

# Trabajo Fin de Grado

## Grado en Ingeniería de las Tecnologías de Telecomunicación

### Esquema de linealización de doble banda concurrente con arquitectura de aprendizaje directo

Autor: Miguel Nogales González-Regueral

Tutores: Juan Antonio Becerra González, María José Madero Ayora

Dpto. Teoría de la Señal y Comunicaciones  
Escuela Técnica Superior de Ingeniería  
Universidad de Sevilla

Sevilla, 2021





Trabajo Fin de Grado  
Grado en Ingeniería de las Tecnologías de Telecomunicación

# **Esquema de linealización de doble banda concurrente con arquitectura de aprendizaje directo**

Autor:

Miguel Nogales González-Regueral

Tutores:

Juan Antonio Becerra González

Profesor Ayudante Doctor

María José Madero Ayora

Profesora Titular de Universidad

Dpto. Teoría de la Señal y Comunicaciones  
Escuela Técnica Superior de Ingeniería  
Universidad de Sevilla

Sevilla, 2021



Trabajo Fin de Grado: Esquema de linealización de doble banda concurrente con arquitectura de aprendizaje directo

Autor: Miguel Nogales González-Regueral  
Tutores: Juan Antonio Becerra González  
María José Madero Ayora

El tribunal nombrado para juzgar el trabajo arriba indicado, compuesto por los siguientes profesores:

Presidente:

Vocal/es:

Secretario:

acuerdan otorgarle la calificación de:

El Secretario del Tribunal

Fecha:



# Agradecimientos

---

**E**l fin de una increíble etapa. Gracias a todas las personas con las que he estado hasta hoy, aunque haya sido poco tiempo, por influenciarme y hacerme ser quién soy.

Primero de todos, gracias a los que sin duda más relación han tenido con este trabajo, al grupo de Predistorsión (llamado así por mí), concretamente a mis tutores Juan Antonio y María José, que han estado ahí apoyando en las partes más oscuras del TFG donde nada funcionaba y al final ha salido todo genial, y ojalá trabajar codo con codo en el futuro y a Elías, que hemos hecho un trabajazo. Gracias también a mis compañeros de piso Ángel, Ricardo y Antonio por los dos últimos años y esperemos que haya muchos más... y muchas series, películas, videojuegos... y de todo. A las tres Marías, por las quedadas y por la fiesta, en especial a Majo, de verdad, muchísimas gracias por estos años. Muchas gracias a mi familia, que tanto tiempo en Sevilla no me deja verlos tanto como quisiera pero aun así les quiero lo mismo o más incluso. También gracias a todos mis compañeros de la facultad, que lamentablemente nos hemos visto mucho menos debido al CoViD, pero pronto estaremos de vuelta en la ETSI. Y en especial a nuestro grupo, que pronto volveremos a conquistar la sala de estudio e ir a comer a la cafetería. A mi grupo de Badajoz...muchas gracias, que igual que con mi familia, demasiado tiempo en Sevilla pero estoy deseando de veros y quedar allí, pienso en todos y especialmente a los que no lo están pasando del todo bien, tenéis todo mi apoyo. Y como último, pero no menos importante, gracias a Abraham, que indirectamente me ha ayudado muchísimo en este trabajo.

Y para acabar, ahora si de verdad, una pequeña ruptura de la cuarta pared. A tí, que estas leyendo esto sin haber sido nombrado en lo anterior, si necesitas algo de mí por que el tema que estás leyendo te interesa, quieres ayuda, o cualquier otra cosa busca mi correo en Google y estoy para lo que quieras, que si has llegado hasta aquí es que interés desde luego tienes.

*Miguel Nogales González-Regueral  
Sevilla, 2021*



# Resumen

---

Este trabajo fin de grado presenta un esquema directo de predistorsión digital en doble banda concurrente, diseñado para la disminución de los efectos no lineales causados por los amplificadores de potencia al operar cerca de la zona de saturación.

En las comunicaciones móviles actuales, 5G y próximamente 6G, es cada vez más importante el poder disponer del mayor ancho de banda posible en conjunto con una capacidad de asignación de recursos espectrales muy elevada, lo que da lugar a la necesidad de adaptar los sistemas de predistorsión a la banda dual concurrente, en su aplicación con *carrier aggregation*.

El proyecto comienza con una introducción sobre los pilares principales de la predistorsión digital, el tipo de señal usada, que es OFDM, y los amplificadores de potencia. Posteriormente se profundiza más en este tipo de sistemas y sus partes, haciendo hincapié en las posibles arquitecturas de estos, directa e indirecta. También, como tema central que es para este trabajo, se habla de la implementación del sistema en doble banda junto con la presentación de sus modelos, para ver luego algunas de las técnicas más populares para su mejora de rendimiento. Para acabar se estudian los resultados obtenidos en dos laboratorios, el *online* de la Universidad de Chalmers y el del grupo de Sistemas de Radiocomunicación de la Universidad de Sevilla.



