_3 R_4 C_1 C_2}} \quad (15)$$

$$Q = \sqrt{\frac{R_5 R_4}{R_1 R_3 C_1 C_2}} \cdot \frac{R_1 + R_2}{R_1 R_3 C_3 (R_4 + R_5)} \quad (16)$$

Cuando la frecuencia natural ha de ser baja (cientos de hercios), la constante de tiempo del filtro es muy alta, lo que conlleva a utilizar condensadores y resistencias de gran valor. Si el circuito debe de ser integrado, el área ocupada se incrementa prohibitivamente, lo cual lleva a la búsqueda de otras soluciones.

La alternativa que en el presente trabajo se propone, y es el núcleo de esta investigación, consiste en sustituir los integradores analógicos por bloques PFMI.

3.2. Filtro bicuadrático con VCOs y contadores ideales

Como ya se desarrolló en el apartado 2.5, un VCO seguido de un contador puede actuar como un integrador. El primer paso, por tanto, es sustituir los integradores analógicos por bloques PFMI. Para ello, se deben calcular los coeficientes de las ganancias en los PFMI para que el modelo resultante coincida con el analógico.

Además, es necesario incluir un tercer VCO que sirva de referencia para que los contadores no se saturen. Es decir, si la entrada es nula, las señales de los VCOs se cancelan mutuamente. La frecuencia del VCO de referencia ha de ser igual a la frecuencia de reposo de los demás VCOs. La Figura 14 muestra la configuración propuesta.

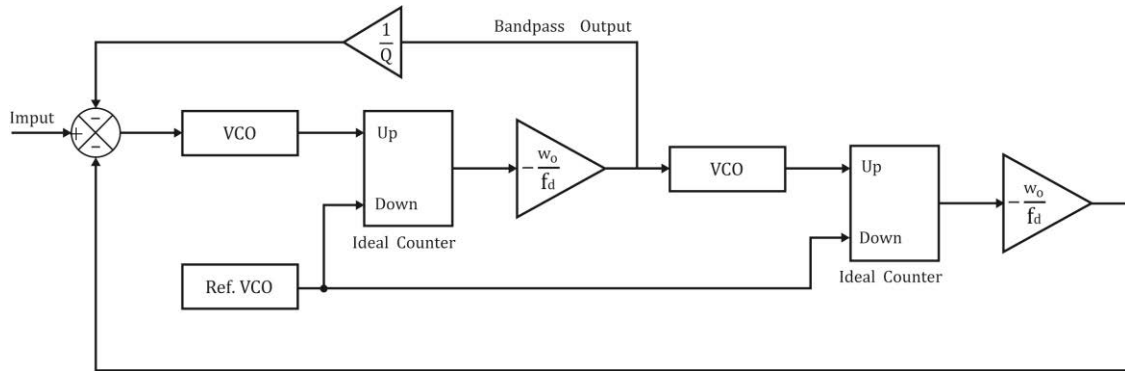


Figura 14 – Filtro con VCOs, single-ended y contadores ideales

Los contadores empleados en primera instancia son ideales, y su estructura interna es la siguiente:

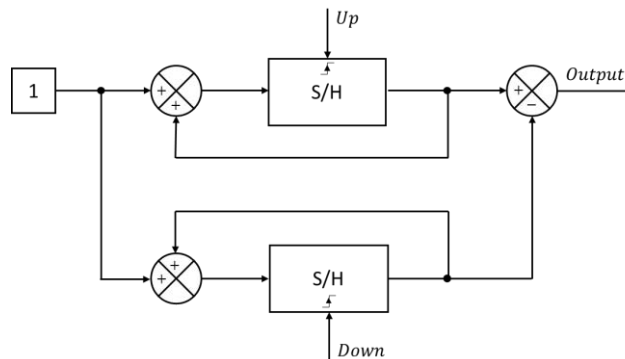


Figura 15 – Contador ideal

La función de este contador es incrementar o disminuir el valor de salida en función de los pulsos que llegan por la entrada Up y la entrada $Down$. Es decir, si el número de pulsos en la entrada Up es mayor que en la entrada $Down$, la salida mostrará una escalera ascendente. Si la entrada $Down$ recibe más pulsos que la entrada Up , la escalera será descendente. Este comportamiento se ve reflejado en la simulación mostrada en la Figura 16.

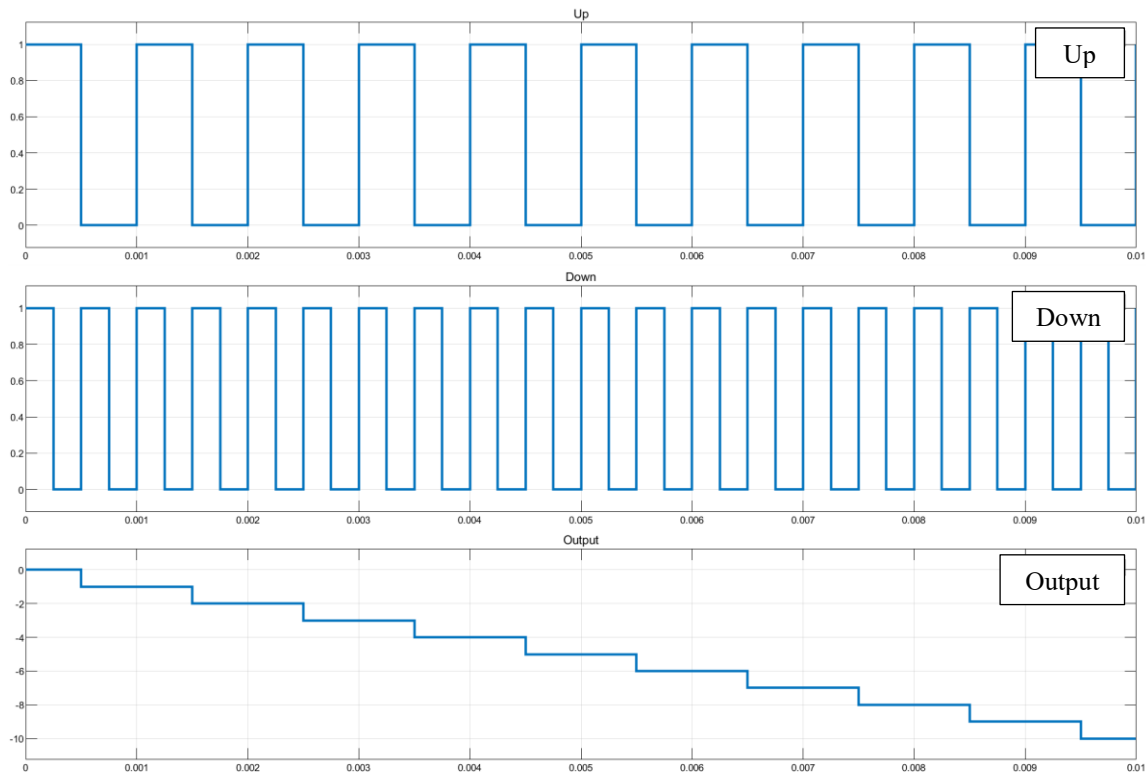


Figura 16 – Simulación del contador ideal

El hecho de que este contador sea ideal impide que desborde en algún momento. En el caso del filtro de la Figura 14, las frecuencias de pulsos en las entradas *Up* y *Down* son muy próximas, por lo que el contador real que se debe diseñar no requiere una gran capacidad, ya que los pulsos procedentes de los VCO se cancelan mutuamente.

3.3. Filtro bicuadrático diferencial con VCOs y contadores reales

3.3.1. Ventajas de la arquitectura diferencial

La arquitectura diferencial permite prescindir del VCO de referencia (aunque para ello se emplean 2 VCOs más). La arquitectura diferencial presenta varias ventajas:

- **Cancelación de los armónicos pares:**

Esta es una de las razones principales por las cuales en circuitos microelectrónicos se emplean arquitecturas diferenciales. Expresando la señal de salida de un amplificador *single-ended* en función de la entrada, desglosada en el conjunto de sus armónicos, se obtiene la siguiente expresión:

$$V_O = a_0 + a_1 \cdot V_i + a_2 \cdot V_i^2 + a_3 \cdot V_i^3 + a_4 \cdot V_i^4 + a_5 \cdot V_i^5 + \dots \quad (17)$$

En el caso de la arquitectura diferencial, las salidas correspondientes son:

$$V_{O+} = a_0 + a_1 \cdot V_i + a_2 \cdot V_i^2 + a_3 \cdot V_i^3 + a_4 \cdot V_i^4 + a_5 \cdot V_i^5 + \dots \quad (18)$$

$$V_{O-} = a_0 - a_1 \cdot V_i + a_2 \cdot V_i^2 - a_3 \cdot V_i^3 - a_4 \cdot V_i^4 + a_5 \cdot V_i^5 - \dots \quad (19)$$

En la salida negativa, solo los armónicos impares tienen signo distinto, ya que en los pares se cancela el signo. De esta manera, la salida diferencial resulta:

$$V_{DIF} = V_{O+} - V_{O-} = 2 \cdot a_1 \cdot V_i + 2 \cdot a_3 \cdot V_i^3 + 2 \cdot a_5 \cdot V_i^5 \quad (20)$$

Quedando así demostrada la cancelación de los armónicos pares, lo cual es una característica favorable de cara a indicadores como la Distorsión Armónica Total (THD).

- Inmunidad al ruido de la alimentación y PSRR

Otra razón fundamental por la cual se emplean circuitos diferenciales en microelectrónica es la mejora en la inmunidad ante ruidos en modo común, procedentes del exterior. En una arquitectura *single-ended*, cuando una interferencia se acopla a la señal de interés, esta se amplifica y perjudica el funcionamiento del sistema.

Sin embargo, cuando se emplea un circuito diferencial, el ruido en modo común aparece en ambas ramas con el mismo signo, de modo que al obtener la salida diferencial, esas interferencias, aun habiendo sido amplificadas previamente, se cancelan (en la Figura 17 se muestra el caso ideal).

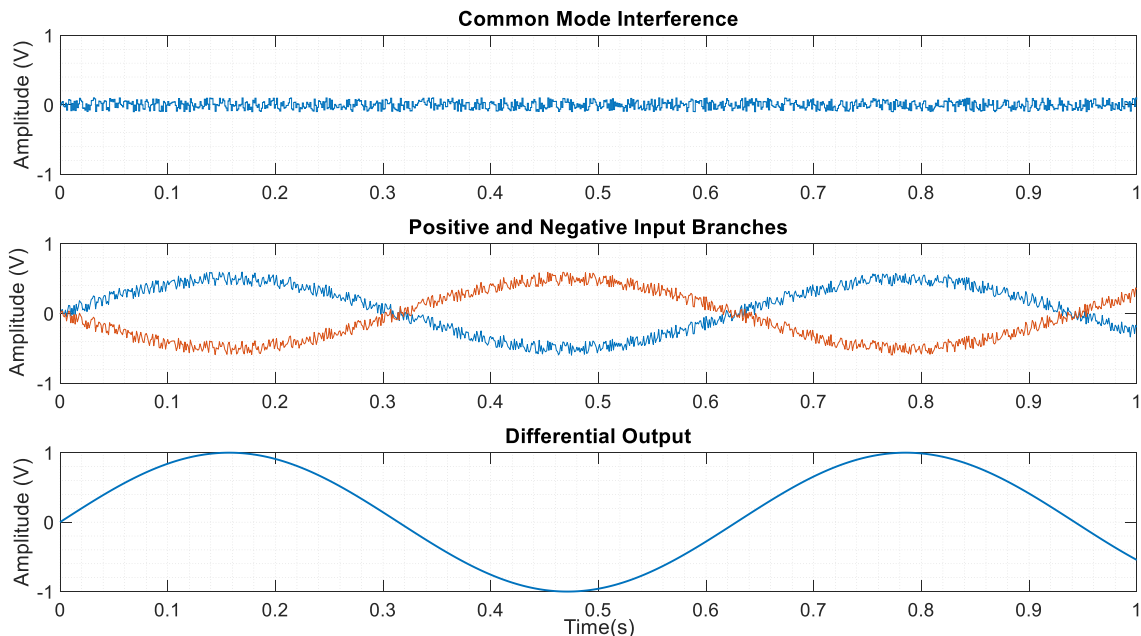


Figura 17 – Cancelación del ruido en modo común en arquitectura diferencial

En la realidad, las interferencias de modo común se atenúan considerablemente, aunque no del todo, debido a las ligeras diferencias que pueden existir en ambas ramas.

- **Mejora de la SNR**

El uso de una estructura diferencial también hace que la SNR (*Signal-to-Noise Ratio*) sea más alta. La SNR, o relación señal a ruido, es un parámetro que indica el ratio entre la potencia de la señal de interés y la potencia del ruido existente. De esta manera, cuanto mayor sea la SNR, mayor será el rango dinámico alcanzado, lo cual es deseable. La expresión de la SNR en el caso *single-ended* es:

$$SNR_{se} = 10 \cdot \log \left(\frac{\left(\frac{V_{in}}{\sqrt{2}} \right)^2}{V_{n-rms}^2} \right) = 10 \cdot \log \left(\frac{\frac{V_{in}^2}{2}}{V_{n-rms}^2} \right) \quad (21)$$

Mientras que para la estructura diferencial:

$$SNR_{dif} = 10 \cdot \log \left(\frac{4 \cdot \frac{V_{in}^2}{2}}{2 \cdot V_{n-rms}^2} \right) = 10 \cdot \log(2) + 10 \cdot \log \left(\frac{V_{in}^2}{V_{n-rms}^2} \right) = 3dB + SNR_{se} \quad (22)$$

La SNR aumenta 3 decibelios con respecto a la estructura *single-ended*.

- **Circuito de polarización del MEMS**

Por último, el uso de una arquitectura diferencial puede ser conveniente si el circuito de polarización del MEMS (*Micro-Electro-Mechanical System*), también es diferencial. El MEMS es el dispositivo transductor por excelencia empleado en los últimos años en aplicaciones móviles, debido a su reducido tamaño. En el caso de un micrófono MEMS, su estructura puede ser un condensador cuyas láminas metálicas se desplazan a causa de las vibraciones. De esta manera, la capacidad del condensador varía con las ondas de sonido. Aunque el MEMS puede generar una señal *single-ended*, generalmente los circuitos de polarización la transforman en diferencial para aprovechar las ventajas anteriormente presentadas.

3.3.2. Diseño del contador asíncrono real

Una vez expuestas las ventajas de un modelo diferencial, el siguiente paso es sustituir los contadores ideales por bloques que permitan la migración al modo diferencial, además de ser implementables en la realidad. Para ello, se propone una arquitectura basada en Flip-Flops de tipo D (Figura 18).

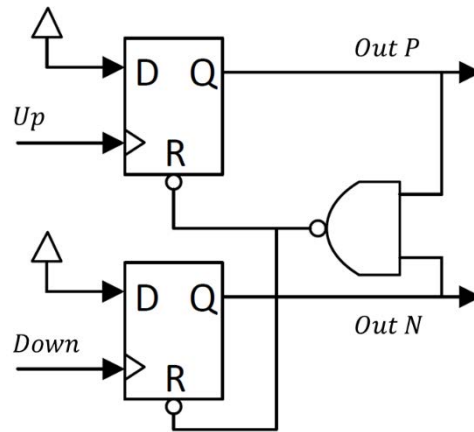


Figura 18 – Contador real de 2 entradas y 2 salidas

El contador dispone de 1 sólo bit. Sin embargo, como ya se ha indicado, las señales que proceden de los distintos VCO tienen frecuencias próximas, por lo que, salvo que la entrada supere cierto valor, el contador no desbordará.

Este contador asíncrono procede de la siguiente manera: Cada vez que la entrada positiva (Up) pasa a nivel alto, la salida positiva ($OutP$) pasa a nivel alto y la salida negativa ($OutN$) se resetea y queda a nivel bajo (o se mantiene ahí si ya lo estaba). Por el contrario, si el pulso llega por la entrada negativa ($Down$), la salida negativa ($OutN$) pasa a nivel alto y la salida positiva ($OutP$) es la que se resetea y pasa a nivel bajo).

Es importante destacar que, en este caso, el contador tiene 2 salidas, mientras que el ideal tan solo tenía 1. Esta cualidad facilita el diseño de una arquitectura diferencial como la que se muestra en la siguiente figura.

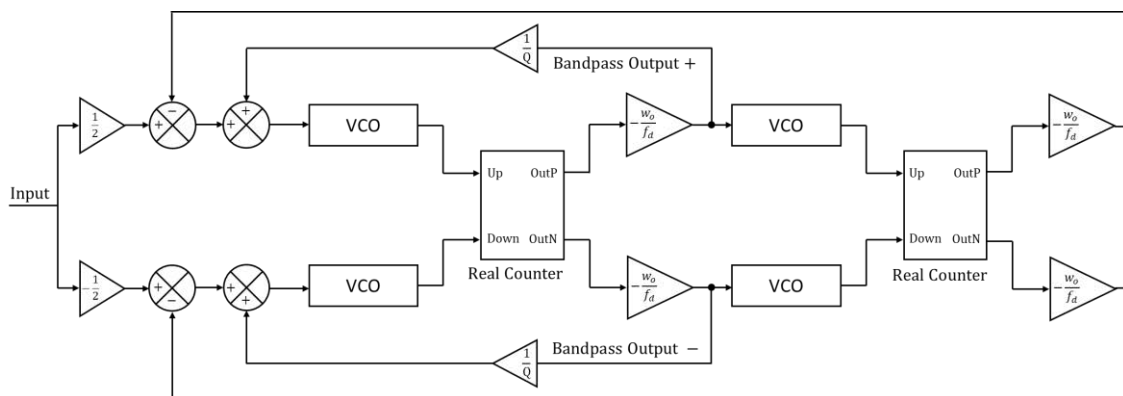


Figura 19 – Filtro con VCOs, diferencial y contadores reales

3.4. Filtro con VCO común y DCOs

Uno de los inconvenientes que presenta el modelo anterior es que, en el caso de un banco de N filtros, se requieren N VCOs (2N en el caso diferencial) que convierten la señal analógica de entrada en pulsos. Los osciladores internos son en realidad DCOs (*Digitally Controlled Oscillator*). La principal diferencia entre un VCO y un DCO es que, en el caso del primero, las tensiones de entrada que recibe son analógicas, mientras que en el caso del DCO, recibe una señal discretizada con un número limitado de niveles. Como consecuencia de esto, el tamaño, complejidad y consumo de un VCO es mayor que en el DCO.

Así pues, se ha diseñado una arquitectura que permite el uso de un solo VCO para todo el banco de filtros, que convierte la señal analógica en pulsos (Figura 20). El resto de osciladores internos de los filtros son DCOs.

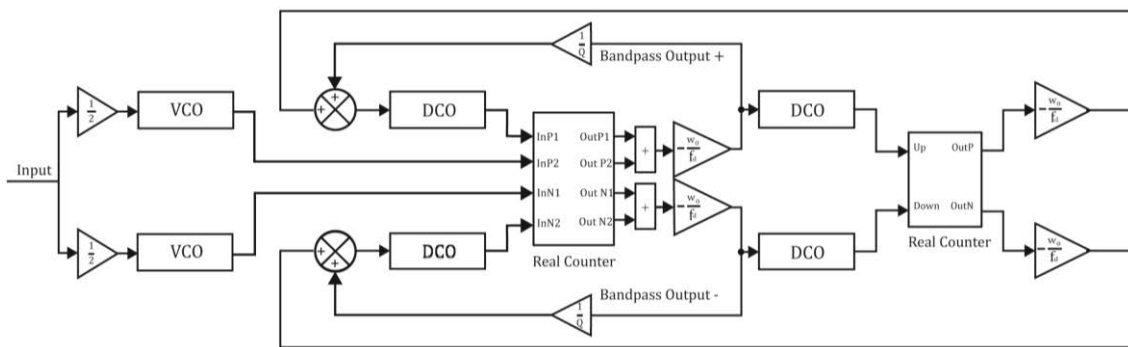


Figura 20 – Filtro con VCO común, DCOs, diferencial, y contadores ideales

El uso de esta arquitectura requiere un contador asíncrono más complejo que el de 2 entradas y 2 salidas, ya que recibe pulsos tanto de la señal de entrada como de los lazos de realimentación. La función de este contador es resetear el estado de las salidas complementarias. Es decir, cuando llega un pulso a cualquiera de las entradas negativas, una de las salidas positivas ha de pasar a nivel bajo (si las hubiere), y viceversa. En el caso de que llegasen 2 flancos positivos y 2 negativos, se cancelarían mutuamente. Las respectivas salidas positivas y negativas se suman, dando lugar a una sola salida positiva y una sola negativa. La segunda parte del filtro es idéntica al modelo anterior, por lo que el contador asíncrono simple es suficiente.

3.4.1. Diseño del contador asíncrono de múltiples entradas y salidas

La arquitectura propuesta para este contador asíncrono es una extensión del contador simple. Mediante 4 Flip-Flops y un bloque combinacional, las salidas se activan y desactivan atendiendo a las entradas correspondientes (Figura 21).

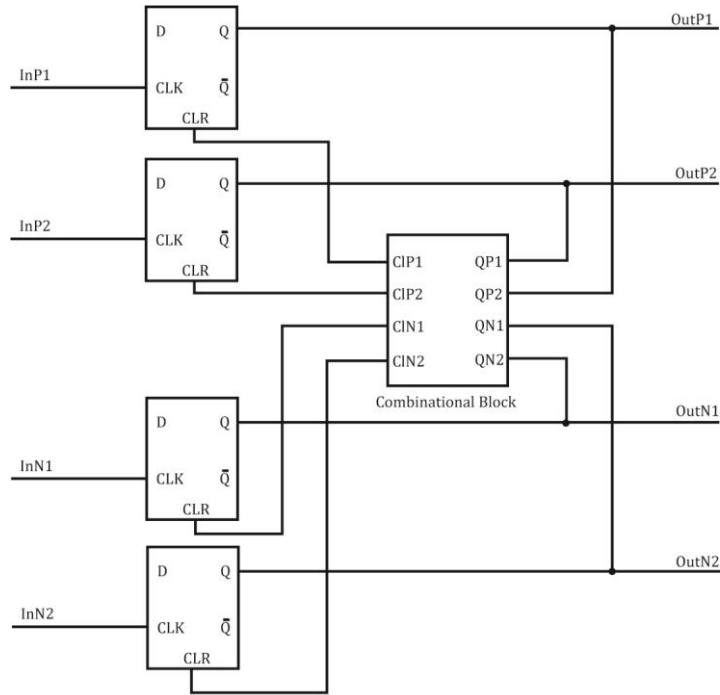


Figura 21 – Contador real de 4 entradas

El bloque combinacional es el encargado de mandar las señales de *reset* a los Flip-Flops. La tabla de verdad que define su comportamiento se muestra a continuación.

Q1N	Q2N	Q1P	Q2P	C1N	C2N	C1P	C2P
0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	1	0	0	0	0	0
0	0	1	1	0	0	0	0
0	1	0	0	0	1	0	1
0	1	0	1	0	1	1	0
0	1	1	0	0	1	1	0
0	1	1	1	0	1	0	1
1	0	0	0	0	0	0	0
1	0	0	1	1	0	0	1
1	0	1	0	1	0	1	0
1	0	1	1	1	0	1	0
1	1	0	0	0	0	0	0
1	1	0	1	0	1	0	1
1	1	1	0	1	0	1	0
1	1	1	1	1	1	1	1

Tabla 1 – Tabla de verdad del bloque combinacional

Existen diversas maneras de sintetizar una tabla de verdad en su circuito lógico correspondiente. En circuitos simples se puede recurrir a los métodos clásicos de simplificación, como el mapa de Karnaugh o diagrama de Veitch. Para casos más complejos, se recurre a herramientas de descripción hardware. Estas herramientas permiten sintetizar circuitos a partir de un código comportamental, de tal forma que el usuario no

necesita conocer los todos los detalles de bajo nivel. Esta opción es, en la mayoría de los casos, preferible, debido a la gran reducción de tiempo de diseño, en comparación con el método manual.