# Abstract

---

In this final degree project a concurrent double-band direct learning digital predistortion scheme is presented, designed with the aim of lowering the non-linearities caused by the power amplifiers when working close to saturation.

In current mobile communications systems, 5G and soon 6G, having the greater band-width available is increasingly important, in conjunction with a high spectrum management capability, which creates the need of adapting the predistortion systems to the dual band scheme: this is the case of carrier aggregation.

This project begins with an introduction on the principal basis of digital predistortion, the type of signal used, OFDM, and the power amplifiers. Next, this kind of systems and its parts are studied in depth, emphasizing the two principal architectures choices, direct and indirect. Also, as this work's principal subject, double band systems' implementation is discussed along with the presentation of its models, after that, some of the most relevant techniques for its boost on performance are reviewed. Lastly, experimental results are analysed for online laboratory of University of Chalmers and the Radiocommunications Systems laboratory from Universidad de Sevilla.



# Índice Abreviado

---

<i>Resumen</i>	III
<i>Abstract</i>	V
<i>Índice Abreviado</i>	VII
<i>Índice</i>	IX
<i>Notación</i>	XI
<b>1 Introducción</b>	<b>1</b>
<b>2 Linealización</b>	<b>3</b>
2.1 Señales OFDM	3
2.2 Amplificadores de potencia	4
2.3 Sistemas de predistorsión digital	9
2.4 Series de Volterra y su aplicación en el modelado no lineal	9
<b>3 Sistemas de predistorsión digital</b>	<b>13</b>
3.1 Arquitecturas de DPD	13
3.2 Bloques constitutivos del sistema DPD	15
3.3 Elección de los parámetros óptimos del DPD	17
3.4 Posibles problemas del DPD	18
<b>4 DPD de doble banda</b>	<b>21</b>
4.1 DPD en banda dual, qué es y qué aporta	21
4.2 Dificultades añadidas del modelo dual	22
4.3 Modelos de doble banda	24
4.4 Esquema usado	26
<b>5 Técnicas complementarias del DPD</b>	<b>29</b>
5.1 Técnicas de coeficiente de aprendizaje dinámico	29
5.2 Técnicas de selección de coeficientes	30
5.3 Técnicas de regularización	33
<b>6 Pruebas experimentales</b>	<b>35</b>
6.1 Banco de medida experimental	35
6.2 Laboratorio de la Universidad de Chalmers	36
6.3 Laboratorio de Sistemas de Radiocomunicación de la Universidad de Sevilla	38
6.4 Uso de técnicas de mejora del DPD dual	39
<b>7 Conclusiones</b>	<b>43</b>
<b>Apéndice A Códigos MatLab</b>	<b>45</b>

<b>Apéndice B Modelos 2D-GMP Usados</b>	<b>53</b>
<b>Apéndice C Artículo Técnicas de CFR</b>	<b>57</b>
<b>Apéndice D Artículo URSI DPD Dual con Ridge</b>	<b>61</b>
<i>Índice de Figuras</i>	65
<i>Índice de Tablas</i>	67
<i>Bibliografía</i>	69
<i>Glosario</i>	73

# Índice

---

<i>Resumen</i>	III
<i>Abstract</i>	V
<i>Índice Abreviado</i>	VII
<i>Índice</i>	IX
<i>Notación</i>	XI
<b>1 Introducción</b>	<b>1</b>
<b>2 Linealización</b>	<b>3</b>
2.1 Señales OFDM	3
2.1.1 Generación	3
2.1.2 Características de la señal OFDM	4
2.1.3 5G-NR	4
2.2 Amplificadores de potencia	4
2.2.1 Efectos no lineales	5
2.2.2 Efectos de memoria	6
2.2.3 Caracterización del amplificador de potencia	6
2.3 Sistemas de predistorsión digital	9
2.4 Series de Volterra y su aplicación en el modelado no lineal	9
2.4.1 Serie de Volterra	9
2.4.2 Modelado con las series	10
<b>3 Sistemas de predistorsión digital</b>	<b>13</b>
3.1 Arquitecturas de DPD	13
3.1.1 Arquitectura de aprendizaje indirecto	13
3.1.2 Arquitectura de aprendizaje directo	14
3.1.3 Diferencias entre arquitecturas	15
3.2 Bloques constitutivos del sistema DPD	15
3.2.1 Bloque de generación de la señal predistorsionada	16
3.2.2 Amplificador	16
3.2.3 Normalización de ganancia y cálculo del error	16
3.2.4 Cálculo del vector de pesos, pseudoinversa de Moore-Penrose	16
3.2.5 Bloque de inicialización	17
3.3 Elección de los parámetros óptimos del DPD	17
3.3.1 Ritmo de aprendizaje	17
3.3.2 Número de iteraciones	17
3.3.3 Modelo del sistema	18
3.4 Posibles problemas del DPD	18
3.4.1 No reduce el NMSE, ACPR o EVM	19
3.4.2 Alta PAPR	19
3.4.3 Divergencia	19