Para el diseño de este bloque combinacional, se ha utilizado el método manual, ya que la función a sintetizar no excesivamente compleja y el circuito resultante es el más óptimo en recursos, al haberse diseñado a bajo nivel. Las expresiones lógicas simplificadas que se han obtenido son las siguientes:

$$C_{1N} = Q_{1N} \cdot Q_{1P} + Q_{1N} \cdot \overline{Q_{2N}} \cdot Q_{2P}$$

$$C_{2N} = Q_{2N} \cdot Q_{2P} + \overline{Q_{1N}} \cdot Q_{2N} \cdot Q_{1P}$$

$$C_{1P} = Q_{1N} \cdot Q_{1P} + Q_{2N} \cdot Q_{1P} \cdot \overline{Q_{2P}}$$

$$C_{2P} = Q_{2N} \cdot Q_{2P} + Q_{1N} \cdot \overline{Q_{1P}} \cdot Q_{2P}$$

Para comprobar el correcto funcionamiento del circuito diseñado, se han realizado simulaciones funcionales del filtro completo (ver capítulo de resultados).

3.5. Filtro programable

Una de las ventajas de la codificación PFM que se realiza en los modelos propuestos es que, mediante el uso de bloques prescaladores (también conocidos como divisores de frecuencia), algunas propiedades del filtro pueden verse modificadas, entre ellas: la frecuencia natural o central del filtro. Si a la salida de cada VCO la frecuencia se divide entre 2, la frecuencia central del filtro también se ve reducida a la mitad. Si se divide entre 4, la frecuencia central será 4 veces menor que la original, y así sucesivamente. De esta manera, colocando un bloque prescalador detrás del VCO/DCO, es posible implementar un filtro programable por octavas. Una octava es un intervalo de dos frecuencias que están separadas por una relación 2:1. A diferencia de la década, el término octava está más vinculado al ámbito musical que al ingenieril. Teniendo en cuenta que el ámbito de aplicación de este filtro es el audio, las ventajas de un filtro programable por octavas son numerosas.

A continuación, se muestra el diseño propuesto para el divisor de frecuencias:

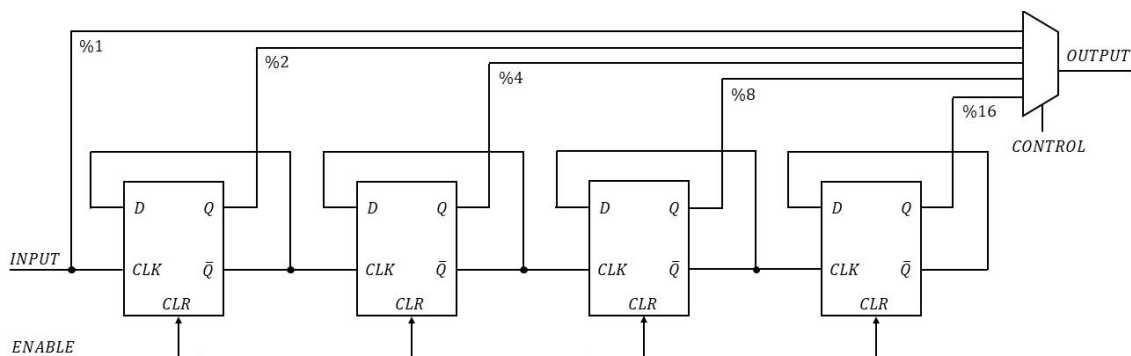


Figura 22 – Circuito divisor de frecuencias

La señal procedente del VCO entra en el primer biestable y a medida que la señal pasa por cada uno de ellos, su frecuencia se reduce a la mitad. Finalmente, todas las salidas de los biestables, junto con la del VCO, se introducen en un multiplexor, donde se selecciona la señal deseada. Como se puede apreciar, la complejidad y el área del circuito son mínimas, pero su versatilidad es máxima. Mientras que, en una arquitectura de filtro analógico convencional, para realizar un filtro programable serían necesarios múltiples condensadores (y el área que ello supone), en la arquitectura propuesta tan solo son necesarios varios Flip-Flops.

A continuación, se muestran los resultados de varias simulaciones (Figura 23), modificando en cada una de ellas el factor de prescalado. La frecuencia central inicial es de 2 kHz. Se observa que, a medida que el factor de prescalado aumenta, los productos de modulación (*sidebands*) se aproximan a la banda de audio. Esto es un fenómeno indeseado, ya que la información de interés, contenida en la banda de audio, se ve interferida por dichos armónicos. En la simulación llevada a cabo (Filtro 7 en los anexos), la frecuencia de oscilación de los VCOs es de 100kHz, por lo que al dividir la frecuencia entre 8, el armónico de menor frecuencia se encuentra en 12,5 kHz, dentro de la banda de audio, como se puede observar en la imagen.

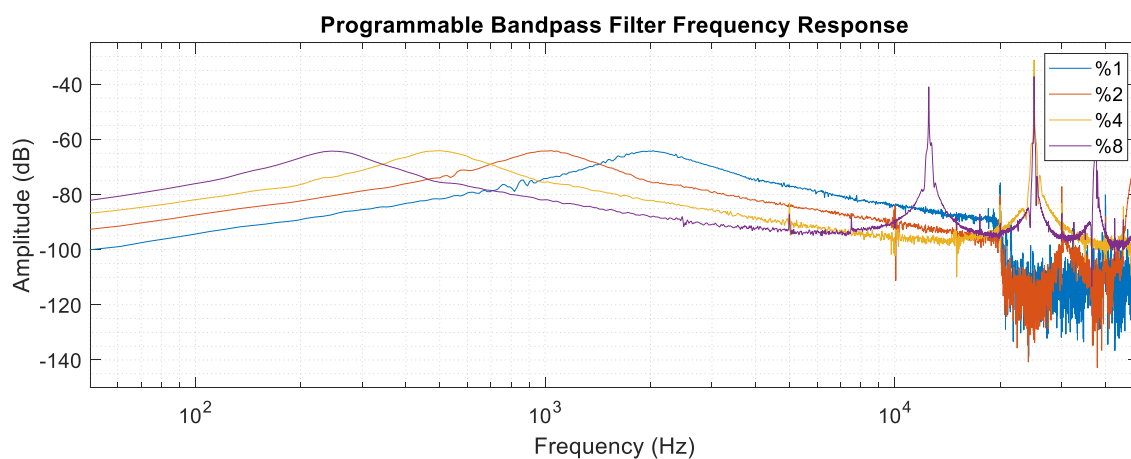


Figura 23 – Respuesta en frecuencia del filtro con distintos valores de prescalado

Para evitar este fenómeno, se puede incrementar la frecuencia de oscilación de los VCOs, alejando así el armónico de menor frecuencia de los 20 kHz. Si la frecuencia de oscilación es de 1MHz, se puede utilizar un factor de prescalado de hasta 32, aunque el aumento de la frecuencia de oscilación viene acompañado de un aumento del consumo del sistema.

3.6. Banco de filtros

Una vez diseñado el filtro, ya se puede formar un banco. Tal y como se advirtió en el capítulo 2, el número de filtros que debe contener el banco es arbitrario y admite múltiples respuestas en función de los requisitos del sistema. Generalmente se utilizan entre 10 y 20 filtros. Ocurre lo mismo con el espaciamiento entre cada filtro (es decir, entre las frecuencias centrales). Bien pueden espaciarse equitativamente en la escala de Hertz o en su escala logarítmica, o recurrir a las escalas perceptivas presentadas en el estado del arte: la escala Mel y la escala Bark.

Mediante la extracción de energía en cada uno de los filtros del banco se genera un vector de energías, que se actualiza en cada *frame*. Se obtiene así una imagen o patrón espectral de la muestra, que puede ser procesada por la red neuronal posteriormente. En el capítulo de resultados se muestran estas imágenes obtenidas con el filtro diseñado.

En la Figura 24 se aprecia la estructura del banco de filtros a nivel de sistema.

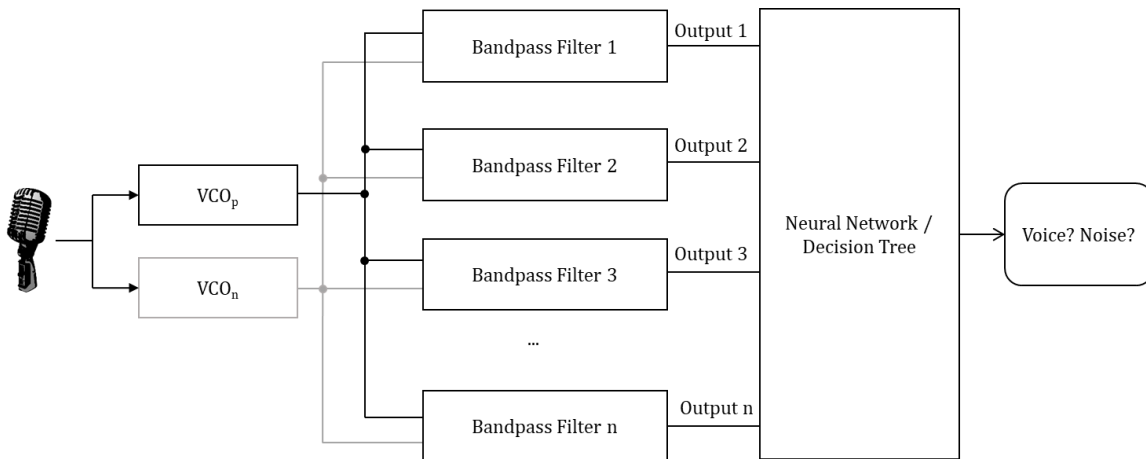


Figura 24 – Arquitectura del banco de filtros

CAPÍTULO 4

DISEÑO DE HARDWARE

4.1. Circuito con VCOs ideales

Una vez finalizado el diseño del sistema, se plantean algunos enfoques para implementar el circuito. Como primera aproximación, se ha realizado un modelo en la herramienta de diseño y simulación de circuitos *LTspice*. Para ello, se ha utilizado el bloque *modulator* (Figura 25). Este bloque posee dos entradas: FM y AM, y una salida Q. Además, tiene dos parámetros que definen su comportamiento: *mark* y *space*.

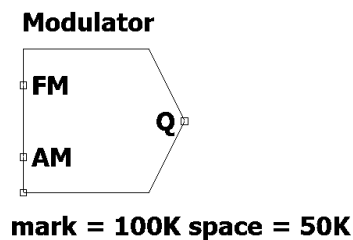


Figura 25 – Bloque modulador (LTspice)

Como su propio nombre indica, mediante este bloque se puede modular una señal, tanto en amplitud (AM) como en frecuencia (FM), dependiendo de qué entrada se utilice. Los términos *mark* y *space* proceden de la modulación por desplazamiento de frecuencia (FSK), donde *mark* es la frecuencia utilizada para definir el 1 lógico, y *space* la frecuencia para el 0 lógico.

Para que este bloque funcione igual que el VCO empleado en Simulink, la señal de entrada debe conectarse al terminal FM, y la entrada AM sólo debe recibir una constante, que corresponderá a la amplitud de la señal de salida Q. La expresión del VCO de Simulink es:

$$f_{VCO} = f_0 + K_{VCO} \cdot V_{IN} \quad (23)$$

Por lo que, ante una entrada de -1V, la frecuencia de salida ha de ser 50kHz, ante una entrada nula, la frecuencia debe ser la de reposo (100 kHz) y ante una entrada de 1V, de 150 kHz. La Figura 26 muestra el circuito implementado para comportarse como el VCO de Simulink.

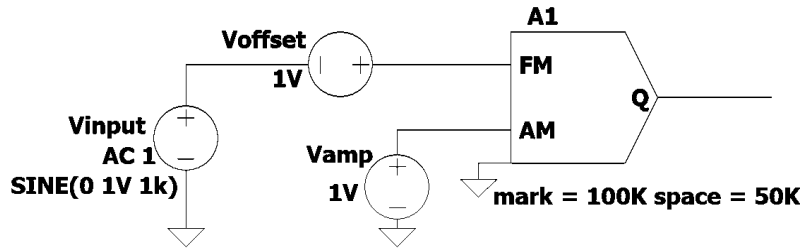


Figura 26 – Configuración del bloque modulador como VCO

Es necesaria la tensión de offset de 1V para que, ante una entrada de valor negativo, el modulador no se sature (ya que espera valores entre 0 y 1). A continuación, se muestra la simulación del circuito:

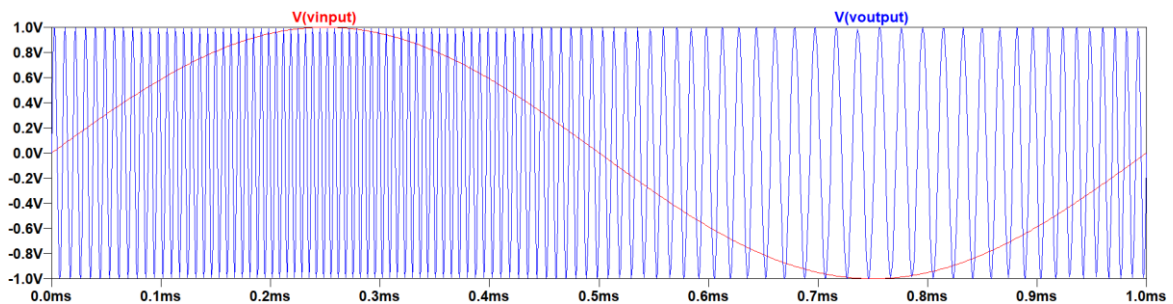


Figura 27 – Entrada y salida del modulador configurado como VCO

El esquemático de LTspice completo del filtro implementado con los moduladores se encuentra en los anexos (Filtro 8).

4.2. Circuito con VCOs reales

Una de las arquitecturas más empleadas para realizar un VCO es el oscilador en anillo. Un oscilador en anillo es una estructura formada por un número impar de inversores formando un lazo cerrado (Figura 28). Cada inversor, debido a sus capacidades parásitas intrínsecas, introduce un retraso en la propagación de la señal, es decir, un retraso cuando la señal pasa de nivel alto a nivel bajo (t_{PHL}) y cuando pasa de nivel bajo a nivel alto (t_{PLH}). Consecuentemente, cuando la señal ha atravesado 2 veces el lazo, se genera un periodo completo de oscilación, cuya frecuencia correspondiente es:

$$f_{VCO} = \frac{1}{N \cdot t_{DELAY}} = \frac{1}{N \cdot (t_{PHL} + t_{PLH})} \quad (24)$$

Donde N es el número de etapas del oscilador y t_{DELAY} es el retraso total introducido por cada inversor.

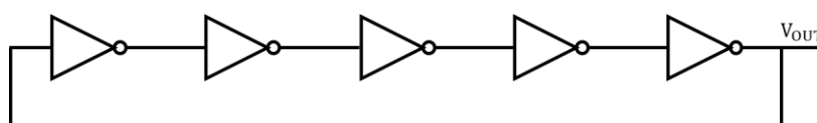


Figura 28 – Oscilador en anillo

El oscilador en anillo implementado tradicionalmente consta de una cadena de inversores CMOS. Para poder analizar su funcionamiento, primero se ha de introducir el inversor CMOS.

Inversor CMOS

El inversor CMOS se forma mediante un transistor PMOS y un NMOS (Figura 29).

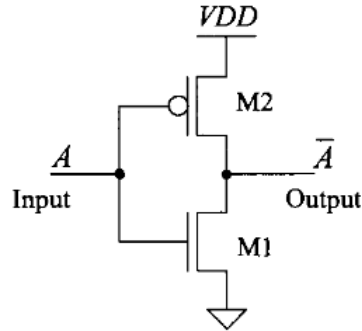


Figura 29 – Inversor CMOS [19]

Cuando la entrada se encuentra a nivel alto, el transistor M1, que es de tipo N, comienza a conducir, fijando así la salida a nivel bajo (GND). Por el contrario, cuando la entrada se encuentra a nivel bajo, es el transistor M2 el que entra en modo de conducción y fija la salida a nivel alto (VDD). En cada una de las transiciones de la entrada, los condensadores parásitos, que se encuentran en paralelo con los transistores (Figura 30), deben cargarse o descargarse. Dicho fenómeno es el causante de los tiempos de retraso mencionados anteriormente. Estos retrasos pueden estimarse mediante las siguientes expresiones:

$$t_{PLH} = 0.7 \cdot R_{p2} \cdot C_{out} \quad (25)$$

$$t_{PHL} = 0.7 \cdot R_{n1} \cdot C_{out} \quad (26)$$

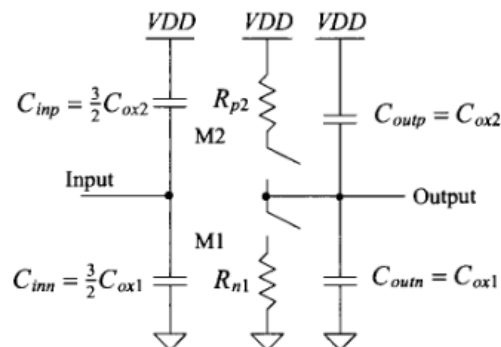


Figura 30 – Componentes parásitas del inversor CMOS [19]

No obstante, las expresiones anteriores tan sólo son una estimación teórica. La frecuencia de oscilación del oscilador se calculará posteriormente mediante simulación, la cual tiene en cuenta muchos más parámetros de los transistores (almacenados en sus librerías correspondientes).

Oscilador en anillo CMOS

Una vez presentada la estructura y comportamiento del inversor CMOS, formar el oscilador en anillo es tan sencillo como encadenar un número impar de estos (Figura 31).

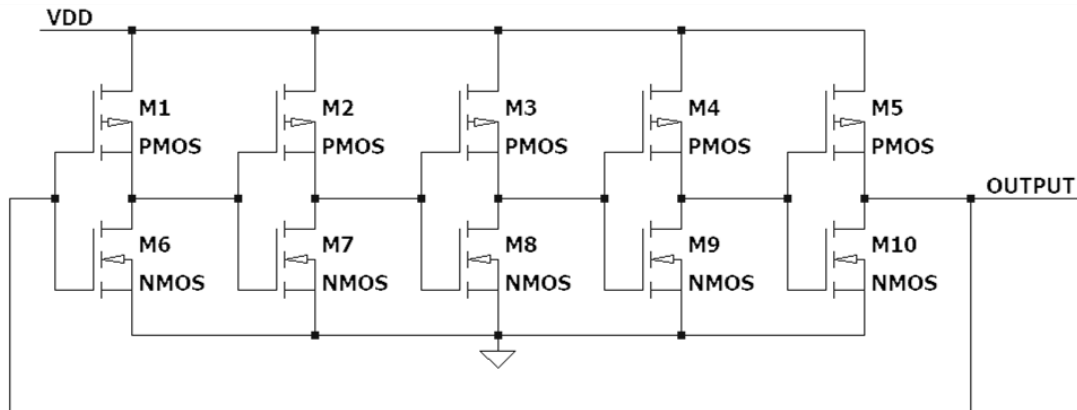


Figura 31 – Oscilador en anillo implementado en CMOS

La frecuencia de oscilación se puede variar de 3 formas:

- Número de transistores:

En este caso se ha elegido arbitrariamente una cadena de 5 transistores. Si se reduce el número de transistores, la frecuencia de oscilación aumenta, mientras que si el número de transistores es mayor, la frecuencia disminuye. Esto es una deducción directa de la expresión (24).

- Tamaño de los transistores:

Como ya se indicó anteriormente, la frecuencia de oscilación depende de las capacidades y resistencias parásitas. Consecuentemente, si se reduce el tamaño de los transistores, sus componentes parásitas también se verán reducidas, lo cual trae consigo un aumento en la frecuencia de oscilación.

- Alimentación:

Los métodos anteriores son permanentes. Es decir, una vez diseñado el circuito, ya no puede variarse ni el número ni el tamaño de los MOSFET. Por esta razón, el método utilizado para variar la frecuencia de oscilación es mediante la alimentación. Como se puede apreciar en la Figura 32, cuando aumenta la tensión de alimentación, aumenta la frecuencia de oscilación. Esto se debe, principalmente, a una reducción de los tiempos de retraso de cada inversor.

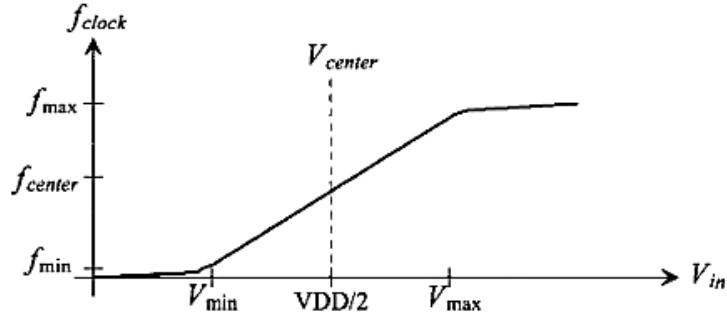


Figura 32 – Gráfica frecuencia-tensión de un oscilador en anillo [19]

Tal y como se desarrolla en [20], las expresiones de los tiempos de propagación son complejas, debido a la influencia de las componentes parásitas. En este artículo se deducen estos tiempos a partir de la expresión de la carga y descarga del condensador parásito que se sitúa en paralelo con la salida de cada inversor.

$$\tau_{dhl} = -C \int_{V_{DD}}^{V_{DD}-v_{TN}} \frac{dv_{out}}{I_{DNS}} - C \int_{V_{DD}-v_{TN}}^{V_{DD}/2} \frac{dv_{out}}{I_{DNL}} \quad (27)$$

$$\tau_{alh} = C \int_0^{-v_{TP}} \frac{dv_{out}}{I_{DPS}} + C \int_{-v_{TP}}^{V_{DD}/2} \frac{dv_{out}}{I_{DNL}} \quad (28)$$

De donde se deduce:

$$\tau_{dhl} = -\frac{C}{g_N(V_{DD} + v_{TN})} \left\{ \frac{2v_{TN}}{V_{DD} + v_{TN}} + \ln \left(\frac{3V_{DD} + 4v_{TN}}{V_{DD}} \right) \right\} \quad (29)$$

$$\tau_{alh} = -\frac{C}{g_P(V_{DD} + v_{TP})} \left\{ \frac{2v_{TP}}{V_{DD} + v_{TP}} + \ln \left(\frac{3V_{DD} + 4v_{TP}}{V_{DD}} \right) \right\} \quad (30)$$

Quedando así demostrado que los tiempos de propagación de los inversores dependen inversamente de la tensión de alimentación aplicada. Consecuentemente, un aumento en la tensión de alimentación supone un aumento en la frecuencia de oscilación.

Fuentes de corriente del oscilador en anillo

Como ya se vio en el capítulo anterior, el filtro diseñado hace uso de DCOs. Un DCO puede realizarse mediante un oscilador en anillo y un número finito de fuentes de corriente. Dependiendo de cuánta corriente se inyecte en el oscilador, la frecuencia resultante será una u otra. La Figura 33 muestra el esquemático de un DCO de 5 etapas y 3 fuentes de corriente (es decir, con 3 niveles de frecuencia).