<b>4</b>	<b>DPD de doble banda</b>	<b>21</b>
4.1	DPD en banda dual, qué es y qué aporta	21
4.1.1	¿Por qué separar en dos bandas?	21
4.2	Dificultades añadidas del modelo dual	22
4.3	Modelos de doble banda	24
4.3.1	<i>Memory Polynomial</i> (MP)	24
4.3.2	<i>Generalized Memory Polynomial</i> (GMP)	24
4.3.3	<i>2-Dimensional Generalized Memory Polynomial</i> (2D-GMP)	25
4.4	Esquema usado	26
<b>5</b>	<b>Técnicas complementarias del DPD</b>	<b>29</b>
5.1	Técnicas de coeficiente de aprendizaje dinámico	29
5.2	Técnicas de selección de coeficientes	30
5.2.1	Orthogonal Matching Pursuit	30
5.2.2	Doubly OMP	31
5.2.3	Criterio Bayesiano	31
5.3	Técnicas de regularización	33
<b>6</b>	<b>Pruebas experimentales</b>	<b>35</b>
6.1	Banco de medida experimental	35
6.2	Laboratorio de la Universidad de Chalmers	36
6.2.1	Montaje de Chalmers	36
6.2.2	Medidas experimentales	36
6.3	Laboratorio de Sistemas de Radiocomunicación de la Universidad de Sevilla	38
6.3.1	Montaje de Sevilla	38
6.3.2	Medidas experimentales	38
6.4	Uso de técnicas de mejora del DPD dual	39
<b>7</b>	<b>Conclusiones</b>	<b>43</b>
	<b>Apéndice A Códigos MatLab</b>	<b>45</b>
	<b>Apéndice B Modelos 2D-GMP Usados</b>	<b>53</b>
	<b>Apéndice C Artículo Técnicas de CFR</b>	<b>57</b>
	<b>Apéndice D Artículo URSI DPD Dual con Ridge</b>	<b>61</b>
	<i>Índice de Figuras</i>	65
	<i>Índice de Tablas</i>	67
	<i>Bibliografía</i>	69
	<i>Glosario</i>	73

# Notación

---

$\mathbb{R}$	Cuerpo de los números reales
$\mathbb{C}$	Cuerpo de los números complejos
$\ \mathbf{v}\ $	Norma del vector $\mathbf{v}$
$\langle \mathbf{v}, \mathbf{w} \rangle$	Producto escalar de los vectores $\mathbf{v}$ y $\mathbf{w}$
$ \mathbf{A} $	Determinante de la matriz cuadrada $\mathbf{A}$
$\det(\mathbf{A})$	Determinante de la matriz (cuadrada) $\mathbf{A}$
$\mathbf{A}^\top$	Transpuesto de $\mathbf{A}$
$\mathbf{A}^{-1}$	Inversa de la matriz $\mathbf{A}$
$\mathbf{A}^\dagger$	Matriz pseudoinversa de la matriz $\mathbf{A}$
$\mathbf{A}^H$	Transpuesto y conjugado de $\mathbf{A}$
$\mathbf{A}^*$	Conjugado
c.t.p.	En casi todos los puntos
c.q.d.	Como queríamos demostrar
■	Como queríamos demostrar
□	Fin de la solución
e.o.c.	En cualquier otro caso
$e$	número $e$
$e^{jx}$	Exponencial compleja
$e^{j2\pi x}$	Exponencial compleja con $2\pi$
$e^{-jx}$	Exponencial compleja negativa
$e^{-j2\pi x}$	Exponencial compleja negativa con $2\pi$
$\text{Re}$	Parte real
$\text{Im}$	Parte imaginaria
$\text{sen}$	Función seno
$\text{tg}$	Función tangente
$\text{arctg}$	Función arco tangente
$\sin^y x$	Función seno de $x$ elevado a $y$
$\cos^y x$	Función coseno de $x$ elevado a $y$
Sa	Función sampling
sgn	Función signo
rect	Función rectángulo
Sinc	Función sinc
$\frac{\partial y}{\partial x}$	Derivada parcial de $y$ respecto a $x$
$x^\circ$	Notación de grado, $x$ grados.
$\text{Pr}(A)$	Probabilidad del suceso $A$
$E[X]$	Valor esperado de la variable aleatoria $X$
$\sigma_X^2$	Varianza de la variable aleatoria $X$
$\sim f_X(x)$	Distribuido siguiendo la función densidad de probabilidad $f_X(x)$
$\mathcal{N}(m_X, \sigma_X^2)$	Distribución gaussiana para la variable aleatoria $X$ , de media $m_X$ y varianza $\sigma_X^2$

$\mathbf{I}_n$	Matriz identidad de dimensión $n$
$\text{diag}(\mathbf{x})$	Matriz diagonal a partir del vector $\mathbf{x}$
$\text{diag}(\mathbf{A})$	Vector diagonal de la matriz $\mathbf{A}$
SNR	Signal-to-noise ratio
MSE	Minimum square error
:	Tal que
$\stackrel{\text{def}}{=}$	Igual por definición
$\ \mathbf{x}\ $	Norma-2 del vector $\mathbf{x}$
$ \mathbf{A} $	Cardinal, número de elementos del conjunto $\mathbf{A}$
$\mathbf{x}_i, i = 1, 2, \dots, n$	Elementos $i$ , de 1 a $n$ , del vector $\mathbf{x}$
$dx$	Diferencial de $x$
$\leq$	Menor o igual
$\geq$	Mayor o igual
$\backslash$	Backslash
$\Leftrightarrow$	Si y sólo si
$x = a + 3 \underset{a=1}{=} 4$	Igual con explicación
$\frac{a}{b}$	Fracción con estilo pequeño, $a/b$
$\Delta$	Incremento
$b \cdot 10^a$	Formato científico
$\rightarrow_x$	Tiende, con $x$
$\mathcal{O}$	Orden
$\text{TM}$	Trade Mark
$\mathbb{E}[x]$	Esperanza matemática de $x$
$\mathbf{C}_x$	Matriz de covarianza de $\mathbf{x}$
$\mathbf{R}_x$	Matriz de correlación de $\mathbf{x}$
$\sigma_x^2$	Varianza de $x$

# 1 Introducción

---

*El comienzo es la parte más importante de la obra.*

PLATÓN

Este Trabajo de Fin de Grado tratará las técnicas de procesado de señal pertenecientes al círculo de las comunicaciones móviles actuales, concretamente la quinta generación de comunicaciones móviles (5G), empleadas para la reducción de los efectos no lineales de los amplificadores de potencia. Esta generación se caracteriza principalmente por una virtud: la flexibilidad en la asignación de los anchos de banda y frecuencias en el espectro. Pero, a su vez, sufre de fuertes efectos no lineales debidos a la combinación entre el comportamiento de los amplificadores de potencia y las técnicas de modulación usadas. Ambas situaciones sugieren la necesidad de aplicar técnicas de corrección contra no linealidades junto a la capacidad de tener un gran ancho de banda donde operar.

Los amplificadores de potencia, también llamados PA, del inglés *power amplifier*, tienen una característica interesante y es que dependiendo del punto de operación donde se trabaje se obtendrá diferente eficiencia, siendo mayor mientras mayor sea la potencia de salida. Por supuesto, el objetivo principal es obtener el máximo rendimiento del amplificador aprovechando todo lo posible la alimentación que se le está otorgando, y así obtener a la salida una señal de la mayor potencia posible. Este hecho tiene una doble implicación, el disponer de esta mayor potencia no solo permitirá que las ondas puedan llegar al receptor en mejores condiciones para su demodulación, sino que se estará aprovechando mejor la energía empleada, se estará siendo eficiente. Esta es una característica fundamental para las tecnologías del siglo XXI, viviendo en una época de importante transformación ecológica y preocupación por el correcto aprovechamiento de la energía.

Pero trabajar de esta manera, maximizando la potencia de entrada, trae varios inconvenientes. El problema es que mientras mayor sea la potencia que se le aplique al amplificador más efecto tendrán las distorsiones no lineales que afectarán a la señal que esté siendo amplificada. Este efecto deteriorará la señal de varias maneras, por ejemplo, como una deformación de la constelación, efecto que se puede medir mediante la magnitud del vector error (*error magnitud vector*, EVM). Pero este no es el único efecto, puesto que uno de los mayores problemas de este fenómeno es la aparición de componentes espectrales fuera del ancho de banda de la señal original, pudiendo interferir en comunicaciones que estén ocurriendo en canales adyacentes, el cual se mide mediante la relación de potencia del canal adyacente (*adjacent channel power ratio*, ACPR). Habiendo sido nombradas EVM y ACPR es importante comentar la existencia de otra figura de mérito llamada NMSE, de error cuadrático medio normalizado, *normalized mean squared error*, la cual mide el grado de similitud entre dos señales, en el caso que ocupa a este proyecto, la señal original y la distorsionada por el PA.

Debido a estos problemas, varias soluciones fueron propuestas. Algunas de las primeras, y también menos potentes, se limitaban a hacer uso de *backoff*, técnica en la cual se limita el punto de operación, lo que conlleva una reducción en eficiencia, o hacían uso de técnicas de reducción de factor de cresta (CFR, *crest factor reduction*) para reducir la relación potencia de pico a potencia promedio (*peak to average power ratio*, PAPR) y así poder mantenerse en la zona lineal por más tiempo. Con el avance de la tecnología e investigación llegaron las técnicas de predistorsión digital (DPD) cuya filosofía principal es concebir los efectos negativos del amplificador como cierta función invertible de manera que antes de pasar la señal por el dispositivo no lineal se le pueda aplicar la función inversa en el bloque destinado al DPD, de manera que los efectos combinados de ambos se anulen y la salida sea solamente la señal amplificada en potencia, que es el objetivo principal. Para poder invertir la función del amplificador será necesario hacer uso de modelos matemáticos

que aproximen el comportamiento no lineal del amplificador. Es aquí donde entran en juego las series de Volterra, base para los sistemas de predistorsión, las cuales ofrecerán diferentes formas de expresar los efectos del amplificador, tanto las no linealidades como la memoria.