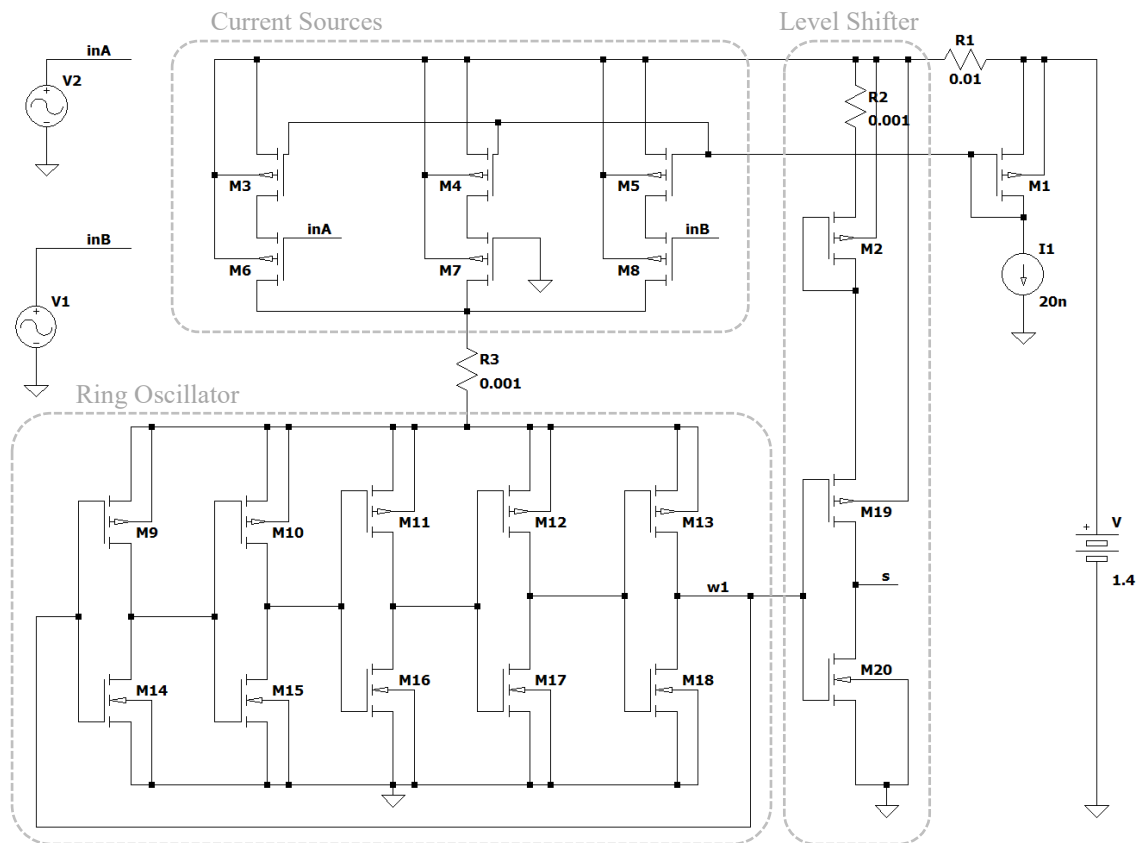


Figura 33 – DCO de 5 etapas con 3 posibles frecuencias

Los transistores M3-M8 conforman las 3 fuentes de corriente, encargadas de inyectar la corriente al oscilador. Una de ellas es permanente (M4 y M7), mientras que las otras dos se activan mediante las señales inA e inB. Cada fuente de corriente copia la referencia de corriente de M1, ya que forman una configuración de espejo de corriente. Los transistores M9-M18 conforman el oscilador en anillo, y su salida (w1) se conecta al elevador de nivel (*level shifter*). Esta estructura, conformada por los transistores M2, M19 y M20, se encarga de amplificar la señal oscilante de tal forma que su amplitud no dependa de la frecuencia (como sí ocurre a la salida del VCO, debido a efectos reales del circuito).

La Figura 34 muestra la simulación del circuito. En ella se pueden apreciar las 3 frecuencias. Dado que las fuentes de corrientes dependientes están formadas por transistores PMOS, se activan a nivel bajo. Así, cuando las señales inA e inB están a nivel bajo, todas las fuentes de corriente están activas y, consecuentemente, la frecuencia de oscilación es máxima. Cuando una de las señales de control está a nivel bajo y la otra está a nivel alto, la frecuencia resultante es intermedia. Por último, cuando ambas señales están a nivel alto, la frecuencia de salida es mínima.

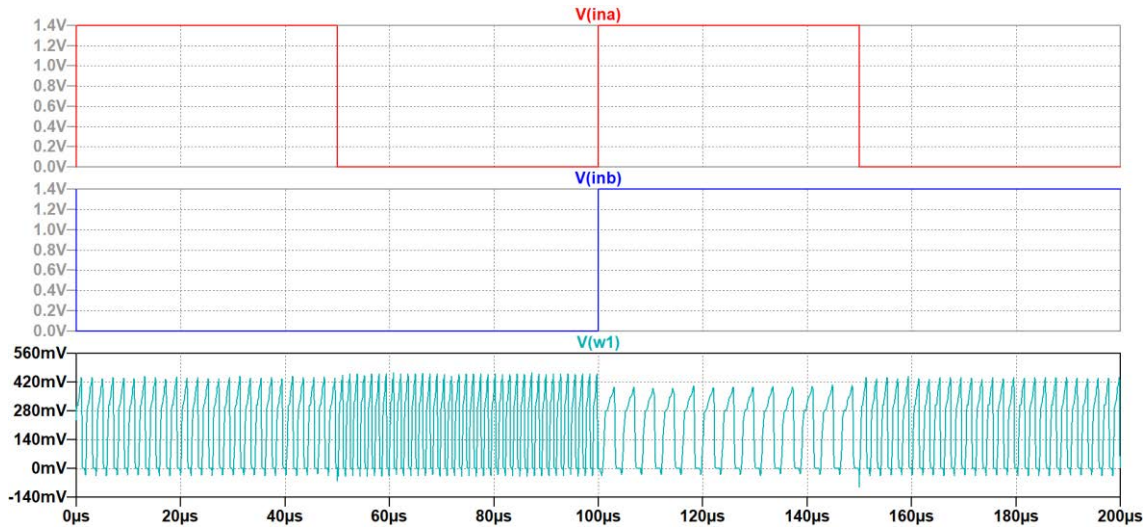


Figura 34 – Simulación del DCO de 5 etapas

La Figura 35 muestra la señal de entrada y de salida en el elevador de nivel, donde se puede apreciar la amplificación de la señal y el ajuste de la amplitud llevados a cabo.

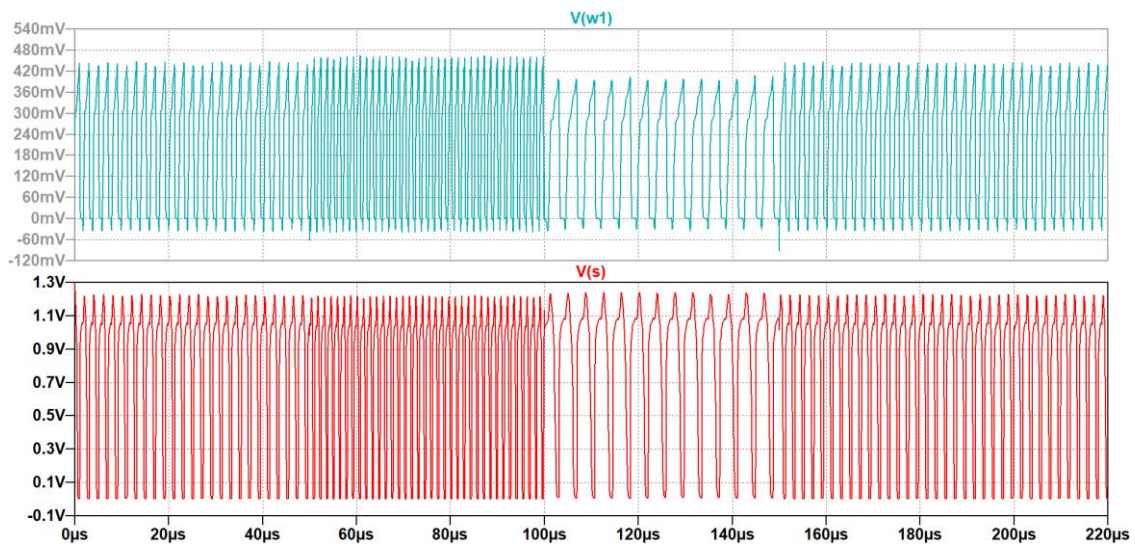


Figura 35 – Entrada y salida del Level Shifter

Otra manera de amplificar y además convertir la señal de salida en una señal cuadrada es mediante el uso del comparador de Schmitt. Ajustando el umbral en la mitad de la señal, se obtienen los siguientes resultados.

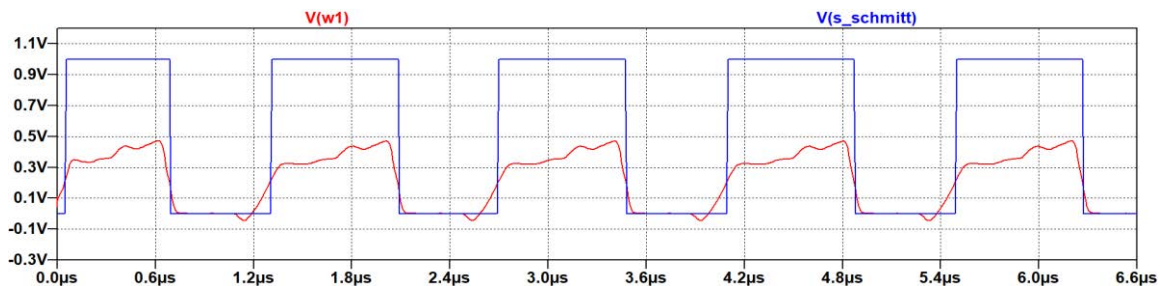


Figura 36 – Simulación del comparador de Schmitt conectado a la salida del oscilador

Una vez comprobado el diseño del DCO con transistores MOS, el siguiente paso es configurar sus valores de corriente para que el filtro tenga la frecuencia central donde se desee.

4.3. Caracterización del oscilador en anillo

Para poder elegir el valor de las fuentes de corriente que alimentarán el oscilador en anillo, es imprescindible caracterizarlo primero. Para ello, se han realizado simulaciones sobre el oscilador en anillo (Figura 37), inyectando cierta corriente y hallando la frecuencia de salida, obteniendo así el gráfico frecuencia-corriente que se muestra en la Figura 38.

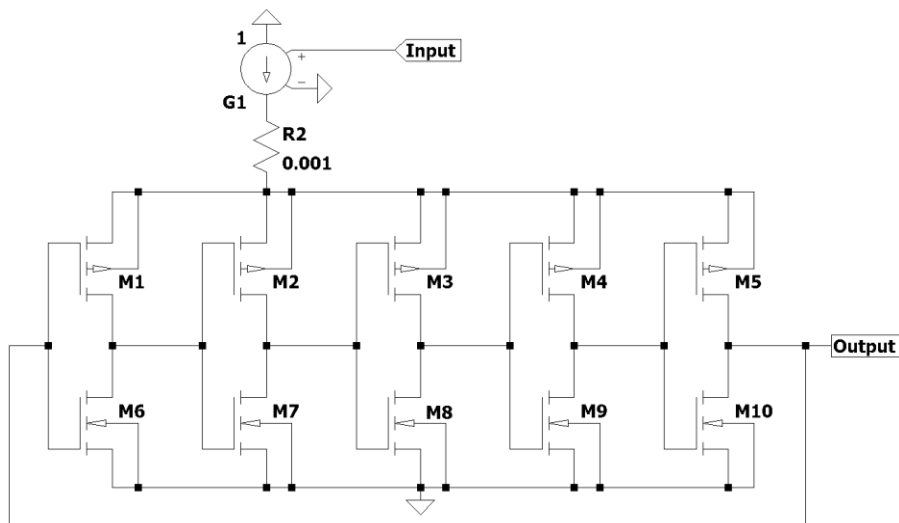


Figura 37 – Circuito de caracterización del VCO

A continuación, se muestran los resultados de simulación del VCO, entre 200kHz y 2MHz.

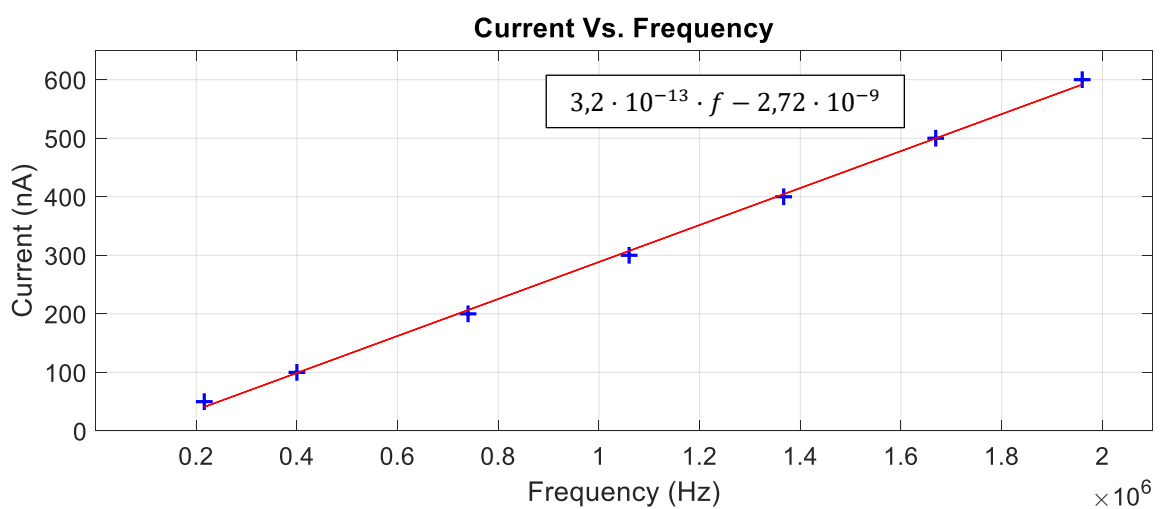


Figura 38 – Gráfica Frecuencia-Corriente (200kHz – 2MHz)

La ecuación exacta de la recta es:

$$I = 3,158071 \cdot 10^{-13} \cdot f - 2,725172 \cdot 10^{-9} \quad (31)$$

Mediante esta expresión, se puede conocer qué corriente ha de inyectarse para obtener una determinada frecuencia.

4.4. Diseño de los DCO

Tal y como se comentó en el capítulo 2, los osciladores internos de la arquitectura propuesta no son VCO sino DCO. A diferencia de los primeros, los DCO solo reciben determinados niveles de tensión (niveles cuantificados).

El DCO se puede alimentar con un número limitado de fuentes de corriente dependientes de tensión, para obtener las diferentes frecuencias. Para calcular el número de fuentes que necesita cada DCO y su valor, es necesario acudir de nuevo al diseño a nivel de sistema.

4.4.1. DCO 1

En el caso del primer DCO, en la Figura 39 se muestran los bloques previos mediante los cuales se puede calcular la expresión de la tensión de entrada V_{BP} del DCO en función de los bits OutP1, OutP2, OutN1 y OutN2.

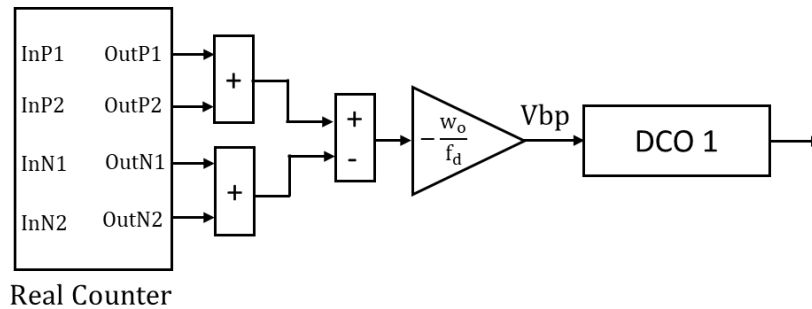


Figura 39 – Sección DCO 1 del filtro digital a nivel sistema

Dado que la expresión de la frecuencia de oscilación es:

$$f_{DCO1} = f_o + f_D \cdot V_{BP} \quad (32)$$

Donde V_{BP} es:

$$V_{BP} = -\frac{\omega_o}{f_d} \cdot (OutP1 + OutP2 - OutN1 - OutN2) \quad (33)$$

La expresión de la frecuencia del DCO es entonces:

$$f_{DCO1} = f_o - \frac{\omega_o}{f_d} \cdot f_d \cdot (OutP1 + OutP2 - OutN1 - OutN2) \quad (34)$$

$$f_{DCO1} = f_o - \omega_o \cdot (OutP1 + OutP2 - OutN1 - OutN2) \quad (35)$$

Teniendo en cuenta que, por la naturaleza del contador, no pueden coexistir salidas positivas y negativas a nivel alto, las posibles frecuencias a la salida del DCO son las mostradas en la siguiente tabla:

OUTP1	OUTP2	OUTN1	OUTN2	f_VCO
0	0	0	0	f_o
0	0	0	1	$f_o + \omega_o$
0	0	1	0	$f_o + \omega_o$
0	0	1	1	$f_o + 2\omega_o$
0	1	0	0	$f_o - \omega_o$
0	1	0	1	----
0	1	1	0	----
0	1	1	1	----
1	0	0	0	$f_o - \omega_o$
1	0	0	1	----
1	0	1	0	----
1	0	1	1	----
1	1	0	0	$f_o - 2\omega_o$
1	1	0	1	----
1	1	1	0	----
1	1	1	1	----

Tabla 2 – Frecuencias del DCO1

Es decir, en el circuito debe haber 5 fuentes de corriente para poder conseguir las 5 posibles frecuencias de oscilación:

$$f_o - 2\omega_o, \quad f_o - \omega_o, \quad f_o, \quad f_o + \omega_o, \quad f_o + 2\omega_o \quad (36)$$

En el caso de un filtro centrado en 1kHz y siendo la frecuencia de oscilación en reposo 500kHz, las frecuencias del DCO son:

$$f_{1_DCO1} = 500 \cdot 10^3 - 2 \cdot 2\pi \cdot 1000 = 487,430 \text{ kHz} \quad (37)$$

$$f_{2_DCO1} = 500 \cdot 10^3 - 2\pi \cdot 1000 = 493,720 \text{ kHz} \quad (38)$$

$$f_{3_DCO1} = 500 \cdot 10^3 = 500 \text{ kHz} \quad (39)$$

$$f_{4_DCO1} = 500 \cdot 10^3 + 2\pi \cdot 1000 = 506,280 \text{ kHz} \quad (40)$$

$$f_{5_DCO1} = 500 \cdot 10^3 + 2\pi \cdot 1000 = 512,570 \text{ kHz} \quad (41)$$

El siguiente paso es obtener, a partir de la caracterización del oscilador en anillo, los valores que deben tener las fuentes de corriente. Dada la ecuación (31), los valores de corriente que debe recibir el DCO 1 han de ser:

$$i_{1_DCO1}(487 \text{ kHz}) = 126,7 \text{ nA} \quad (42)$$

$$i_{2_DCO1}(494 \text{ kHz}) = 128,7 \text{ nA} \quad (43)$$

$$i_{3_DCO1}(500,00 \text{ kHz}) = 130,7 \text{ nA} \quad (44)$$

$$i_{4_DCO1}(506 \text{ kHz}) = 132,7 \text{ nA} \quad (45)$$

$$i_{5_DCO1}(513 \text{ kHz}) = 134,7 \text{ nA} \quad (46)$$

Para obtener esos valores se deberían emplear 4 fuentes de corriente dependientes de tensión, de 2nA (la diferencia entre cada nivel de corriente) y una fuente de corriente fija, de 31,73 nA. La Figura 40 muestra la disposición de las fuentes.

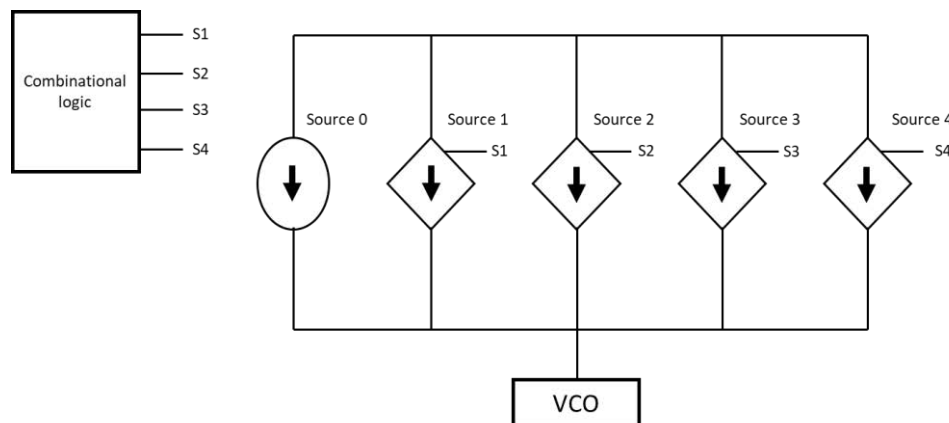


Figura 40 – Disposición de las fuentes de corriente en el DCO 1

Sin embargo, dado que los valores de las fuentes dependientes son muy pequeños e imposibles de implementar en la realidad, se ha de buscar una alternativa. Existen dos posibles soluciones a este problema:

- Modificar el tamaño de los transistores

Aumentando el tamaño de los transistores, la corriente que debe inyectarse es mayor para obtener los mismos valores de frecuencia. Consecuentemente, también aumenta el consumo del circuito. Por tanto, se trata de una labor de optimización, en la que se busca un *trade-off* entre consumo y funcionalidad del circuito. Esta tarea no es sencilla y excede los objetivos propuestos en este trabajo, de modo que no se ha llevado a cabo.

- Hacer uso del bloque prescalador

Como ya se describió en el capítulo 3.5, mediante el uso de un prescalador, la frecuencia central del filtro también se divide por el factor de prescalado. Esto se puede aprovechar para que los valores de corriente estén más distantes entre sí, ya que para obtener la misma frecuencia se ha de inyectar más corriente. A continuación se realiza el mismo proceso para calcular las fuentes de corriente, pero esta vez con un prescalado de 16. La frecuencia de oscilación se ha aumentado a 1 MHz para conseguir un mayor distanciamiento entre las fuentes.

$$f_{1_DCO1} = 10^6 - 2 \cdot 2\pi \cdot 16 \cdot 1000 \approx 798 \text{ kHz} \quad (47)$$

$$f_{2_DCO1} = 10^6 - 2\pi \cdot 16 \cdot 1000 \approx 900 \text{ kHz} \quad (48)$$

$$f_{3_DCO1} = 10^6 = 1 \text{ MHz} \quad (49)$$

$$f_{4_DCO1} = 10^6 + 2\pi \cdot 16 \cdot 1000 \approx 1,1 \text{ MHz} \quad (50)$$

$$f_{5_DCO1} = 10^6 + 2\pi \cdot 16 \cdot 1000 \approx 1,2 \text{ MHz} \quad (51)$$

Resultando en unas corrientes de:

$$i_{1_DCO1}(798 \text{ kHz}) = 225 \text{ nA} \quad (52)$$

$$i_{2_DCO1}(900 \text{ kHz}) = 257 \text{ nA} \quad (53)$$

$$i_{3_DCO1}(1 \text{ MHz}) = 289 \text{ nA} \quad (54)$$

$$i_{4_DCO1}(1,1 \text{ MHz}) = 320 \text{ nA} \quad (55)$$

$$i_{5_DCO1}(1,2 \text{ MHz}) = 352 \text{ nA} \quad (56)$$

Se observa que las diferencias entre los niveles de corriente son de un orden de magnitud superior. En este caso se emplearían **4 fuentes de corriente dependientes** de tensión, **de 32 nA** y una **fuentes de corriente fija de 225 nA**.