En este trabajo en concreto tendrá como objetivo la implementación de los sistemas DPD de doble banda concurrente, debido a su importancia en los sistemas móviles actuales en su aplicación mediante *Carrier Aggregation (CA)*, empleando una arquitectura de aprendizaje directo para el desarrollo del DPD. El trabajo comenzará con conceptos principales sobre linealización en el Capítulo 2 como la modulación usada o generalidades de los PA, seguido de un análisis en profundidad de los sistemas de predistorsión digital en el Capítulo 3, donde se verán sus principales componentes y arquitecturas. Posteriormente, en el Capítulo 4, se explorará la particularización del sistema a la doble banda concurrente y por ciertos detalles que serán explicados en esa misma sección, será necesario aplicar técnicas complementarias para mejorar el DPD, las cuales serán revisadas en el Capítulo 5. Para acabar, en el Capítulo 6 se mostrarán los resultados obtenidos con diferentes configuraciones propuestas, probadas tanto en el laboratorio de la Escuela Técnica Superior de Ingeniería en Sevilla del Grupo de Sistemas de Radiocomunicación como en el laboratorio online de la Universidad de Chalmers, y se compararán resultados. Para acabar, se cerrará con conclusiones y posibles líneas futuras en el Capítulo 7, haciendo un breve resumen de los resultados del trabajo.

## 2 Linealización

---

*Cree en ti mismo. No en mi creyendo en ti, ni en ti creyendo en mi. Cree en ti y solo ti.*

KAMINA EN TENGEN TOPPA GURREN LAGANN

Para comenzar con la explicación teórica de las partes principales de este Trabajo Fin de Grado, como es el DPD o sus técnicas complementarias, convendría repasar primero los pilares fundamentales sobre los que se sustenta. Estos elementos son el amplificador de potencia, la parte tangible del sistema y causa del problema a solucionar, y el otro la propia señal, en este caso es una modulación OFDM. Sin más dilación, se comentarán las características más importantes de la señal mencionada, seguido de un breve repaso de los efectos negativos causados por los amplificadores para después enlazar con la solución elegida para estos problemas y como ayudará.

### 2.1 Señales OFDM

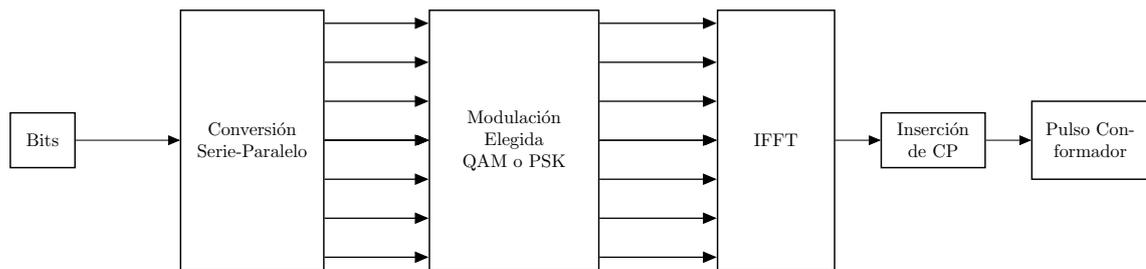
Las señales OFDM [1], del inglés *orthogonal frequency division multiplexing*, surgen al aplicar una técnica de multiplexación a varias portadoras, de manera que se asigne una modulación a cada una de estas y se aplique una transformación que haga que sean ortogonales en frecuencia.

Este tipo de señal ha sido muy importante para la evolución de las comunicaciones inalámbricas, concretamente las móviles. Según avanzaban las tecnologías móviles, de la conmutación de circuitos al uso de la red IP y en adelante crecía a su vez la demanda de mayores tasas y eficiencia espectral. Con el objetivo de poder alojar esa demanda en las frecuencias de radio se adoptó este tipo de multiplexación en la cuarta generación de comunicaciones móviles como la elegida de entre las propuestas. Sus ventajas eran evidentes, no solo ofrecía las tasas y eficiencia requeridas, también aportaba una protección muy eficaz contra el multitrayecto y una gran flexibilidad a la hora de asignar recursos.

#### 2.1.1 Generación

Un aspecto muy importante de la señal OFDM es la generación, debido a que a cambio de sus grandes prestaciones es costosa de crear, motivo por el cual tardó en ver la luz como opción para las comunicaciones inalámbricas. Aun así, el avance de las tecnologías junto con el uso de algoritmos muy eficientes como la FFT [2] (*Fast Fourier Transform*) logró su implementación en los sistemas de comunicaciones, volviéndose desde ese momento la reina en el terreno móvil hasta este momento.