Una vez diseñadas las fuentes de corriente del DCO 1, es necesario un bloque combinatorial que active y desactive las fuentes de corriente dependiendo de los bits procedentes del contador asíncrono complejo (Figura 40).

Diseño del bloque combinacional del DCO 1

Para diseñar la lógica combinacional que dispare las fuentes de corriente, se elabora una tabla con las entradas posibles (que son las salidas OutP1, OutP2, OutN1 y OutN2) y las salidas esperadas (que son S1, S2, S3 y S4).

OUTP1	OUTP2	OUTN1	OUTN2	Frecuencia	S1	S2	S3	S4
0	0	0	0	f_0	1	1	0	0
0	0	0	1	$f_0 + \omega_0$	1	1	1	0
0	0	1	0	$f_0 + \omega_0$	1	1	1	0
0	0	1	1	$f_0 + 2 \cdot \omega_0$	1	1	1	1
0	1	0	0	$f_0 - \omega_0$	1	0	0	0
1	0	0	0	$f_0 - \omega_0$	1	0	0	0
1	1	0	0	$f_0 - 2 \cdot \omega_0$	0	0	0	0

Tabla 3 – Configuración de las fuentes de corriente para cada frecuencia posible (DCO1)

Mediante Karnaugh se extraen las expresiones lógicas para disparar cada una de las fuentes. Se han tenido en cuenta los casos prohibidos para usarse como casos neutrales (y así simplificar más las expresiones):

$$S_1 = \overline{OP1} + \overline{OP2} \quad (57)$$

$$S_2 = \overline{OP1} \cdot \overline{OP2} \quad (58)$$

$$S_3 = ON1 + ON2 \quad (59)$$

$$S_4 = ON1 \cdot ON2 \quad (60)$$

En la simulación que se muestra en la Figura 41 se puede comprobar que los valores de S1, S2, S3 y S4 corresponden con los de la Tabla 3.

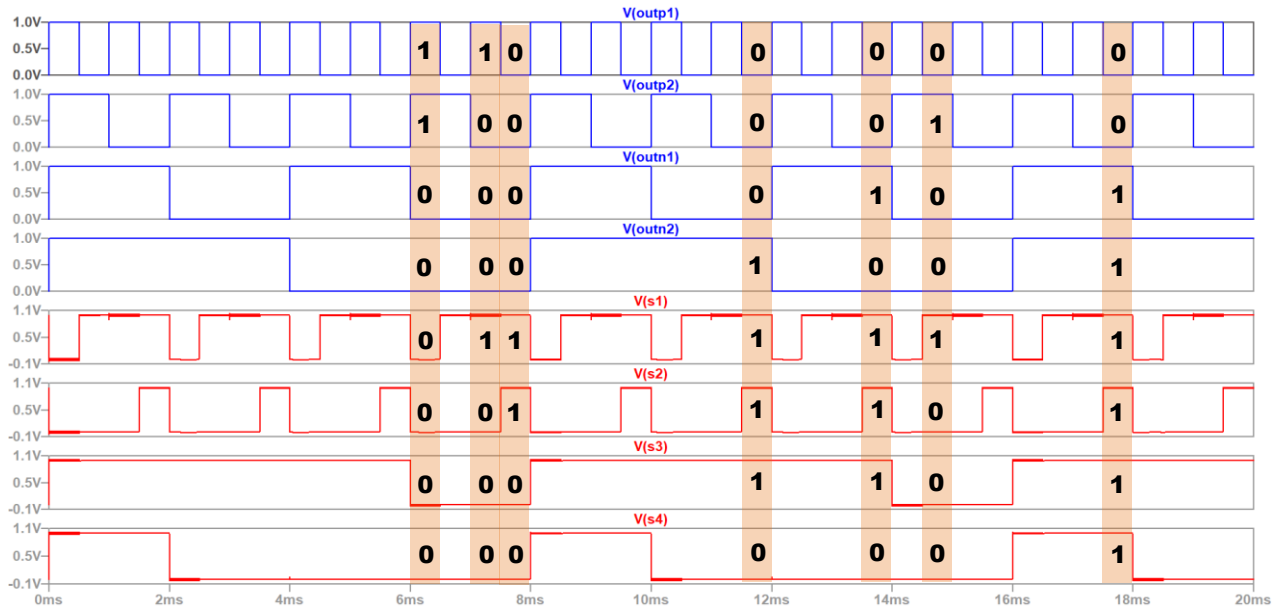


Figura 41 – Simulación de las fuentes del bloque combinacional del DCO 1

4.4.2. DCO 2

El proceso para calcular el circuito del segundo DCO es muy similar. En este caso, el DCO recibe información procedente de la salida paso banda (aplicándole la inversa del factor Q) y también de la salida paso bajo. Al tener más bits de entrada, se requerirán más fuentes de corriente.

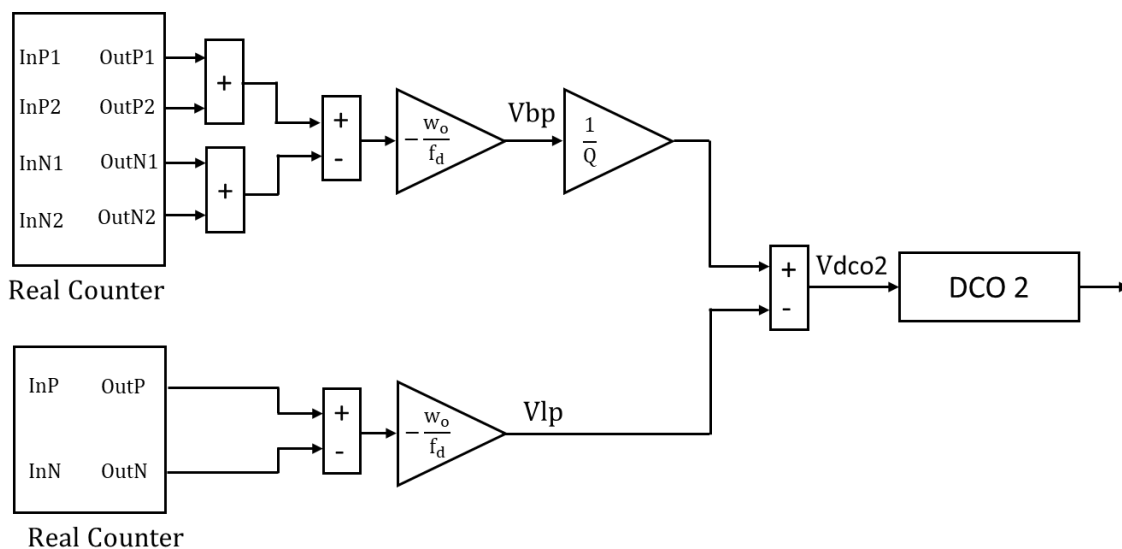


Figura 42 – Sección DCO 2 del filtro digital a nivel sistema

Dado que la expresión del oscilador es:

$$f_{VCO} = f_o + f_D \cdot V_{DCO2} \quad (61)$$

Donde V_{DCO2} es:

$$V_{DCO2} = -\frac{\omega_o}{Q \cdot f_d} \cdot (OutP1 + OutP2 - OutN1 - OutN2) + \frac{\omega_o}{f_d} \cdot (OutP - OutN) \quad (62)$$

La expresión de la frecuencia del DCO es entonces:

$$f_{VCO} = f_o - \frac{\omega_o}{Q} \cdot (OutP1 + OutP2 - OutN1 - OutN2) + \omega_o \cdot (OutP - OutN) \quad (63)$$

Considerando un factor $Q=1$, la expresión de la frecuencia se simplifica a:

$$f_{VCO} = f_o + \omega_o \cdot (OutP - OutN + OutN1 + OutN2 - OutP1 - OutP2) \quad (64)$$

Debido a que la expresión contiene 6 bits, el número de combinaciones asciende a 64. Por este motivo, sólo se han incluido en la siguiente tabla las combinaciones que pueden darse en la realidad. El resto de casos se han tomado como casos neutrales, usados favorablemente para simplificar las expresiones lógicas.

OUTP	OUTN	OUTP1	OUTP2	OUTN1	OUTN2	f VCO
0	0	0	0	0	0	f_o
0	0	0	0	0	1	$f_o + \omega_o$
0	0	0	0	1	0	$f_o + \omega_o$
0	0	0	0	1	1	$f_o + 2 \cdot \omega_o$
0	0	0	1	0	0	$f_o - \omega_o$
0	0	1	0	0	0	$f_o - \omega_o$
0	0	1	1	0	0	$f_o - 2 \cdot \omega_o$
0	1	0	0	0	0	$f_o - \omega_o$
0	1	0	0	0	1	f_o
0	1	0	0	1	0	f_o
0	1	0	0	1	1	$f_o + \omega_o$
0	1	0	1	0	0	$f_o - 2 \cdot \omega_o$
0	1	1	0	0	0	$f_o - 2 \cdot \omega_o$
0	1	1	1	0	0	$f_o - 3 \cdot \omega_o$
1	0	0	0	0	0	$f_o + \omega_o$
1	0	0	0	0	1	$f_o + 2 \cdot \omega_o$
1	0	0	0	1	0	$f_o + 2 \cdot \omega_o$
1	0	0	0	1	1	$f_o + 3 \cdot \omega_o$
1	0	0	1	0	0	f_o
1	0	1	0	0	0	f_o
1	0	1	1	0	0	$f_o - \omega_o$

Tabla 4 – Frecuencias del DCO2

Es decir, en el circuito debe haber 7 fuentes de corriente para poder conseguir las 7 posibles frecuencias de oscilación:

$$f_0 - 3\omega_0, \quad f_0 - 2\omega_0, \quad f_0 - \omega_0, \quad f_0, \quad f_0 + \omega_0, \quad f_0 + 2\omega_0, \quad f_0 + 3\omega_0 \quad (65)$$

En el caso de un filtro centrado en 1kHz, siendo la frecuencia de oscilación en reposo 1MHz y haciendo uso del prescalador en 16 (para evitar el mismo problema que en el DCO 1), las frecuencias a las que trabajará el DCO son:

$$f_{1_DCO2} = 10^6 - 3 \cdot 2\pi \cdot 16 \cdot 1000 \approx 698 \text{ kHz} \quad (66)$$

$$f_{2_DCO2} = 10^6 - 2 \cdot 2\pi \cdot 16 \cdot 1000 \approx 798 \text{ kHz} \quad (67)$$

$$f_{3_DCO2} = 10^6 - 2\pi \cdot 1000 \approx 900 \text{ kHz} \quad (68)$$

$$f_{4_DCO2} = 10^6 = 1 \text{ MHz} \quad (69)$$

$$f_{5_DCO2} = 10^6 + 2\pi \cdot 1000 \approx 1,1 \text{ MHz} \quad (70)$$

$$f_{6_DCO2} = 10^6 + 2\pi \cdot 1000 \approx 1,2 \text{ MHz} \quad (71)$$

$$f_{7_DCO2} = 10^6 + 3 \cdot 2\pi \cdot 1000 \approx 1,3 \text{ MHz} \quad (72)$$

El siguiente paso es obtener, mediante la caracterización del oscilador en anillo, los valores que deben tener las fuentes de corriente. Dada la ecuación (31), los valores de corriente que debe recibir el DCO 1 han de ser:

$$i_{1_DCO2}(698 \text{ kHz}) = 193 \text{ nA} \quad (73)$$

$$i_{2_DCO2}(798 \text{ kHz}) = 224 \text{ nA} \quad (74)$$

$$i_{3_DCO2}(900 \text{ kHz}) = 257 \text{ nA} \quad (75)$$

$$i_{4_DCO2}(1 \text{ MHz}) = 289 \text{ nA} \quad (76)$$

$$i_{5_DCO2}(1,1 \text{ MHz}) = 320 \text{ nA} \quad (77)$$

$$i_{6_DCO2}(1,2 \text{ MHz}) = 352 \text{ nA} \quad (78)$$

$$i_{7_DCO2}(1,3 \text{ MHz}) = 383 \text{ nA} \quad (79)$$

Para obtener esos valores se deberían emplear 6 fuentes dependientes de tensión, de 32 nA (la diferencia entre cada nivel de corriente) y una fuente de corriente fija, de 193 nA. La Figura 43 muestra la disposición de las fuentes.

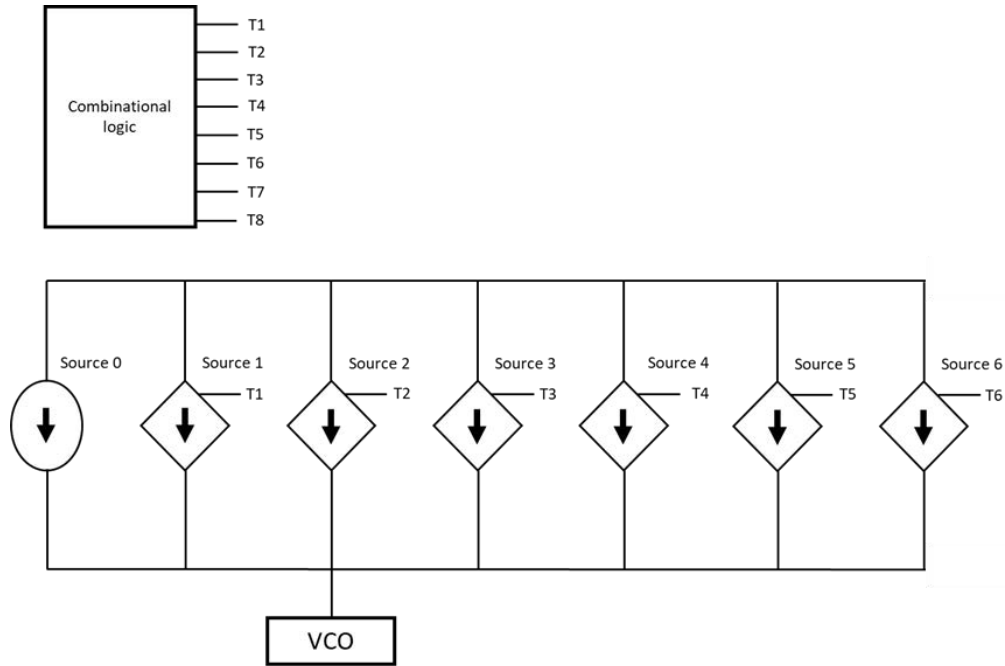


Figura 43 – Disposición de las fuentes de corriente en el DCO 2

Diseño del bloque combinacional del DCO 2

Al igual que para el caso anterior, se ha de diseñar el circuito combinacional que activa y desactiva las fuentes dependientes, con el objetivo de obtener las frecuencias adecuadas cuando se activan los bits correspondientes de salida de los contadores.

OUTP	OUTN	OUTP1	OUTP2	OUTN1	OUTN2	Frecuencia	t1	t2	t3	t4	t5	t6
0	0	0	0	0	0	f_0	1	1	1	0	0	0
0	0	0	0	0	1	$f_0 + \omega_0$	1	1	1	1	0	0
0	0	0	0	1	0	$f_0 + \omega_0$	1	1	1	1	0	0
0	0	0	0	1	1	$f_0 + 2 \cdot \omega_0$	1	1	1	1	1	0
0	0	0	1	0	0	$f_0 - \omega_0$	1	1	0	0	0	0
0	0	1	0	0	0	$f_0 - \omega_0$	1	1	0	0	0	0
0	0	1	1	0	0	$f_0 - 2 \cdot \omega_0$	1	0	0	0	0	0
0	1	0	0	0	0	$f_0 - \omega_0$	1	1	0	0	0	0
0	1	0	0	0	1	f_0	1	1	1	0	0	0
0	1	0	0	1	0	f_0	1	1	1	0	0	0
0	1	0	0	1	1	$f_0 + \omega_0$	1	1	1	1	0	0
0	1	0	1	0	0	$f_0 - 2 \cdot \omega_0$	1	0	0	0	0	0
0	1	1	0	0	0	$f_0 - 2 \cdot \omega_0$	1	0	0	0	0	0
0	1	1	1	0	0	$f_0 - 3 \cdot \omega_0$	0	0	0	0	0	0
1	0	0	0	0	0	$f_0 + \omega_0$	1	1	1	1	0	0
1	0	0	0	0	1	$f_0 + 2 \cdot \omega_0$	1	1	1	1	1	0
1	0	0	0	1	0	$f_0 + 2 \cdot \omega_0$	1	1	1	1	1	0
1	0	0	0	1	1	$f_0 + 3 \cdot \omega_0$	1	1	1	1	1	1
1	0	0	1	0	0	f_0	1	1	1	0	0	0
1	0	1	0	0	0	f_0	1	1	1	0	0	0
1	0	1	1	0	0	$f_0 - \omega_0$	1	1	0	0	0	0

Tabla 5 – Configuración de las fuentes de corriente para cada frecuencia posible (DCO2)

Al igual que en el caso del DCO 1, mediante el método de simplificación de Karnaugh se extraen las expresiones lógicas que forman el circuito combinacional.

$$t_1 = \overline{ON} + \overline{OP1} + \overline{OP2} \quad (80)$$

$$t_2 = OP + \overline{OP1} \cdot \overline{OP2} + \overline{ON} \cdot \overline{OP1} + \overline{ON} \cdot \overline{OP2} \quad (81)$$

$$t_3 = ON \cdot \overline{OP1} \cdot \overline{OP2} + ON1 + ON2 + OP \cdot \overline{OP1} + OP \cdot \overline{OP2} \quad (82)$$

$$t_4 = ON1 \cdot ON2 + OP \cdot \overline{OP1} \cdot \overline{OP2} + \overline{ON} \cdot ON1 + \overline{ON} \cdot ON2 \quad (83)$$

$$t_5 = \overline{ON} \cdot ON1 \cdot ON2 + OP \cdot ON2 + OP \cdot ON1 \quad (84)$$

$$t_6 = OP \cdot \overline{ON} \cdot \overline{OP1} \cdot \overline{OP2} \cdot ON1 \cdot ON2 \quad (85)$$

4.5. Estimación de consumo

Una vez diseñados los bloques, se ha de estimar el consumo de cada uno de ellos. Los bloques principales son: Los DCO (incluyendo sus fuentes de corriente), los bloques contadores y los bloques combinacionales que activan las fuentes de corriente.

Debido a que el software empleado para el diseño de los circuitos (LTspice) no dispone de modelos de puertas lógicas implementadas con transistores, estas no consumen corriente. Consecuentemente, sólo se ha estimado el consumo de los osciladores en anillo. Sin embargo, en el artículo elaborado sobre este trabajo (adjuntado en los anexos) se incluye una estimación de los 2 bloques contadores, realizada con el software de diseño de circuitos integrados: *Virtuoso Analog Design (Cadence)*.

Para calcular la potencia que consumen los DCO, se simulan todas las posibles combinaciones en sus fuentes de corriente y se extrae la corriente media consumida en cada caso. Por último, se halla el valor medio de todos ellos. La siguiente tabla muestra los resultados de corrientes en el DCO 1.

s1	s2	s3	s4	Corriente media (nA)
0	0	0	0	130,88
0	0	0	1	98,34
0	0	1	1	65,68
0	1	1	1	33,1
1	1	1	1	0,5
				65,7

Tabla 6 – Corrientes medias en el DCO 1

Para calcular la potencia que demanda, se multiplica la corriente media por la tensión de alimentación del circuito (1.2V).

$$P_{DCO1} = I_{avg1} \cdot V_{supply} = 65,7 \text{ nA} \cdot 1,2 \text{ V} = 78,4 \text{ nW} \quad (86)$$

La siguiente tabla muestra los resultados de corrientes en el DCO 2.

t1	t2	t3	t4	t5	t6	Corriente media (nA)
0	0	0	0	0	0	196,36
0	0	0	0	0	1	163,91
0	0	0	0	1	1	157,5
0	0	0	1	1	1	131,13
0	0	1	1	1	1	98,58
0	1	1	1	1	1	65,94
1	1	1	1	1	1	33,35
						118,16

Tabla 7 – Corrientes medias en el DCO 2

La potencia en el DCO 2 es:

$$P_{DCO2} = I_{avg2} \cdot V_{supply} = 118,16 \text{ nA} \cdot 1,2 \text{ V} = 141,79 \text{ nW} \quad (87)$$

Atendiendo a los resultados del artículo publicado (adjunto en los anexos), las corrientes medias consumidas por los contadores de 2 y 4 bits son 11,94 nA y 39,71 nA, respectivamente. Así pues, la potencia consumida aproximada un filtro es de:

$$P_{FILTER} = 2 \cdot 78,4 \text{ nW} + 2 \cdot 141,79 \text{ nW} + (39,71 \text{ nA} + 11,94 \text{ nA}) \cdot 1,2 \text{ V} \quad (88)$$

$$P_{FILTER} = 502,36 \text{ nW} \quad (89)$$

El consumo de potencia total de un banco de 16 filtros es de:

$$P_{FILTER_BANK} = 16 \cdot P_{FILTER} = 16 \cdot 502,36 \text{ nW} = 8,01 \text{ } \mu\text{W} \quad (90)$$

CAPÍTULO 5

RESULTADOS

5.1. Resultados de simulación del filtro

En este apartado se muestran los resultados de simulación de la arquitectura de filtro diseñada.

Simulación con entrada multi-senoidal

En primer lugar, se han introducido 3 señales sinusoidales distanciadas una década (100Hz, 1kHz y 10kHz). La Figura 44 muestra la señal de entrada y de salida del filtro en el dominio del tiempo (modelo single-ended con contadores reales y VCO común). El factor Q es 1 en todos los casos.

El filtro simulado corresponde al modelo 7 de los anexos.

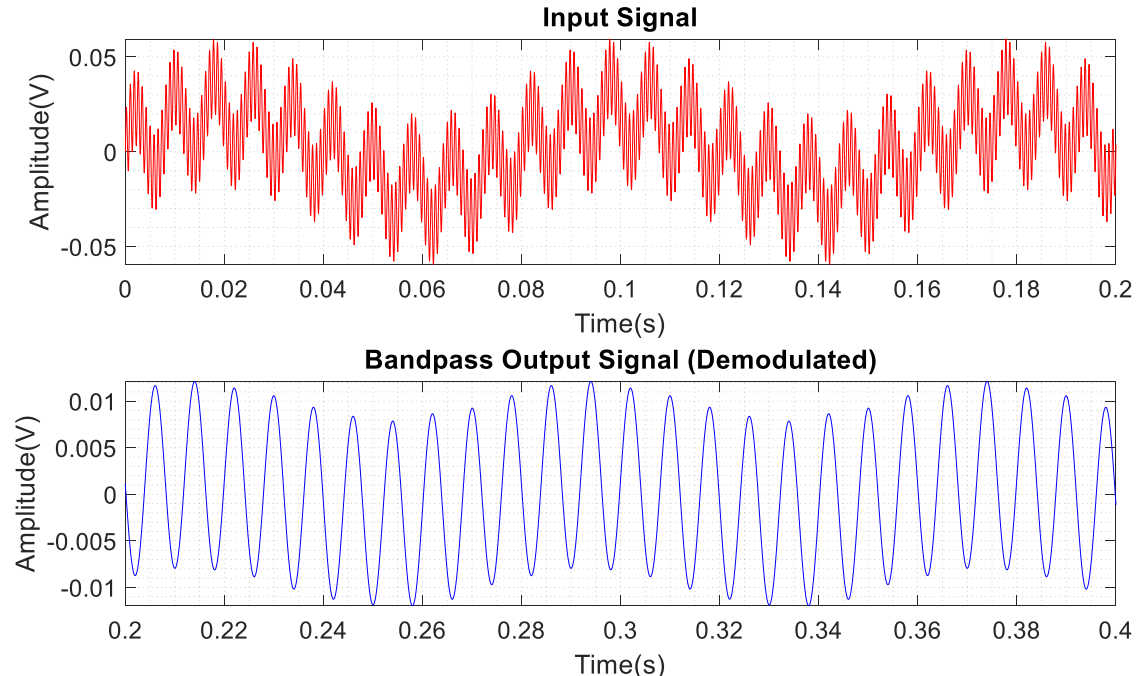


Figura 44 – Simulación temporal del filtro digital (Filtro 7, Simulink)

La Figura 45 muestra la simulación del filtro en LTspice. El filtro simulado corresponde al modelo 9 de los anexos.

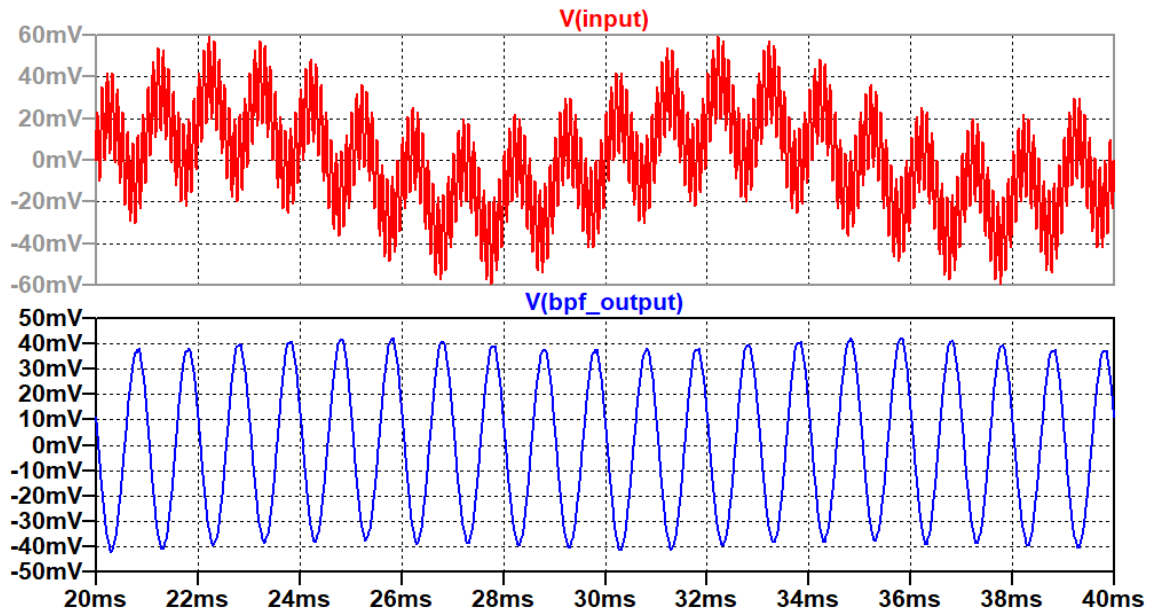


Figura 45 – Simulación temporal del filtro digital (Filtro 9, LTspice)

Los resultados obtenidos en ambas herramientas de simulación son prácticamente idénticos. A continuación, se muestran los resultados de la simulación en el dominio de la frecuencia, donde se pueden apreciar los tres tonos de igual amplitud en la señal de entrada. A la salida se percibe la atenuación de al menos 20 dB en el tono de baja y de alta frecuencia.

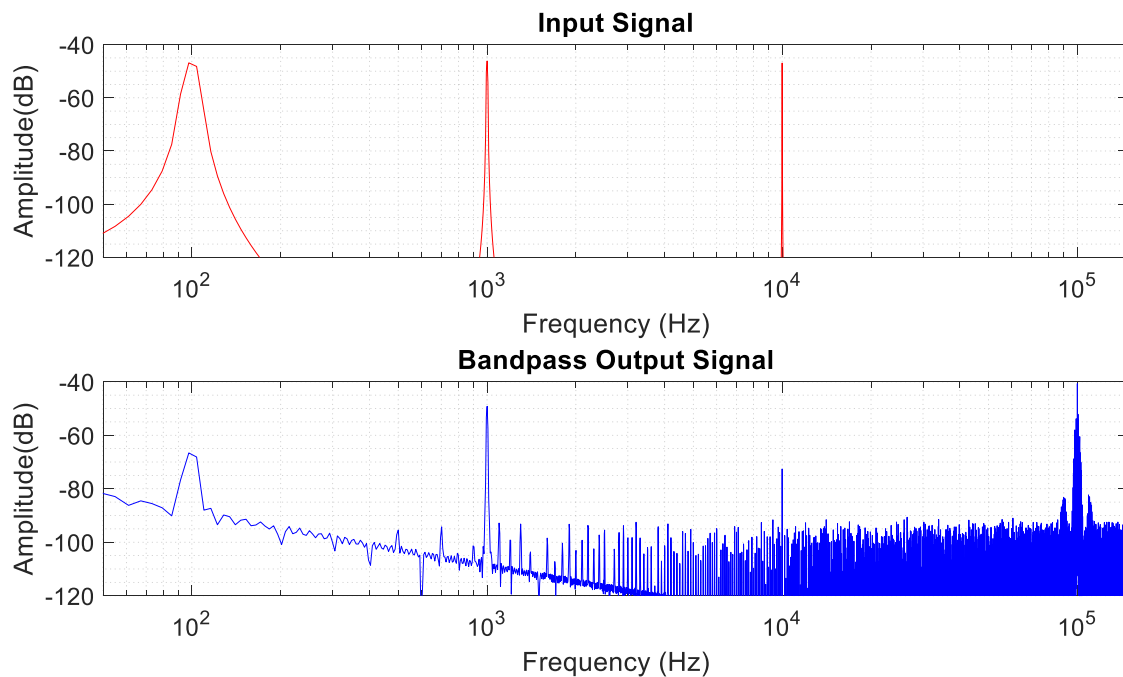


Figura 46 – Espectro de la señal de entrada y de salida del filtro (Filtro 7, Simulink)

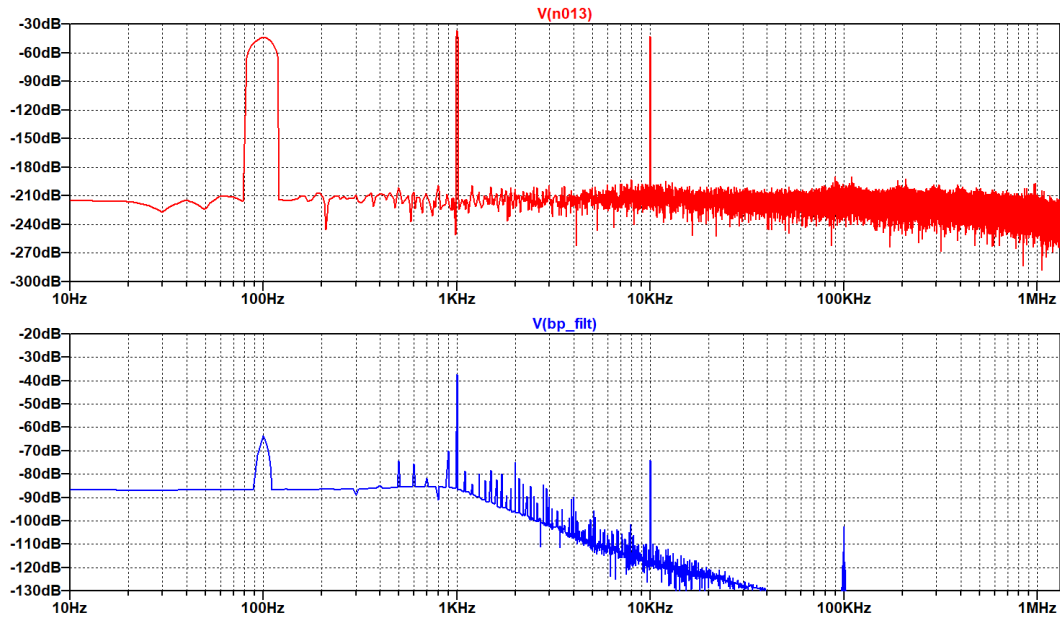


Figura 47 – Espectro de la señal de entrada y de salida del filtro (Filtro 9, LTspice)

Respuesta en frecuencia del filtro

La mejor manera de caracterizar el filtro es mediante su respuesta en frecuencia. La Figura 48 muestra la respuesta en frecuencia del filtro diferencial real con VCO común (Filtro 7 en los anexos) junto con la del filtro analógico inicial (Filtro 1 en los anexos). Para obtener esta respuesta, se ha empleado una señal seno hiperbólico que, tal y como se indicó anteriormente, posee una energía limitada a un ancho de banda determinado. Este ancho de banda se ha fijado en 22 kHz, que es la frecuencia máxima que percibe el oído humano.

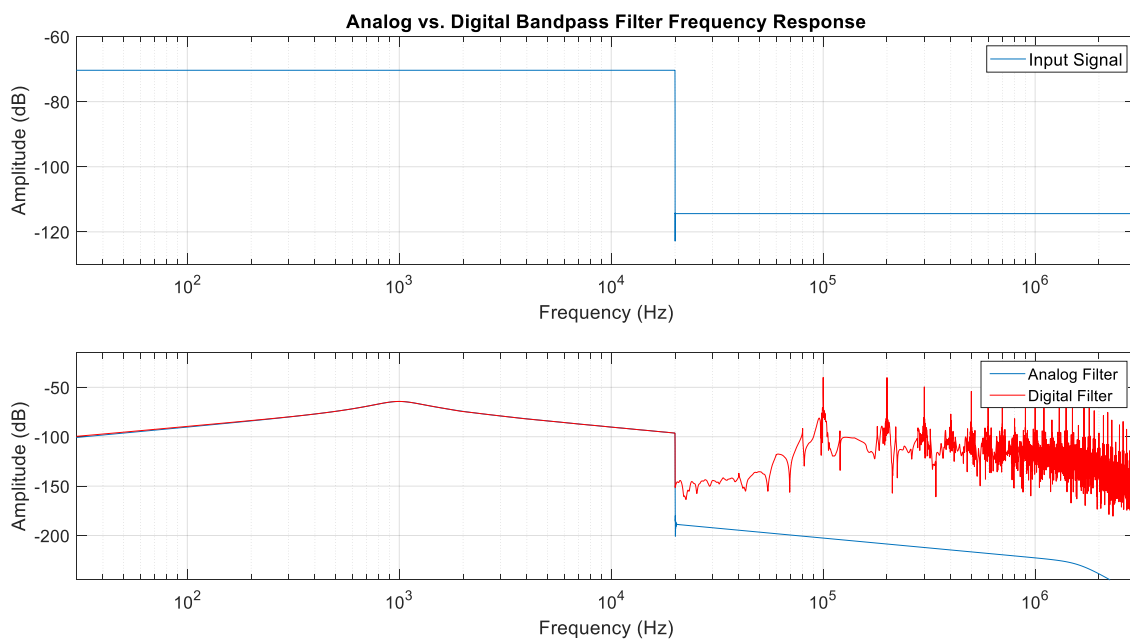


Figura 48 – Respuesta en frecuencia del filtro analógico y digital

En la figura se pueden apreciar las bandas laterales o *sidebands*. Se trata de productos de modulación que aparecen a frecuencias armónicas respecto a la de oscilación (en este caso, 100 kHz). En la banda de interés (200 Hz–22 kHz), la respuesta en frecuencia del modelo analógico ideal y del modelo digital coincide.

5.2. Extracción del patrón de energías

Como se indicó en el estado del arte, uno de los métodos más empleados para el reconocimiento del habla es transformar las muestras de audio en imágenes/patrones a partir de la energía en el dominio de la frecuencia y del tiempo, simultáneamente. Para ello, el siguiente paso tras filtrar la señal de entrada es calcular la energía en cada uno de los filtros del banco. La señal de entrada se divide en ventanas o *frames* de varios milisegundos (normalmente, entre 5 y 20 ms), y se extrae la energía en cada uno de ellos. De esta manera, si el banco tiene M filtros y el fragmento de audio tiene N *frames*, se forma una imagen de M filas y N columnas. En la Figura 49, la muestra de entrada es una señal de tipo chirp, que comienza a una frecuencia baja (cientos de hercios) y termina a 10kHz. En ese caso se ha utilizado el filtro analógico ideal. Se puede apreciar que la energía en los momentos iniciales se detecta en los filtros de menor frecuencia, y progresivamente se activan los de mayores frecuencias a medida que la señal de entrada varía. De esta manera, se puede representar una muestra de audio como una imagen, siendo intuitivo su comportamiento incluso a simple vista, en casos sencillos como una señal chirp. En la Figura 49 se puede ver el resultado a la salida de un banco de filtros analógicos (Filtro 1 en los anexos). Cada frame es de aproximadamente 5 ms. El espaciamiento entre filtros es de 500 Hz. No se ha indicado la escala exacta en la figura debido a que el aspecto relevante no son las unidades sino el patrón que forma la muestra.

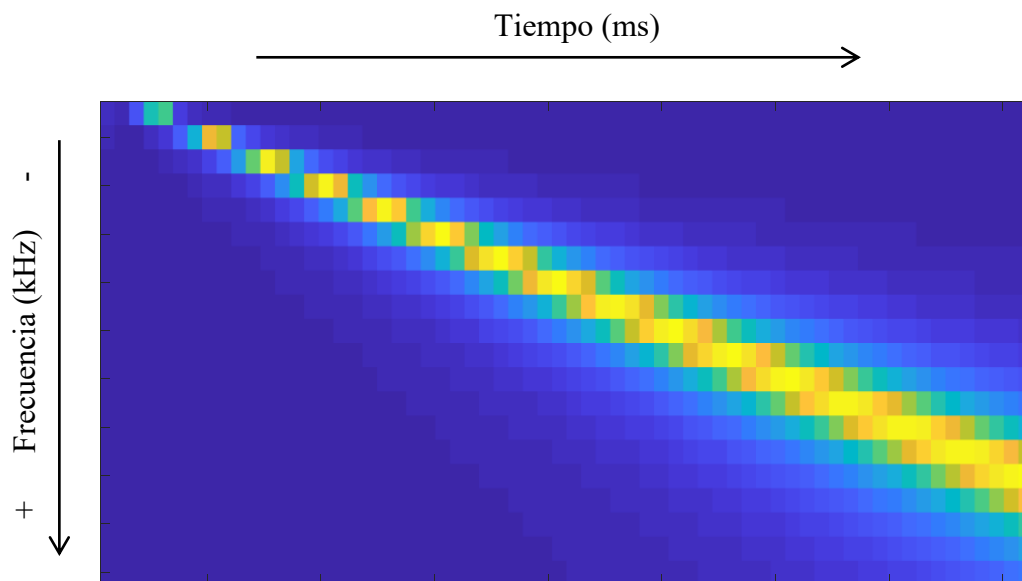


Figura 49 – Patrón de energías del banco de filtros (analógicos)

Una vez extraída esta imagen, el siguiente paso es normalizar el valor de cada elemento de la matriz y obtener así unos coeficientes con valores entre 0 y 1. Adicionalmente, se puede aplicar la transformada de coseno discreta (DCT), un proceso utilizado ampliamente en el ámbito de la compresión de imágenes que, a efectos prácticos, actúa como un filtro paso bajo, de tal forma que el peso de la imagen se ve considerablemente reducido, sin perder información. Tras ello, los coeficientes tratados se introducen en un árbol de decisión o una red neuronal que, tras ser entrenada, determina si en el fragmento de audio hay alguien hablando o no.

A continuación, se muestran los patrones extraídos para diversas señales de entrada en el filtro digital (Filtro 7 de los anexos). En este caso no se ha aplicado la transformada de coseno discreta. En primer lugar, se introduce una señal chirp.

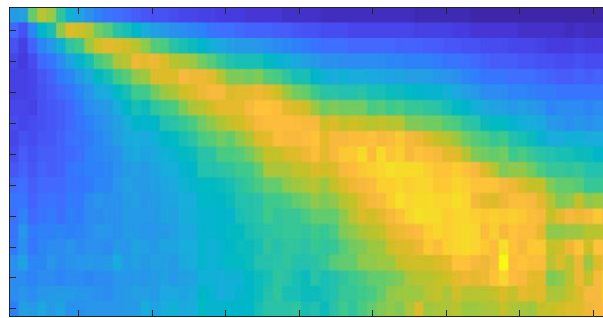


Figura 50 – Señal Chirp

En la figura 51 se observa que, a medida que la señal de entrada aumenta su frecuencia, más filtros se activan simultáneamente. Esto se debe a que, cuando la frecuencia central del filtro aumenta, también aumenta su ancho de banda, para mantener el factor Q constante.

A continuación, se introduce una señal senoidal a frecuencia constante.

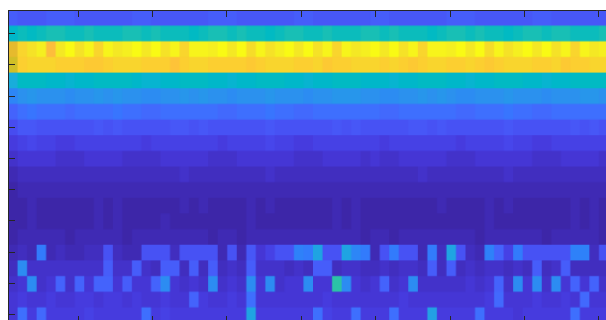


Figura 51 – Señal Senoidal

Se observa cierta actividad a altas frecuencias. Esto puede deberse a efectos de modulación (*aliasing*) del filtro. Este efecto no deseado puede solventarse mediante tareas de post-procesado (DCT).

Por último se introduce una señal aleatoria.

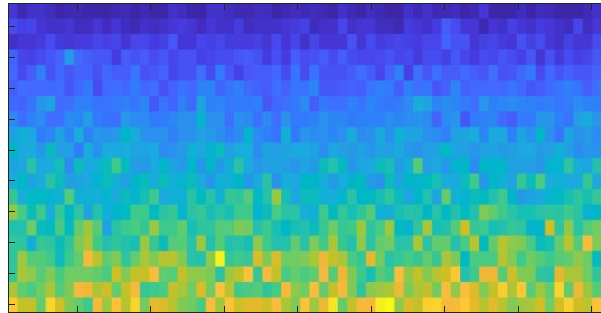


Figura 52 – Señal Aleatoria

Se observa mayor actividad a altas frecuencias que a bajas. Esto se debe al efecto del ancho de banda indicado anteriormente. Dado que en las frecuencias más altas los filtros tienen mayor ancho de banda, estos reciben más energía en promedio que los filtros de baja frecuencia, por lo que se activan con mayor facilidad. Este fenómeno puede resolverse mediante el uso de distintos factores Q en el banco de filtros, o bien mediante un espaciamiento entre filtros más adecuado (Mel, Bark).

Una vez presentados los resultados, se puede concluir que la arquitectura de filtros desarrollada funciona correctamente y los resultados de los patrones de energía son satisfactorios.

5.3. Consumo y área

El objetivo de este trabajo era desarrollar una arquitectura de filtros paso banda que permitiera un ahorro tanto de energía consumida como de área ocupada, en la medida de lo posible. En el capítulo 4 se presentó una primera estimación del consumo de un banco de 16 filtros, obteniendo una potencia consumida de aproximadamente $8 \mu\text{W}$. Se trata de un consumo comparable al de las soluciones analógicas, pero con un área ocupada muy inferior (ya que estas soluciones requieren del uso de condensadores de tamaño considerable).

A pesar de que no se ha podido estimar el área ocupada del sistema (debido a la herramienta de software utilizada), en la siguiente tabla se muestra la comparación de soluciones equivalentes, presentada en el artículo realizado sobre este trabajo (adjuntado en los anexos). Debido a que en este artículo se emplearon transistores de menor tamaño, la potencia estimada del circuito es 8 veces inferior.

	Potencia (uW)	Número de canales	Area (mm2)	Enfoque	ADC incluido
Nuestra solución	1.01	16	0.03	Hibrido (VCO)	Sí
[21]	0.06	16-48	0.73	Hibrido	No
[22]	6	16	2	Analógico	No
[4]	0.38	16	0.13	Analógico	No
[23]	>50	-	-	Digital	No

Tabla 8 – Comparación de soluciones

CAPÍTULO 6

PRESUPUESTO Y PLAN DE TRABAJO

6.1. Presupuesto

A continuación se muestra el desglose del presupuesto para el desarrollo del proyecto. Dado que este trabajo es de carácter teórico y de investigación, no se incluye un presupuesto de coste del prototipo. Se han considerado los costes de personal y de licencias.

Unidad	Descripción	Medición	Precio Unitario (€/Ud.)	Precio total (€)
h	Ingeniero jefe de proyecto Coordinación del proyecto	25	50	1500
h	Ingeniero proyectista Análisis y diseño del sistema	150	30	4500
h	Ingeniero proyectista Generación de la documentación	100	20	2000
				8000

Tabla 9 – Coste de personal

Unidad	Descripción	Medición	Precio Unitario (€/Ud.)	Precio total(€)
año	Licencia MATLAB 2018	1	800	800
año	Licencia LTSpice	1	0	0
				800

Tabla 10 – Coste de licencias

Así pues, el coste total de este proyecto se estima en **8800 €**.

6.2. Plan de trabajo (diagrama de Gantt)

Adicionalmente, se muestra el diagrama de Gantt donde se refleja el tiempo de dedicación para las tareas que han conformado este proyecto.

	OCTUBRE	NOVIEMBRE	DICIEMBRE	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO
FASE DE BÚSQUEDA											
PLANTEAMIENTO DEL PROBLEMA Y SOLUCIONES											
BÚSQUEDA DE FUENTES											
DISEÑO DEL SISTEMA											
DISEÑO A NIVEL SISTEMA DE MODELOS IDEALES											
DISEÑO A NIVEL SISTEMA DE MODELOS REALES											
DISEÑO HARDWARE											
ANÁLISIS DE LOS RESULTADOS											
ESTIMACIONES DE CONSUMO											
ANÁLISIS DE LOS MODELOS DISEÑADOS											
DOCUMENTACIÓN DEL PROYECTO											

CAPÍTULO 7

CONCLUSIONES Y LINEAS FUTURAS DE INVESTIGACIÓN

En este capítulo se enumeran las conclusiones extraídas a partir de los resultados presentados. Posteriormente, se incluyen posibles líneas futuras de investigación acerca de este proyecto.