El esquema para la generación de la señal comienza por una etapa de conversión de serie a paralelo, en la cual los bits ya transformados a símbolos según la modulación usada, ya sea PSK o QAM, se agrupan en segmentos de cierta longitud para aplicarles la IFFT (*inverse FFT*). Esta longitud se busca siempre que sea igual a una potencia de dos, maximizando así la eficiencia del algoritmo FFT. Posteriormente a la transformación se vuelve a pasar a serie, incluyendo en el proceso el llamado prefijo cíclico (CP), un segmento del final de cada símbolo que se coloca delante para añadirle periodicidad y detrás de este bloque se aplicaría el pulso conformador para su paso de digital a analógico. De esta manera, ya estaría la señal OFDM lista para convertirla a analógico y ser subida a la frecuencia de portadora elegida.



**Figura 2.1** Bloques principales del transmisor OFDM.

El esquema de generación se muestra en la Figura 2.1, donde se pueden apreciar los bloques principales del transmisor. En el receptor se realizarían las operaciones contrarias, cambiando serie-paralelo por paralelo-serie y viceversa, a su vez que realizando una FFT en vez de su función inversa.

### 2.1.2 Características de la señal OFDM

Se ha nombrado brevemente en la sección anterior, pero aquí se expondrá en profundidad, ¿Qué hace tan especial a la señal OFDM?

Para empezar, es muy eficiente en el manejo del espectro, ya que, definidos unos parámetros de ancho de banda y frecuencia central, la señal se podrá ajustar perfectamente, sin residuos en otras frecuencias. Esta propiedad fundamental será aportada por la ortogonalidad en frecuencia, y le concede una capacidad de ajuste al espectro muy beneficiosa. Esta adaptación se realiza modificando los parámetros en la generación como son el número de portadoras o separación entre portadoras ( $\Delta f$ ) para ajustarse al hueco que se le asigne. Otro detalle que otorga mucho potencial al sistema es la protección ante el multitrayecto, la cual se le concede por medio de la inclusión del prefijo cíclico el cual provoca que la convolución que ocurre al pasar por el canal dispersivo sea cíclica, siendo de esta manera la ecualización muy sencilla. En este caso, una multiplicación a cada una de las portadoras por una constante bastará, eliminando por completo el problema de la dispersión [3].

Pero hay ciertos efectos negativos adicionales aparte de la generación. Uno de ellos es lo sensible que es la ortogonalidad de la señal, muy comprometida por los efectos no lineales del amplificador, pero este es el tema que se tratará justamente en el proyecto. Otro problema es el elevado factor de cresta de la señal, el cual ocurre debido a la coincidencia de los picos de las portadoras que componen la señal. Este efecto puede llegar a ser muy nocivo, sobre todo al entrar en la zona más no lineal, lo que puede llegar a provocar emisiones fuera de banda, por lo cual debe ser controlado. Para reducir estos efectos se emplean técnicas de reducción de factor de cresta, las cuales pueden ser usadas en pareja con el sistema DPD para mejores resultados. Un artículo resumiendo las técnicas más populares que son usadas en señales OFDM como las de este proyecto, realizado por el mismo autor, puede ser encontrado en el Apéndice C.

### 2.1.3 5G-NR

El último detalle que comentar sobre esta modulación tan peculiar es la tecnología de acceso empleada en 5G, la *5G New Radio* [4, 5] (5G-NR) cuyas directrices son seguidas en este Trabajo Fin de Grado a la hora de asignar valores a los parámetros que se usarán en la generación de la señal.

Desarrollada por el 3GPP, define tres casos de uso dependiendo de la finalidad de la conexión: banda ultra ancha, baja latencia o conexión IoT masiva. Esta variedad de situaciones se logra gracias a la flexibilidad de OFDM para utilizar el espectro, definiendo el estándar varios anchos de banda que pueden llegar hasta los 400 MHz, en el caso de necesitar las mayores tasas posibles. Algunas tecnologías complementarias que son usadas en 5G NR son MIMO [6] o CA [7], esta última teniendo una especial importancia en este trabajo, pues el sistema de linealización tendrá que tratar con una señal que emplee CA.

## 2.2 Amplificadores de potencia

Los amplificadores de potencia son dispositivos activos, esto significa que requieren de cierta tensión de alimentación para funcionar. El hecho de tener que ser alimentados provoca que aparezcan efectos negativos debidos a la saturación, inherentes a este tipo de dispositivos. Aún siendo las no linealidades efectos más importantes, ya que serán los que más deterioren la señal, no se deben pasar por alto los efectos de memoria,