7.1. Conclusiones y objetivos cumplidos

- Se ha presentado el estado del arte acerca de la estructura del oído humano, los principales métodos de extracción de características vocales para el reconocimiento del habla y, por último, el uso de VCOs para el procesamiento de audio así como sus ventajas respecto a los métodos tradicionales.
- Se ha estudiado e implementado, a nivel de sistema, una novedosa arquitectura de filtros paso banda apoyada en el uso de VCOs. Para ello, se ha realizado una migración progresiva desde el modelo analógico clásico hasta el modelo digital programable.
- Se han expuesto las propiedades principales de los osciladores en anillo y su implementación para formar los DCOs del filtro paso banda digital. Se han realizado simulaciones que verifican su funcionamiento. Se ha estimado el consumo de los elementos principales del circuito.
- Se han presentado y comparado los resultados de las principales simulaciones en Simulink y LTspice. Se han extraído las imágenes procedentes de la extracción de energías del banco de filtros, emulando así el sistema empleado en los sistemas de detección de eventos presentados en el estado del arte. También se ha estimado el consumo total del banco de filtros.
- Por último, se ha presentado un artículo con los principales conceptos y resultados de este trabajo en el *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, el cual ha sido aceptado y expuesto en Texas (USA) entre el 4 y 7 de agosto de 2019.

7.2. Líneas futuras de investigación

- Profundizar en los métodos de extracción de características vocales para definir los parámetros del banco de filtros (número de filtros, distribución de frecuencias, factor Q empleado, etc.), con el objetivo de facilitar y optimizar las tareas de extracción de características vocales.
- Finalizar el diseño de circuito del filtro digital real y realizar las simulaciones funcionales correspondientes sobre el circuito completo.
- Implementar el circuito en el software *Virtuoso Analog Design (Cadence)* para realizar una estimación del consumo de los bloques combinacionales.
- Realizar un estudio del ruido de fase en el circuito oscilador en anillo, con la finalidad de optimizar el tamaño de los transistores, manteniendo la funcionalidad del filtro.
- Una vez implementado y optimizado el circuito, realizar el diseño de su *layout*.

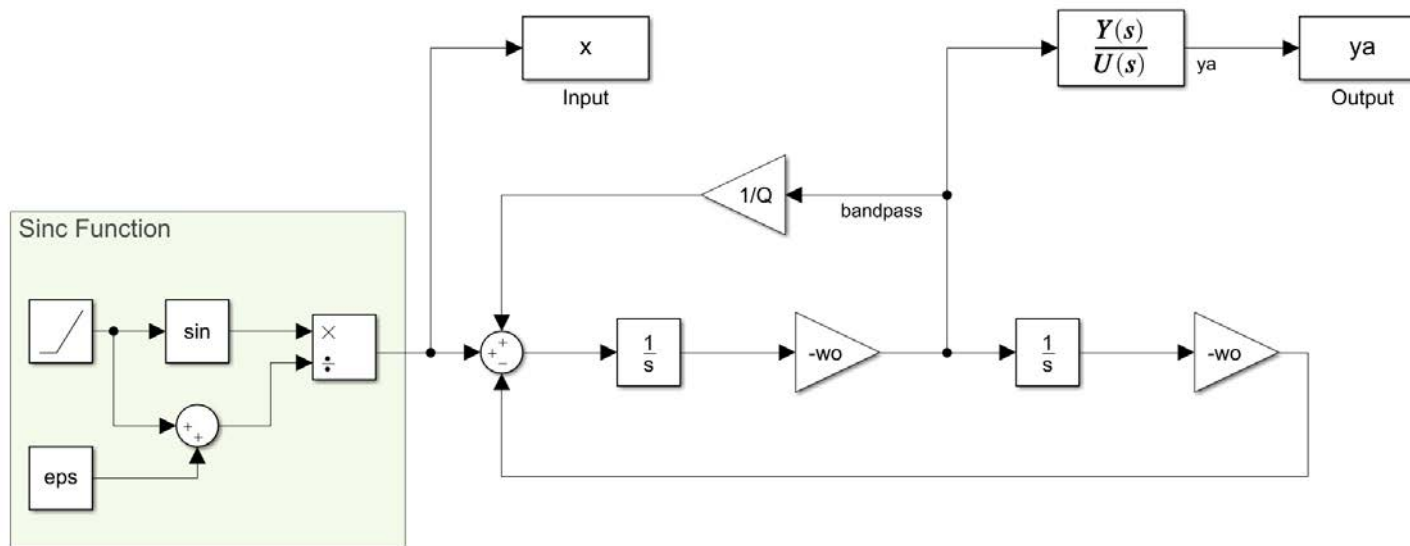
BIBLIOGRAFÍA

- [1] R.L.Klevans y R.D.Rodman, <<Introduction>> de Voice Recognition, Boston, London:Artech House, 1997, pp.1-15.
- [2] M. Price, J. Glass and A. P. Chandrakasan, "14.4 A scalable speech recognizer with deep-neural-network acoustic models and voice-activated power gating," 2017 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, 2017, pp. 244-245.
- [3] A. Raychowdhury, C. Tokunaga, W. Beltman, M. Deisher, J. Tschanz and V. De, "A 2.3nJ/frame Voice Activity Detector based audio front-end for context-aware System-on-Chip applications in 32nm CMOS," Proceedings of the IEEE 2012 Custom Integrated Circuits Conference, San Jose, CA, 2012, pp. 1-4.
- [4] M. Yang, C. Yeh, Y. Zhou, J. P. Cerqueira, A. A. Lazar and M. Seok, "A 1 μ W voice activity detector using analog feature extraction and digital deep neural network," 2018 IEEE International Solid - State Circuits Conference - (ISSCC), San Francisco, CA, 2018, pp. 346-348.
- [5] Ben Gold; Nelson Morgan; Dan Ellis, "Models of Pitch Perception," in *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*, Wiley,2011, pp.
- [6] S. Wang, T. J. Koickal, A. Hamilton, R. Cheung and L. S. Smith, "A Bio-Realistic Analog CMOS Cochlea Filter With High Tunability and Ultra-Steep Roll-Off," in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 9, no. 3, pp. 297-311, June 2015.
- [7] Z. Tüske, P. Golik, R. Schlüter and F. R. Drepper, "Non-stationary feature extraction for automatic speech recognition," 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, 2011, p.204-207.
- [8] X. Valero and F. Alias, "Gammatone Cepstral Coefficients: Biologically Inspired Features for Non-Speech Audio Classification," in *IEEE Transactions on Multimedia*, vol. 14, no. 6, pp. 1684-1689, Dec. 2012.
- [9] J. Qi, D. Wang, Y. Jiang and R. Liu, "Auditory features based on Gammatone filters for robust speech recognition," 2013 IEEE International Symposium on Circuits and Systems (ISCAS2013), Beijing, 2013, pp. 305-308.
- [10] Ben Gold; Nelson Morgan; Dan Ellis, "The Auditory System as a Bank Filter," in *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*, Wiley, 2011, pp.257-270

- [11] J Shannon, Ben & K Paliwal, Kuldip. (2003). A comparative study of filter bank spacing for speech recognition. MICROELECTRONIC ENGINEERING RESEARCH CONFERENCE.
- [12] D. O'Shaughnessy, "Linear predictive coding," in IEEE Potentials, vol. 7, no. 1, pp. 29-32, Feb. 1988.
- [13] Dave, Namrata. (2013). Feature extraction methods LPC, PLP and MFCC in speech recognition. International Journal for Advance Research in Engineering and Technology (ISSN 2320-6802). Volume 1.
- [14] B. Drost, M. Talegaonkar and P. K. Hanumolu, "Analog Filter Design Using Ring Oscillator Integrators," in IEEE Journal of Solid-State Circuits, vol. 47, no. 12, pp. 3120-3129, Dec. 2012.
- [15] L. Hernandez, E. Gutierrez and F. Cardes, "Frequency-encoded integrators applied to filtering and sigma-delta modulation," 2016 IEEE International Symposium on Circuits and Systems (ISCAS), Montreal, QC, 2016, pp. 478-481.
- [16] F. Cardes, E. Gutierrez, A. Quintero, C. Buffa, A. Wiesbauer and L. Hernandez, "0.04-mm² 103-dB-A Dynamic Range Second-Order VCO-Based Audio Sigma Delta ADC in 0.13- μ m CMOS," in IEEE Journal of Solid-State Circuits, vol. 53, no. 6, pp. 1731-1742, June 2018.
- [17] A. Babaie-Fishani and P. Rombouts, "A Mostly Digital VCO-Based CT-SDM With Third-Order Noise Shaping," in IEEE Journal of Solid-State Circuits, vol. 52, no. 8, pp. 2141-2153, Aug. 2017.
- [18] B. Razavi, "The Biquadratic Filter [A Circuit for All Seasons]," in IEEE Solid-State Circuits Magazine, vol. 10, no. 2, pp. 11-109, Spring 2018
- [19] Baker, R.J., CMOS: Circuit Design, Layout, and Simulation. Revised 2nd Ed., Wiley-IEEE, 2008, pp.550-555.
- [20] Mandal, Mrinal & Sarkar, Bishnu Charan. (2010). Ring oscillators: Characteristics and applications. Indian Journal of Pure and Applied Physics. 48. 136-145.]
- [21] M. Cho et al., "17.2 A 142nW Voice and Acoustic Activity Detection Chip for mm-Scale Sensor Nodes Using Time-Interleaved Mixer-Based Frequency Scanning," 2019 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2019, pp. 278-280.
- [22] K. Badami, S. Lauwereins, W. Meert and M. Verhelst, "24.2 Context aware hierarchical information-sensing in a 6W 90nm CMOS voice activity detector," 2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers, San Francisco, CA, 2015, pp. 1-3.
- [23] A. Raychowdhury, C. Tokunaga, W. Beltman, M. Deisher, J. W. Tschanz and V. De, "A 2.3 nJ/Frame Voice Activity Detector-Based Audio Front-End for Context-Aware System-On-Chip Applications in 32-nm CMOS," in IEEE Journal of Solid-State Circuits, vol. 48, no. 8, pp. 1963-1969, Aug. 2013.

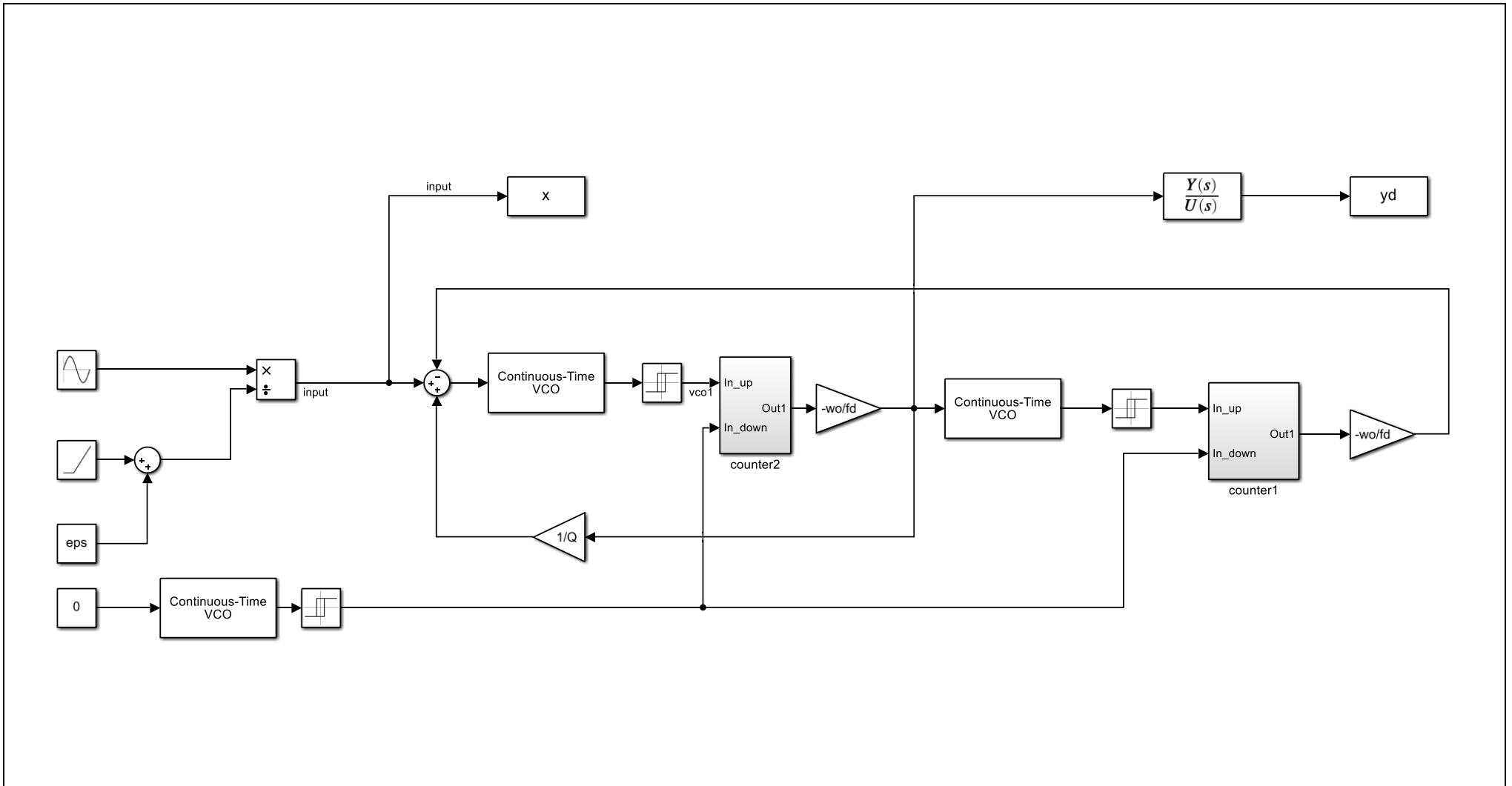
ANEXOS

I. ESQUEMÁTICOS



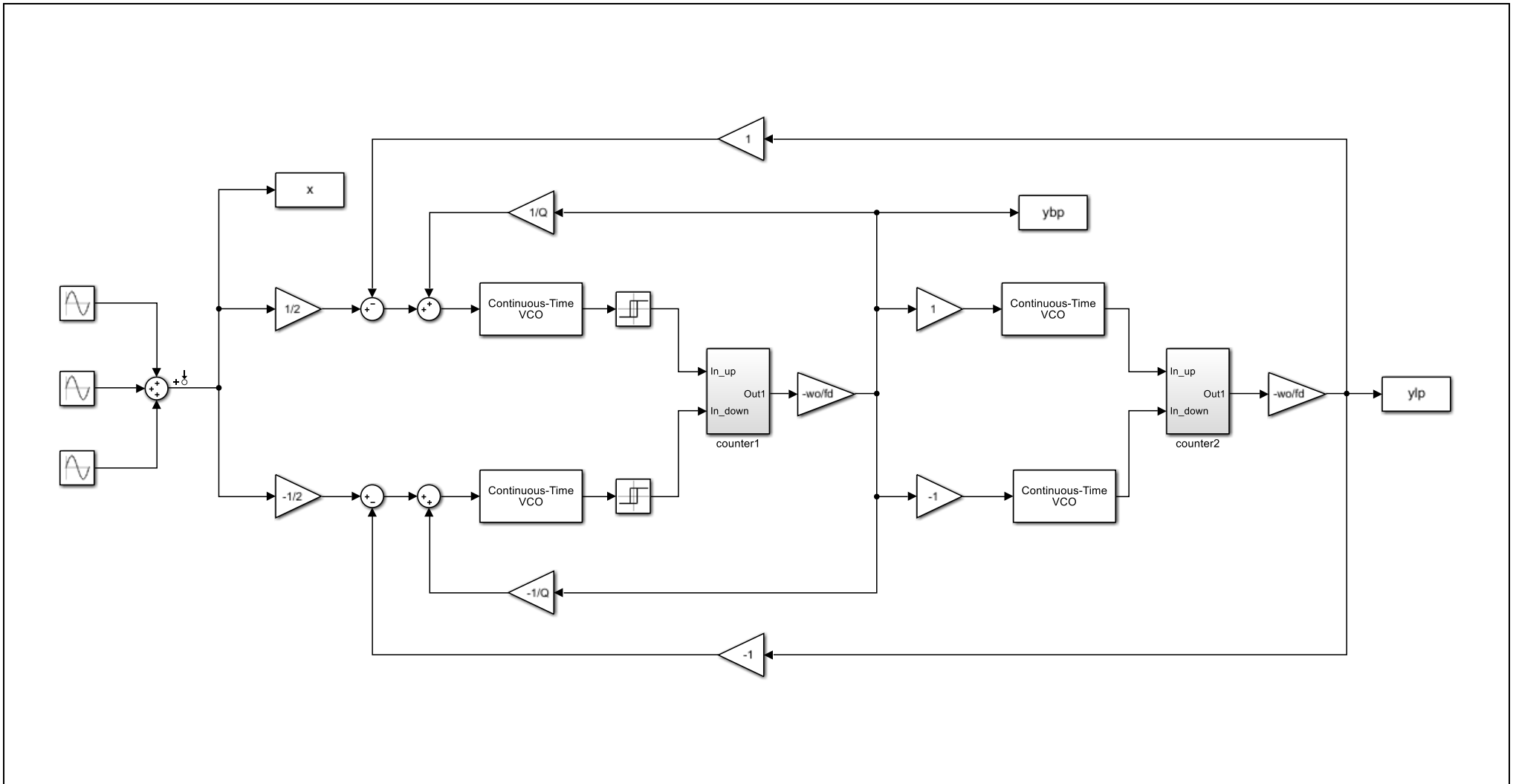
Filtro 1
(Simulink)

Descripción: Estructura clásica de filtro bicuadrático analógico. Función sinc como señal de entrada, para caracterizar el sistema.



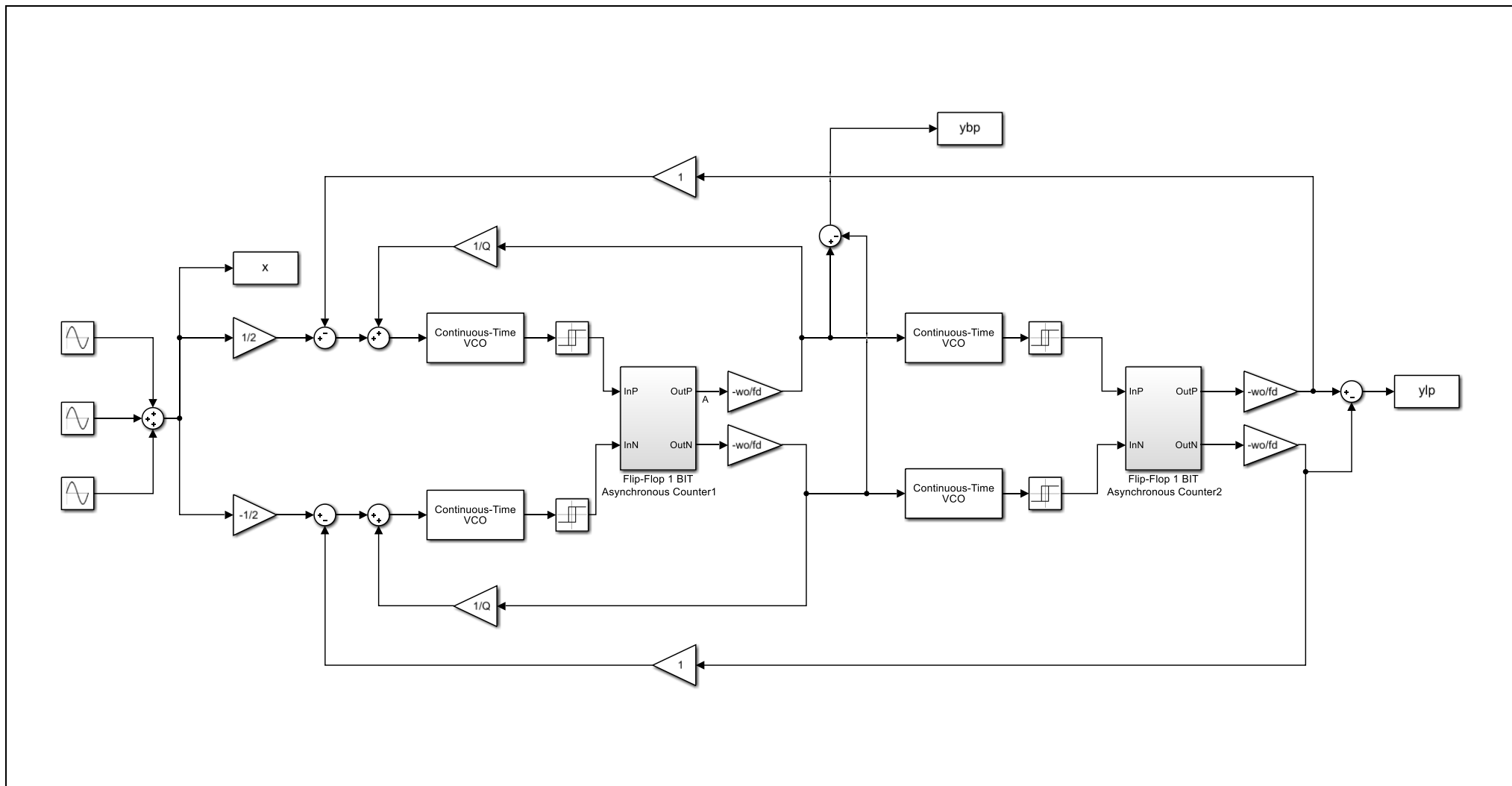
**Filtro 2
(Simulink)**

Descripción: Filtro digital con VCOs.
 Modelo single-ended. Contadores ideales.
 Función sinc como señal de entrada, para caracterizar el sistema.