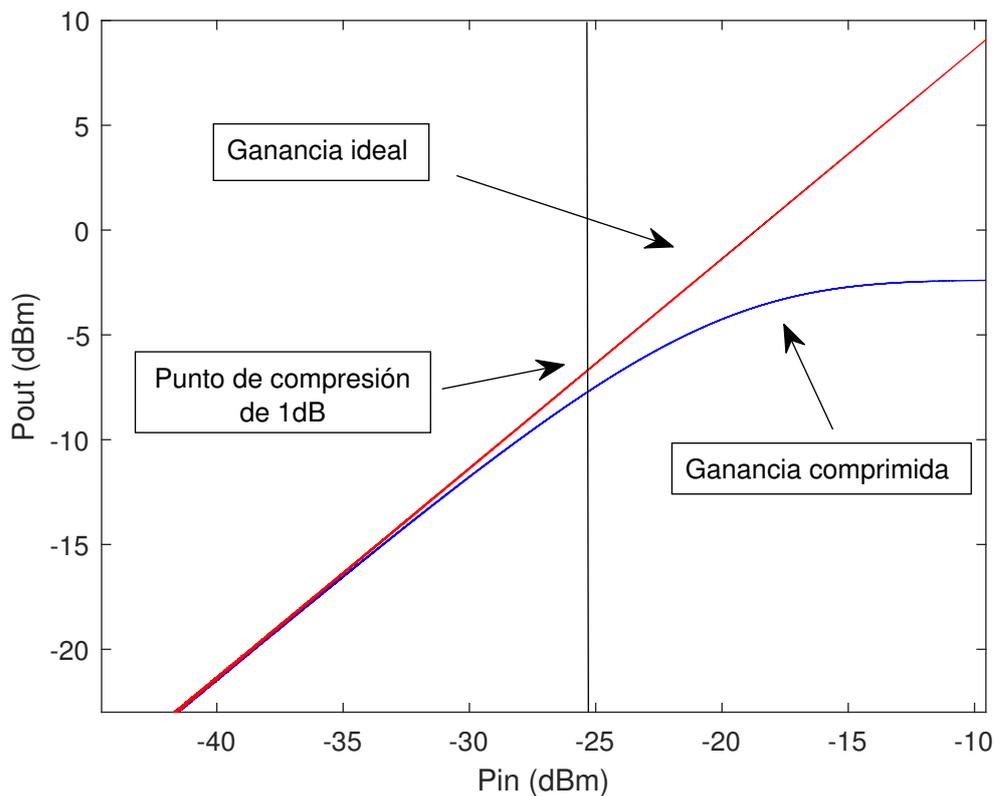
que también tendrán un peso importante en la distorsión. Para poder cuantificar los efectos de ambos tipos de distorsión se usarán ciertas figuras de mérito, las cuales juegan un rol determinante en el sistema y evalúan varios de los escenarios que más preocupan, como las emisiones en otros canales. A continuación, se verán en profundidad estos efectos y la manera de cuantificarlos.

### 2.2.1 Efectos no lineales

Los efectos no lineales son el problema principal que combatir, ya que provocan distorsiones en la señal muy complicadas de mitigar, y por ello requieren de sistemas muy complejos para reducir su influencia. En general, y como se tratará al amplificador como una función dependiente del tiempo y del valor de la señal, se puede definir la no linealidad del amplificador como todo efecto contrario al lineal, el cual sí tiene una definición precisa, siendo lineal toda función que cumpla

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \quad (2.1)$$

donde  $\alpha$  y  $\beta$  son constantes. La propiedad de la linealidad muestra que para toda función que la cumpla podrá ser sumada y escalada siendo estos cambios directamente reflejados en la salida. La complicación principal que acarrea las no linealidades es que generan componentes frecuenciales de la señal que no estaban previamente en esta, lo que trae consecuencias tales como las emisiones fuera de banda.