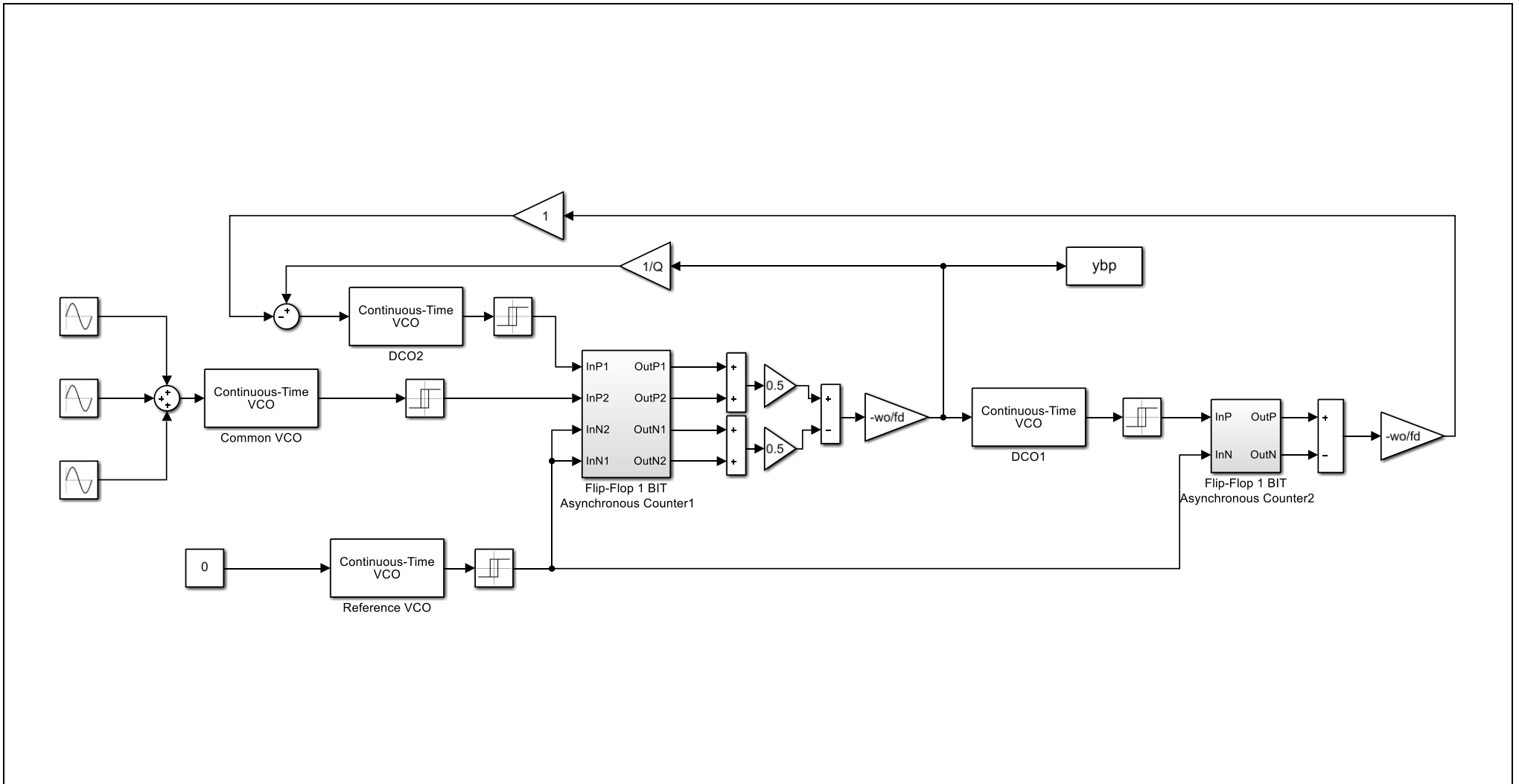
**Filtro 3
(Simulink)**

Descripción: Filtro digital con VCOs.
Modelo diferencial. Contadores ideales.
Funciones sinusoidales de entrada, espaciadas 1 década.



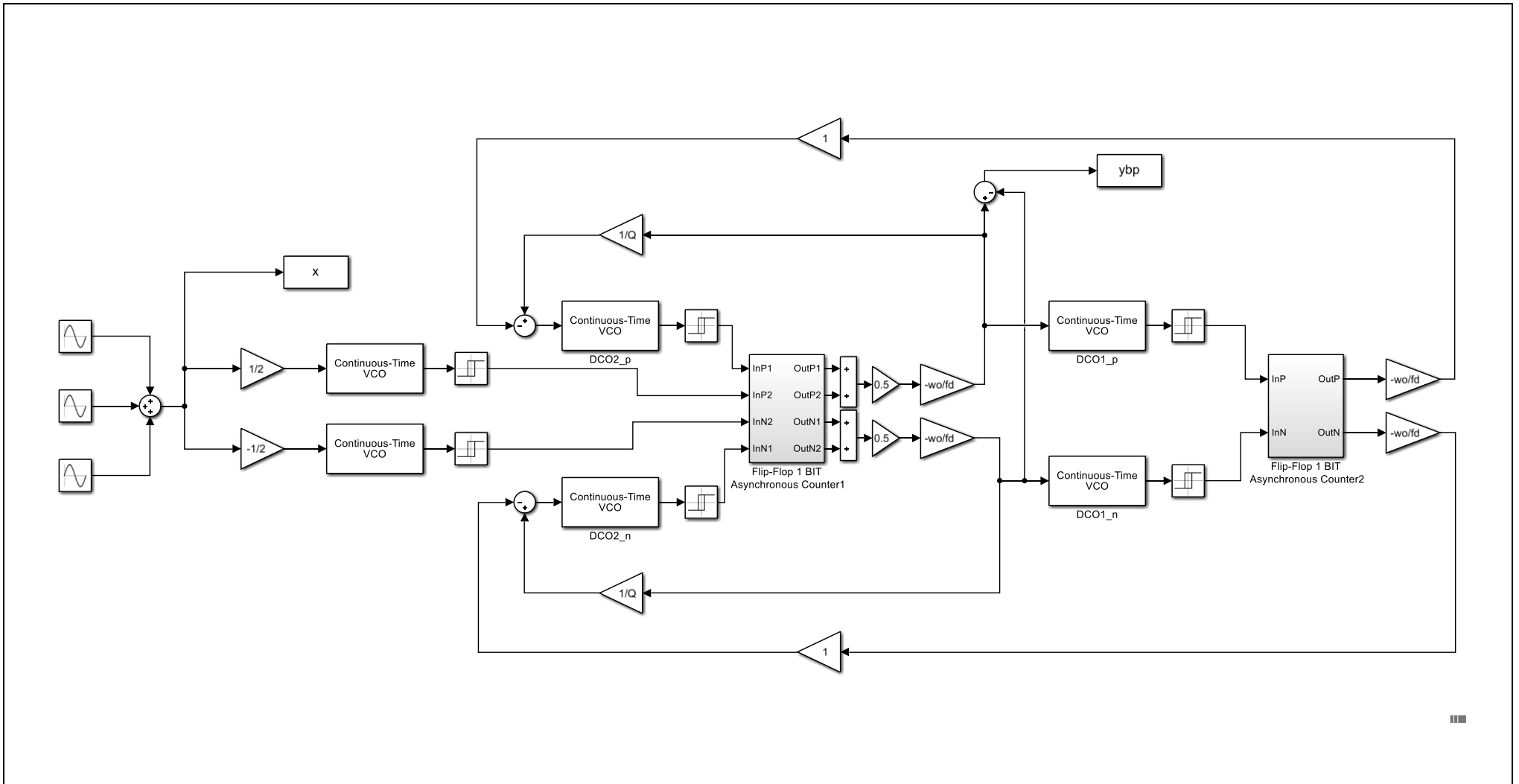
**Filtro 4
(LTspice)**

Descripción: Filtro digital con VCOs.
Modelo diferencial. Contadores reales.
Funciones sinusoidales de entrada, espaciadas 1 década.



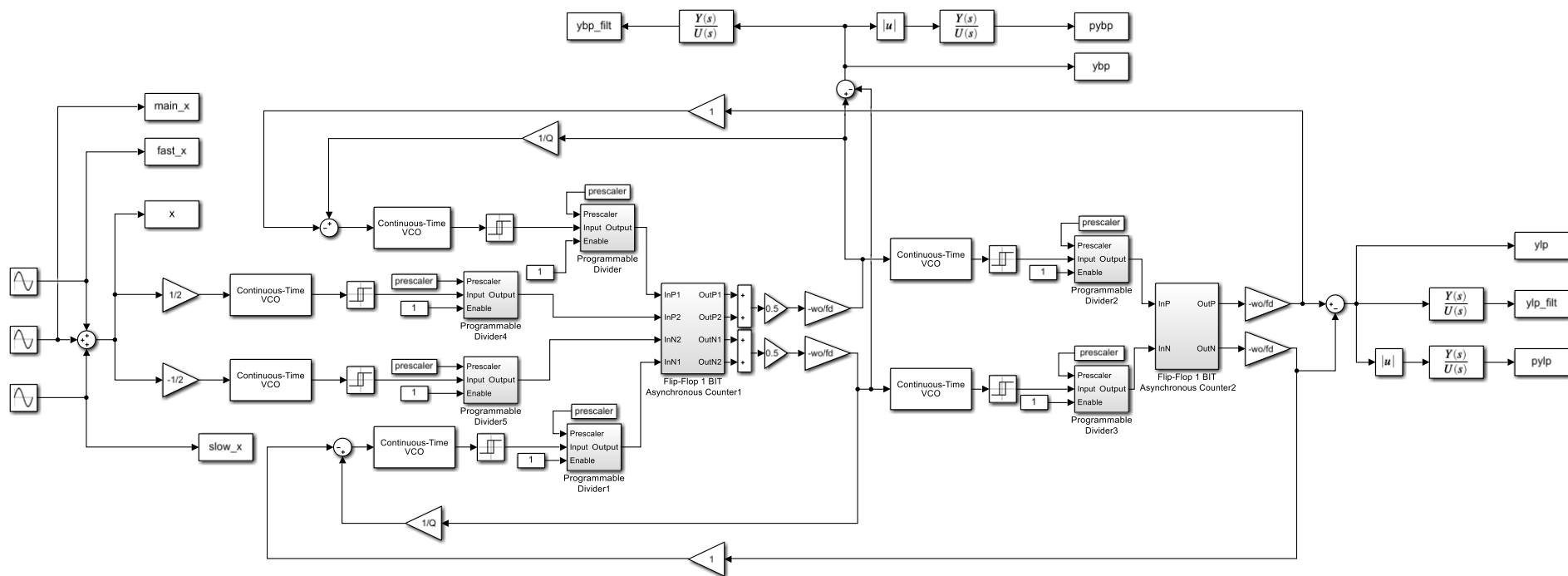
**Filtro 5
(Simulink)**

Descripción: Filtro digital con VCO común y DCOs internos. Modelo single-ended. Contadores reales (de 2 y 4 entradas). Funciones sinusoidales de entrada, espaciadas 1 década.



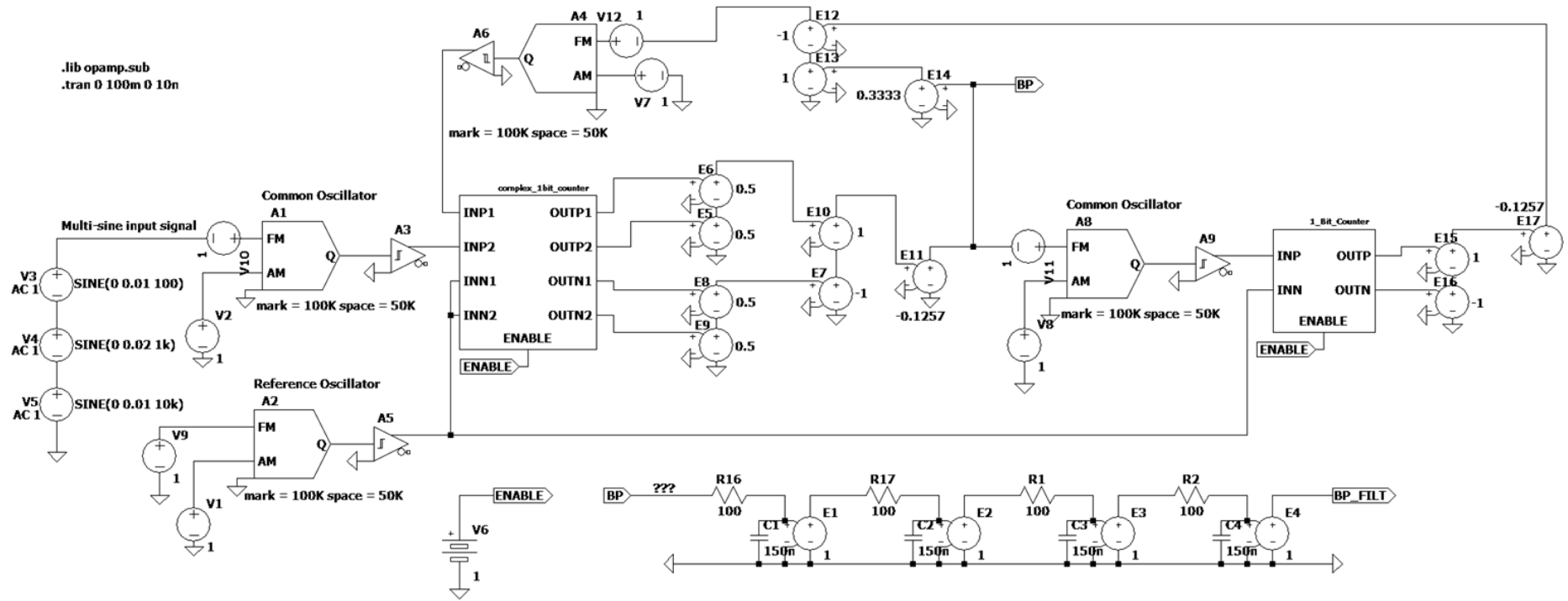
**Filtro 6
(Simulink)**

Descripción: Filtro digital con VCO común y DCOs internos.
 Modelo diferencial. Contadores reales (de 2 y 4 entradas).
 Funciones sinusoidales de entrada, espaciadas 1 década.



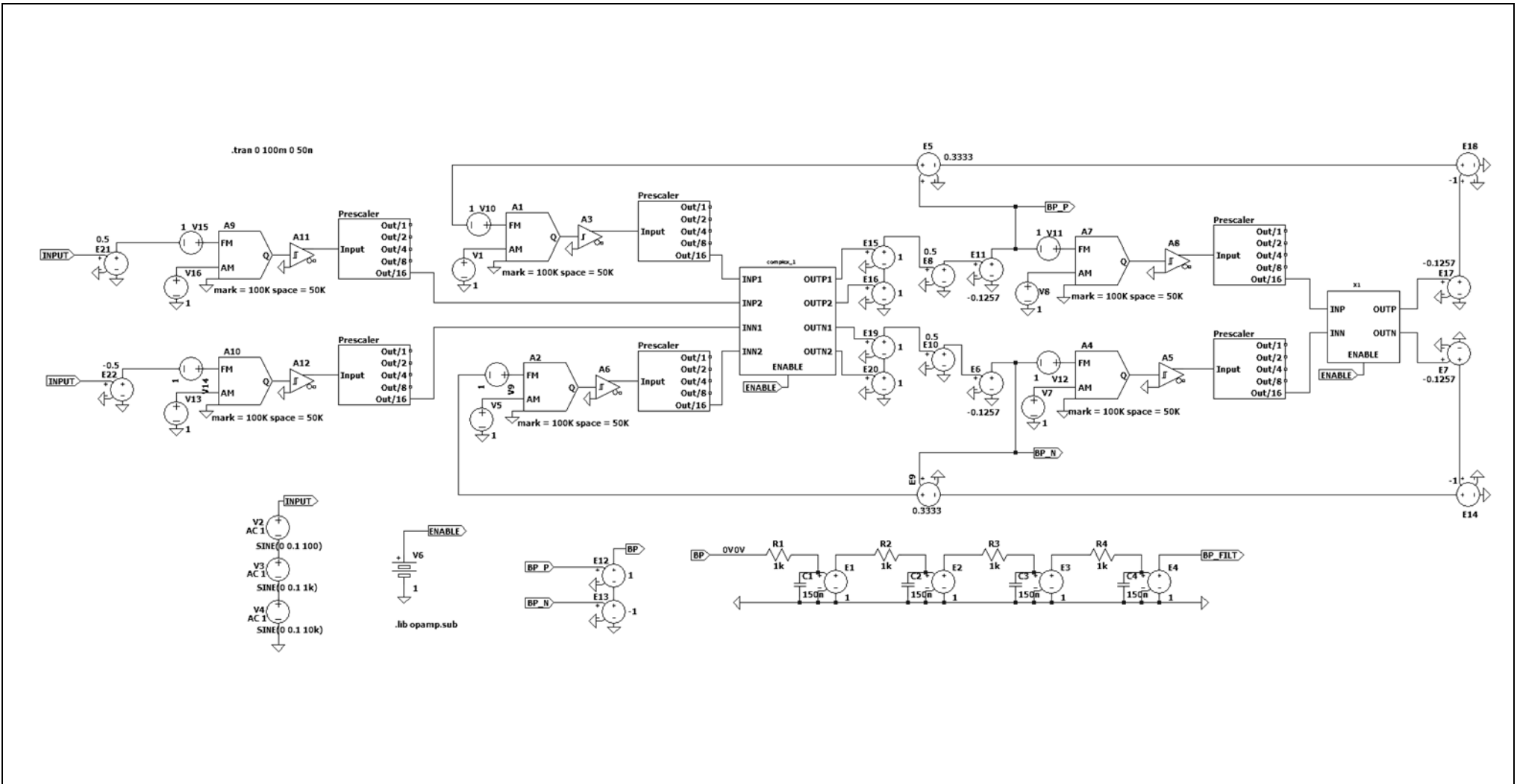
**Filtro 7
(Simulink)**

Descripción: Filtro digital con VCO común y DCOs internos. Filtro PROGRAMABLE.
Modelo diferencial. Contadores reales (de 2 y 4 entradas).
Funciones sinusoidales de entrada, espaciadas 1 década.



**Filtro 8
(LTspice)**

Descripción: Circuito del filtro digital con VCO común y DCOs internos (ideales). Modelo single-ended. Contadores reales (de 2 y 4 entradas). Funciones sinusoidales de entrada, espaciadas 1 década.



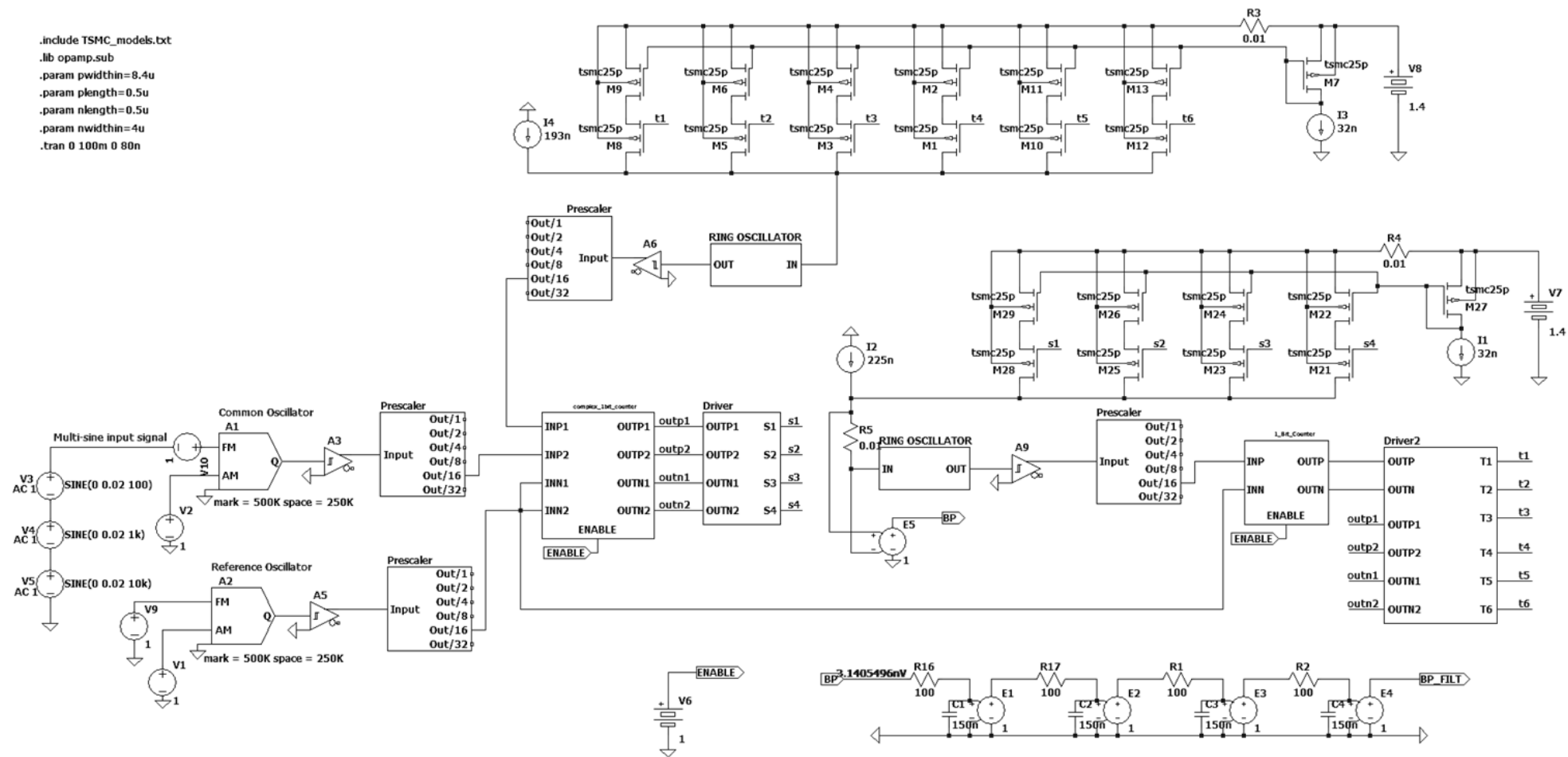
Filtro 9 (LTspice)

Descripción: Circuito del filtro digital con VCO común y DCOs internos (ideales). Filtro PROGRAMABLE. Modelo diferencial. Contadores reales (de 2 y 4 entradas). Funciones sinusoidales de entrada, espaciadas 1 década.


```

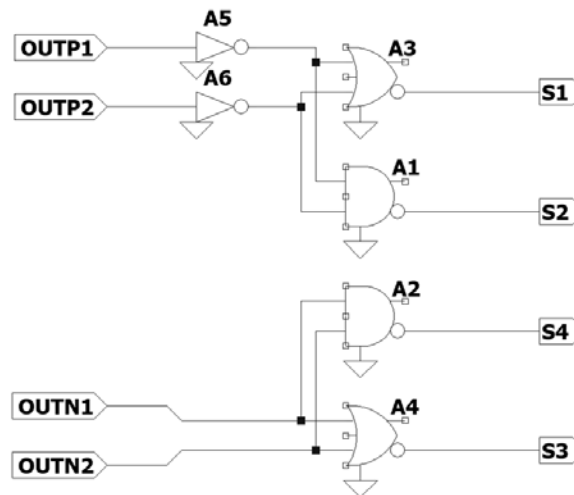
.include TSMC_models.txt
.lib opamp.sub
.param pwidthin=8.4u
.param plength=0.5u
.param nlength=0.5u
.param nwidthin=4u
.tran 0 100m 0 80n

```



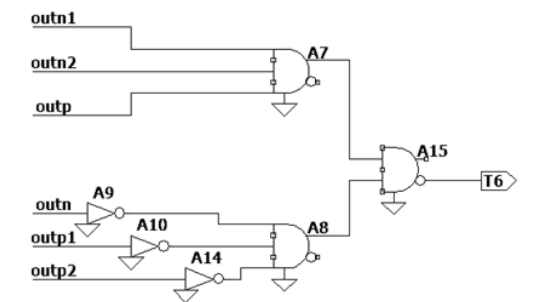
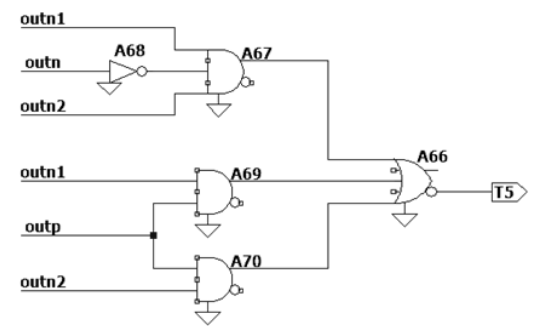
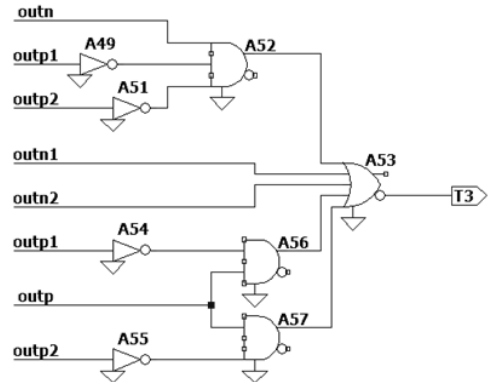
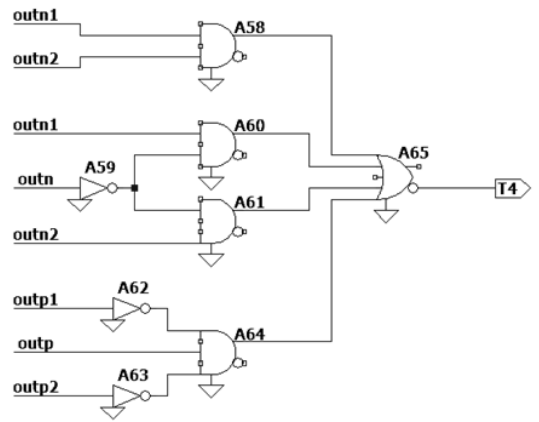
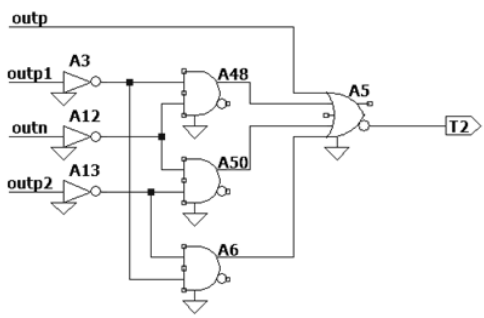
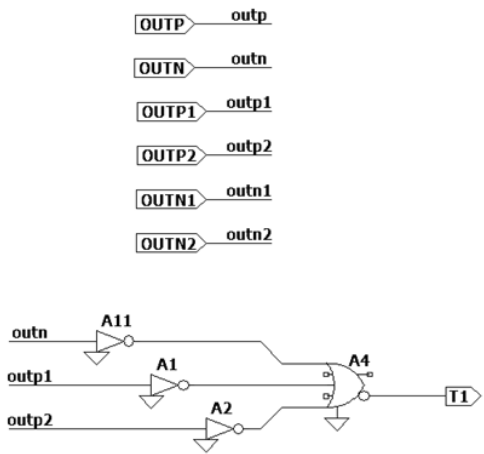
Filtro 10 (LTspice)

Descripción: Circuito del filtro digital con VCO común y DCOs internos (reales). Filtro PROGRAMABLE. Modelo single-ended. Contadores reales (de 2 y 4 entradas). Funciones sinusoidales de entrada, espaciadas 1 década.



**Bloque combinacional
(LTspice)**

Descripción: Bloque combinacional de activación/desactivación de las fuentes de corriente del DCO 1. (Filtro 10)
4 entradas procedentes del primer contador asíncrono.
No confundir con bloque combinacional interno del contador asíncrono.



Bloque combinacional (LTspice)

Descripción: Bloque combinacional de activación/desactivación de las fuentes de corriente del DCO 2. (Filtro 10) 6 entradas procedentes del primer y segundo contador asíncrono. No confundir con bloque combinacional interno del contador asíncrono.

II. ARTÍCULO PUBLICADO

VCO-based Feature Extraction Architecture for Low Power Speech Recognition Applications

Eric Gutierrez, Carlos Perez, Fernando Hernandez, Luis Hernandez
Electronics Technology Department
Carlos III University
Madrid, Spain
eric.gutierrez@uc3m.es

Abstract—This work proposes a new approach for speech recognition and voice activity detection tasks. The main limitation of the state-of-the-art solutions for these type of applications is the high power consumption and occupied area, which strongly limit their implementation on portable devices. In this manuscript, we propose a new solution based on making use of voltage-controlled-oscillators based analog-to-digital converters (VCO-based ADCs). VCO-based ADCs have shown a power and area efficient way to implement audio ADCs. Here, a VCO-based band-pass filter is proposed to extract the features of analog input sound signals and generate a digital output stream encoding the mel-frequency cepstrum coefficients (MFCCs). This digital signal will be processed afterwards by a decision circuit to determine whether the sound signal corresponds to a pattern. Additionally, the same VCO can be reused to implement both the band-pass filters in the decision mode and a standard ADC output once a keyword or voice has been detected, saving significant area. A behavioral model of the filter has been made to validate the performance by simulation. Then, the architecture has been designed using a 130-nm CMOS process to get a realistic outlook of the power consumption and occupied area we may expect on silicon. Finally, the proposed solution is compared to other equivalent solutions.

Index Terms—speech recognition, voice activity detector, voltage-controlled oscillator, analog-to-digital conversion, artificial intelligence, internet-of-things

I. INTRODUCTION

Nowadays there is an increasing interest in voice user interfaces implemented on portable devices, such as speech recognition or keyword spotting [1], [2]. In this context, voice activity detectors (VADs) are critical to take an audio input stream and be able of distinguish between human voice and noise [3]–[5].

The conventional approach for voice activation detection is made by digitizing raw data coming from a sensor and making intensive digital post-processing, including windowing, FFT, filtering and power estimation (Fig.1(a)), to get the features of the data. After that, the digital features feed a decision circuit, such as a feedforward neural network or a decision tree, to decide whether the data correspond to human voice or not. However, the always-on nature and the high power consumption of the digital circuits strongly limit their use in portable devices with battery-life limitation [6].

This work was supported by the CICYT project TEC2017-82653-R, Spain.

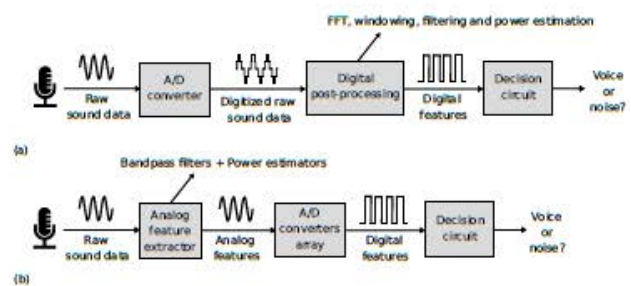


Fig. 1. VAD approaches: (a) feature extraction in the digital domain, and (b) in the analog domain.

The requirement of ultra-low power circuits for VADs has lead the designers to a different perspective, more devoted to “analog” solutions based on extracting the features in the analog domain and perform the analog-to-digital conversion afterwards (Fig.1(b)). To make the analog extraction of the features a bank of band-pass filters and power estimators are used to get the energy of the input signal filtered in different bands of interest. When the features have been digitized, they feed a decision circuit in the same way as in the conventional approach. This new approach reduces considerably the power consumption of the features extractor circuit, making them well-suited for portable devices [4], [5]. The disadvantage of this approach is the larger occupied area in comparison to the digital approach, due mainly to the large capacitors required to implement the analog filters in the audio band. In addition, a conventional ADC is required on a side of the band-pass filters to conventionally digitize the audio data, once the triggering event has been detected.

In this manuscript, we propose a new approach based on the use of voltage-controlled-oscillators based ADCs (VCO-based ADCs) to perform the feature extraction. VCO-based ADCs have recently attracted the attention of audio designer’s community due to their low power and dynamic range suited to human hearing [7], [8]. Then, how to implement bi-quadratic filters with VCOs was proposed in [9]–[11]. Here, we will make use of VCO-based bi-quadratic filters to make the feature extraction and generate a digital signal that could be connected to a decision circuit. Note that the output signal of VCO-based filter is already a digital signal. Therefore, in comparison

to the approach of Fig.1(b), we save power because we do not need the array of A/D converters. Additionally, filters implemented with VCOs do not require of large capacitors to implement long time constants, but digital counters instead that can be implemented with low-size transistors. Consequently, the proposed solution is expected to have much lower area in comparison to both the approaches of Fig.1.

The outline of the manuscript is as follows. Section II describes the theoretical foundations of the proposed VCO-based system and validates the approach by simulation. Section III extends the architecture to practical circuits to get an estimation of the power consumption and occupied area. Finally, Section IV concludes the paper.

II. VCO-BASED BI-QUADRATIC FILTERS

Voice recognition tasks usually make use of mel-frequency cepstrum coefficients (MFCCs) to efficiently distinguish between human voice and other sounds like background noise [12]. These coefficients are commonly calculated through digital circuitry by means of short-time Fourier Transforms, filter banks and discrete cosine transforms (DCTs). The output of the DCTs, typically scaled between 0 and 1, represent the MFCCs that will be used afterwards in the decision stage.

We propose a new approach to extract the MFCCs based on making use of highly scalable and ultra-low power VCOs occupying a minimum area. The proposed approach is based on the well-known notion that a VCO followed by a digital counter works as an integrator [8]. This enables us to build filters with VCOs and, specially for our interest, the bank of band-pass filters required in voice recognition tasks.

To proceed with the implementation of the bank of filters, we have selected the generic bi-quadratic filter (Fig.2(a)). As we are interested in the band-pass output, our transfer-function will look as follows:

$$H(s) = \frac{-K\omega_0 s}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2}, \quad (1)$$

where K is the input gain, Q is the quality factor and ω_0 is the center frequency. Conventional analog bi-quadratic filters can be implemented with opamps operating with very low power consumption. Nevertheless, the lower the center frequency the larger the capacitor and the resistor required. Consequently, the occupied area increases prohibitively.

A. Architecture for VCO-based bi-quadratic filter

The oscillation frequency of a VCO $f_{osc}(t)$ with an input signal $x(t)$ follows:

$$f_{osc}(t) = f_o + K_{VCO} \cdot x(t), \quad x(t) \in [-1, 1], \quad (2)$$

where f_o is the rest oscillation frequency and K_{VCO} is the gain of the VCO. The input signal is assumed to be dimensionless.

In [10], it was shown that an integrator can be built with a pulse frequency modulator (PFM), composed of a VCO (that integrates the phase of the input signal), and an asynchronous digital counter (that quantifies the phase). Taking this equivalence into consideration, we are able to build a

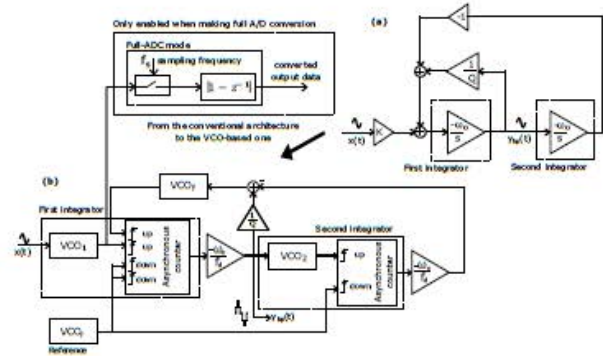


Fig. 2. Conventional (a) and proposed (b) single-ended VCO-based architectures for a bi-quadratic filter.

bi-quadratic filter with VCOs and counters just by replacing the conventional opamp-based integrators with the VCO-based integrator structures.

The proposed architecture of the VCO-based band-pass filter is depicted in Fig. 2(b). As can be observed, each integrator is composed of a VCO and a counter. Each counter has an internal count value that is increased by each rising edge in any “up” input and decreased by each rising edge in any “down” input. The outputs of the counters are a digital representation of the correspondent count value.

Note that in the architecture shown in Fig. 2(b) we have a reference VCO (VCO_r) whose output gets into the “down” inputs of the counters. This VCO is required here to subtract the offset value of the integrated phase of the VCO. In the case of a single-ended structure (as is the case of Fig.2(b)) VCO_r is required, but it is not the case for a differential architecture as we will see later on.

B. Behavioral simulation

A behavioral model of the VCO-based system of Fig. 2(b) was made and simulated to validate the architecture. To check the proper performance of the filter, the analog input signal was composed of three sinusoidal waveforms spaced one decade between them and with a central tone at 1 kHz. All the VCOs have the same oscillation parameters, $f_o = 100$ kHz and $K_{VCO} = 50$ kHz. The central frequency of the filter is equal to 1 kHz. The spectra of both the input and the output signals can be observed in Fig.2 (a) and (b), respectively.

Fig.2 (b) shows the output signal, but low-pass filtered to avoid aliasing phenomena. It can be viewed that the most powerful component corresponds to the closest frequency to the center frequency of the filter. In addition, it can be appreciated that this frequency component is 20 dB more powerful than the other two, as expected from (1).

Finally, it is relevant to notice that the filter output signal is actually a digital signal, corresponding to the output of a counter. This is of high relevance because, firstly, we no longer require an ADC to convert the analog features extracted here and, secondly, the first VCO can be reused to make the

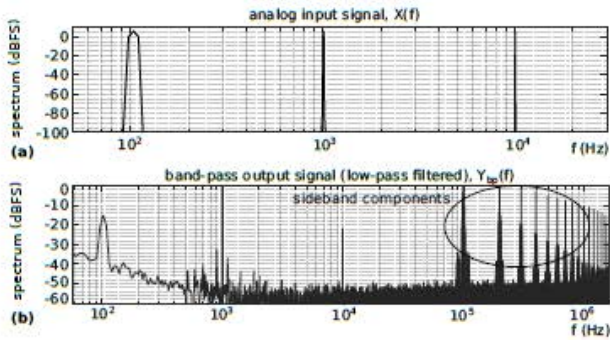


Fig. 3. Spectra of the input (a) and the output (b) signals of the filter.

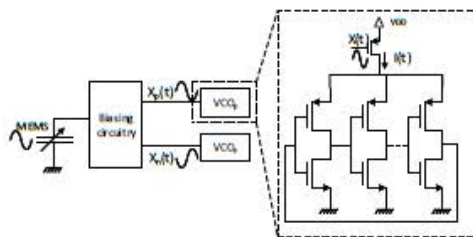


Fig. 4. Front-end of the architecture: MEMS, biasing circuitry and first VCOs.

conventional A/D conversion after the detection stage. This means strongly reducing the power consumption and the area in comparison to other speech recognition solutions.

III. CIRCUIT APPLICATION

The architecture proposed in Fig.2 was designed in a 130-nm CMOS process to get an intuition about the area and power consumption we may expect from a silicon test. In this section we will describe all the designed blocks, especially focusing on the VCOs and the asynchronous counters, and we will get an approach to the area and power consumption of each of them.

A. VCO and front-end circuit

The first element we need in a speech recognition task is a sound source. In this case, we have selected a capacitor-based MEMS sensor in accordance with [8]. The performance of this sensor is based on the variation of a capacitance that depends on the input sound pressure. This capacitance, along with a biasing circuitry, generates a proportional voltage that gets into the first VCO. Although for simplicity we had assumed a single-ended configuration in the previous section, now we will assume a differential configuration which would be the practical architecture for a silicon implementation. Then, we have a differential output at the MEMS sensor that is connected to two identical VCOs, as shown in Fig. 4.

The VCO is built with a ring-oscillator configuration [7], where the input voltage signal $x(t)$ is turned into a current $i(t)$ that feeds the inverters. In our proposal, these first ring-oscillators are composed of eleven taps. However, only one

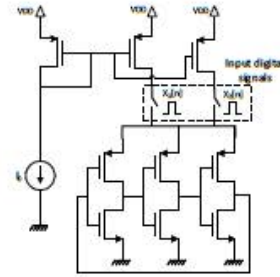


Fig. 5. Designed DCO circuit with three inverters.

tap per ring-oscillator of them will be connected to the asynchronous counter afterwards (the P-side ring-oscillator to the “up” input and the N-side one to the “down” input). The oscillation parameters of the VCO define the number of taps.

B. DCOs

If we look at the VCOs of Fig. 2, we will notice that the only VCO whose input signal is analog is the first one. In the remaining VCOs the input signal is digital, which means that they are digitally-controlled oscillators (DCOs). This supposes that they must be implemented with a different architecture with respect to the circuit of the ring-oscillator depicted in Fig. 4. Again, they are implemented with a ring-oscillator configuration. The designed circuit for the DCOs is shown in Fig. 5. In this case, the structure of the inverters connected in a ring remains, but the current that feeds them can be selected by means of switches. These switches will be closed or not depending on the value of the input digital signal. Finally, the current is the same for all the branches and it is taken from a current mirror structure. For the DCOs, an architecture with only three inverters is chosen.

C. Digital logic design: asynchronous counters

The first counter has two inputs for the P-side and two inputs for the N-side. The output is a 4-bit signal. Note that for a differential architecture we no longer need the reference VCO of Fig. 2 because the phase offset term is subtracted between the first two VCOs of the differential architecture (Fig. 4). Fig. 6(a) depicts the schematic of this counter. The four input signal get into the clock input of four flip-flops and the outputs of these flip-flops are connected to a combinational digital logic that compares the count value for both sides of the differential configuration and makes the subtraction.

The second counter has two inputs, one for each side of the differential configuration. Fig. 6(b) shows the schematic of this counter. In this case, for the selected oscillation parameters, we only require a 2-bit counter, which significantly simplifies the circuit of the counter. For both counters, two very close edges with opposite directions (counting up and down) will not be lost due to the input flip-flops. If metastability occurs, the logic will only take longer time to resolve.

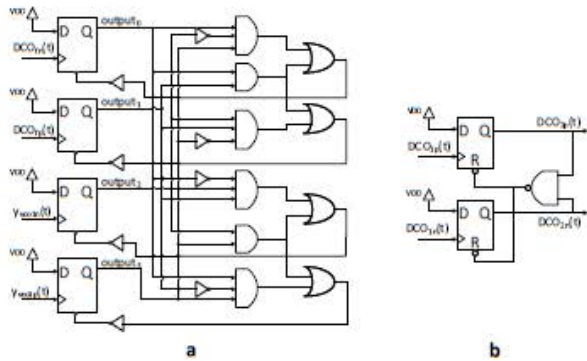


Fig. 6. (a) 4-bit counter and, (b) 2-bit counter.

TABLE I
ESTIMATED POWER CONSUMPTION AND OCCUPIED AREA PER CHANNEL

Component	Current (nA)	Area (μm^2)	#	Total current (nA)	Total area (μm^2)
VCO	0.222	210.97	2	0.444	421.94
DCO	0.222	290.40	4	0.888	1161.60
4-bit counter	39.71	135.78	1	39.710	135.78
2-bit counter	11.94	83.68	1	11.940	83.68
Total				52.982	1803.00

D. Power consumption and area estimation

To validate the proposed circuit, we have designed the previous circuits in a 130-nm CMOS process. The supply voltage for all of the described blocks is 1.2 V. The oscillation parameters for all the oscillators remain from the behavioral simulation of Section II. With these conditions we performed a transient simulation to estimate the power consumption of the architecture. The results for the power consumption and the estimated area are shown in Table I.

The values provided by Table I refers to one single channel of the feature extractor circuit. If we assume to make use of 16 channels to be able of making a proper decision, the whole power consumption will be equal to $1.01 \mu\text{W}$ and the occupied area will be 0.03 mm^2 .

E. Comparison to state-of-the-art applications

In Table II our solution is compared to other equivalent solutions. As can be observed, our proposal shows the best performance in terms of area due to the high scalability of the architecture and its mostly digital implementation. Additionally, the power consumption is competitive with respect to ultra-low power analog solutions.

IV. CONCLUSION

A new approach for circuits devoted to speech recognition and voice activity detection is proposed. This new approach is based on using VCO-based ADC filters, which enable us to implement an architecture with very small area and low power consumption. The architecture is validated by behavioral simulation and designed in a 130-nm CMOS process to estimate the occupied area and the power consumption. The simulated

TABLE II
COMPARISON TO EQUIVALENT SOLUTIONS

	Power (μW)	Channel number (nW)	Area (mm^2)	Approach	ADC included
Our solution	1.01	16	0.03	Hybrid (VCOs)	Yes
[3]	0.06	16-48	0.73 approx.	Hybrid	No
[4]	6 ¹	16	2	Analog	No
[5]	0.38	16	0.16	Analog	No
[6]	>50	-	-	Digital	No

¹ The power of the classification stage is included.

results show that the solution is the best in terms of area with a competitive power consumption, very suitable for portable devices and Internet-of-things (IoT) applications.

REFERENCES

- [1] S. Yin et al., "A 141 uW, 246 PJ/Neuron Binarized Convolutional Neural Network Based Self-Learning Speech Recognition Processor in 28NM CMOS," 2018 IEEE Symposium on VLSI Circuits, Honolulu, HI, 2018, pp. 139-140.
- [2] M. Price, J. Glass and A. P. Chandrakasan, "A Low-Power Speech Recognizer and Voice Activity Detector Using Deep Neural Networks," in IEEE Journal of Solid-State Circuits, vol. 53, no. 1, pp. 66-75, Jan. 2018.
- [3] M. Cho et al., "17.2 A 142nW Voice and Acoustic Activity Detection Chip for mm-Scale Sensor Nodes Using Time-Interleaved Mixer-Based Frequency Scanning," 2019 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2019, pp. 278-280.
- [4] K. Badami, S. Lauwereins, W. Meert and M. Verhelst, "24.2 Context-aware hierarchical information-sensing in a $6\mu\text{W}$ 90nm CMOS voice activity detector," 2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers, San Francisco, CA, 2015, pp. 1-3.
- [5] M. Yang, C. Yeh, Y. Zhou, J. P. Cerqueira, A. A. Lazar and M. Seok, "A $1\mu\text{W}$ voice activity detector using analog feature extraction and digital deep neural network," 2018 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, 2018, pp. 346-348.
- [6] A. Raychowdhury, C. Tokunaga, W. Beltman, M. Deisher, J. W. Tschanz and V. De, "A 2.3 nJ/Frame Voice Activity Detector-Based Audio Front-End for Context-Aware System-On-Chip Applications in 32-nm CMOS," in IEEE Journal of Solid-State Circuits, vol. 48, no. 8, pp. 1963-1969, Aug. 2013.
- [7] E. Gutierrez, L. Hernandez, F. Cardes and P. Rombouts, "A Pulse Frequency Modulation Interpretation of VCOs Enabling VCO-ADC Architectures With Extended Noise Shaping," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 65, no. 2, pp. 444-457, Feb. 2018.
- [8] F. Cardes, E. Gutierrez, A. Quintero, C. Buffa, A. Wiesbauer and L. Hernandez, "0.04- mm^2 103-dB-A Dynamic Range Second-Order VCO-Based Audio $\Sigma\Delta$ ADC in $0.13\text{-}\mu\text{m}$ CMOS," in IEEE Journal of Solid-State Circuits, vol. 53, no. 6, pp. 1731-1742, June 2018.
- [9] B. Drost, M. Talegaonkar and P. K. Hanumolu, "Analog Filter Design Using Ring Oscillator Integrators," in IEEE Journal of Solid-State Circuits, vol. 47, no. 12, pp. 3120-3129, Dec. 2012.
- [10] L. Hernandez, E. Gutierrez and F. Cardes, "Frequency-encoded integrators applied to filtering and sigma-delta modulation," 2016 IEEE International Symposium on Circuits and Systems (ISCAS), Montreal, QC, 2016, pp. 478-481.
- [11] L. B. Leene and T. G. Constandinou, "Time Domain Processing Techniques Using Ring Oscillator-Based Filter Structures," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 64, no. 12, pp. 3003-3012, Dec. 2017.
- [12] Q. Li, H. Zhu, F. Qiao, X. Liu, Q. Wei and H. Yang, "Energy-efficient MFCC extraction architecture in mixed-signal domain for automatic speech recognition," 2018 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH), Athens, 2018, pp. 1-3.

FIN DEL DOCUMENTO