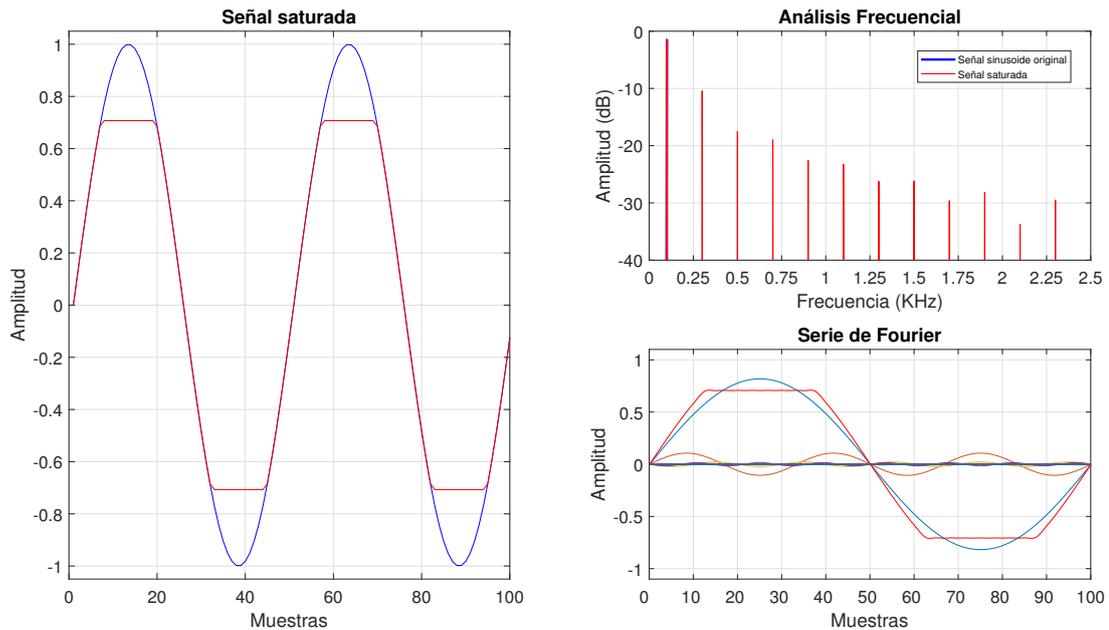


**Figura 2.2** Compresión de ganancia de un amplificador.

La causa de esta problemática se encuentra en el fenómeno de la saturación. Al llegar a las potencias máximas que puede aportar el amplificador se satura, es decir, se comprime la ganancia, muestra de que no dispone de más energía para entregar. Este concepto es fácilmente ilustrable mediante una gráfica relacionando potencia de entrada y salida, como la Figura 2.2, donde se aprecian claramente las zonas donde la ganancia es constante, representada por la pendiente de la recta mientras que cuando comienza a disminuir entra en saturación, encontrando el punto crítico en el punto de compresión de 1dB. Esta parte tan concreta, donde la ganancia ideal y la real se separan justamente 1 dB, es la frontera que separa lo que se considera lineal frente

a la zona no lineal. Es de apreciar que mientras más se adentre uno en la zona no lineal mayores serán sus efectos puesto que mayor será la diferencia entre ideal y real.

Los efectos provocados por este fenómeno son más sencillos de observar aplicados a una señal sinusoidal, estudiando su amplitud y componentes frecuenciales. La señal original se mantiene, pero reduce su amplitud. Por otra parte, se crean componentes espectrales en las frecuencias múltiplos de la original, las cuales van decayendo mientras mayor sea su frecuencia, cada vez menor amplitud, aunque se crean infinitos términos. Toda esta casuística se puede observar en la Figura 2.3. Estos armónicos serían equivalentes a las emisiones fuera de banda en el sistema OFDM, ocasionando recrecimiento espectral, mientras que los efectos dentro de la banda, es decir, el escalado, no será constante dentro de ella, por lo que el efecto llegará a ser aún más negativo.



**Figura 2.3** Comparativa de los efectos de la saturación en tiempo y frecuencia y descomposición de la señal saturada en frecuencias puras por serie de Fourier.

### 2.2.2 Efectos de memoria

El otro efecto importante causado por los amplificadores es el de la memoria. Se entiende como memoria a la influencia de los valores anteriores de la señal en el instante actual, pudiendo afectar de manera constructiva o destructiva. Este tipo de efecto tiene varias implicaciones, como la dispersión de la constelación o de la ganancia y es variable dependiendo del amplificador que se esté usando. El efecto que causará la memoria de cierta muestra se irá desvaneciendo según vayan pasando las muestras, lo cual se denomina *fading*, y hará que a partir de cierto número el efecto sea tan pequeño que no se tenga en cuenta, por lo que se define un parámetro para considerar el máximo de muestras de memoria que podrán afectar a la señal de forma notable. Este concepto se llama *memory depth* y tendrá una gran importancia a la hora de configurar los modelos, que se verán posteriormente.

Dentro de los efectos de memoria se diferencian dos tipos, uno a corto plazo, causado por los componentes físicos que forman el amplificador tales como las capacidades o semiconductores, y el otro a largo plazo, causado por ejemplo por el cambio en la temperatura del dispositivo dependiendo del tiempo que lleve en funcionamiento.

### 2.2.3 Caracterización del amplificador de potencia

Una vez está definido cual es el impacto tanto de las no linealidades como de la memoria, es necesario cuantificarlo. Esto ayudará a realizar un seguimiento del rendimiento del sistema DPD para ver si realmente está funcionando como se desea, reduciendo el error y las emisiones fuera de banda, y será usado por el

propio sistema como parámetros objetivo para mejorar buscando su reducción. Los parámetros o figuras de mérito que se usarán en conjunto con estos sistemas son:

- **AM-AM:** La característica AM-AM relaciona la potencia instantánea de entrada con la de salida, tomando la forma de nube de puntos como se puede observar en Figura 2.4. El efecto de la saturación se observa de forma similar a como se veía en la característica de ganancia del amplificador, después de la recta con pendiente igual a la ganancia del amplificador haya una zona donde bruscamente cae la pendiente hasta volverse igual a cero, la zona más no lineal y de saturación. La dispersión en la nube se debe a los efectos de memoria, por lo que la gráfica perfecta sería una pendiente para todo valor del eje x con la mínima dispersión posible.

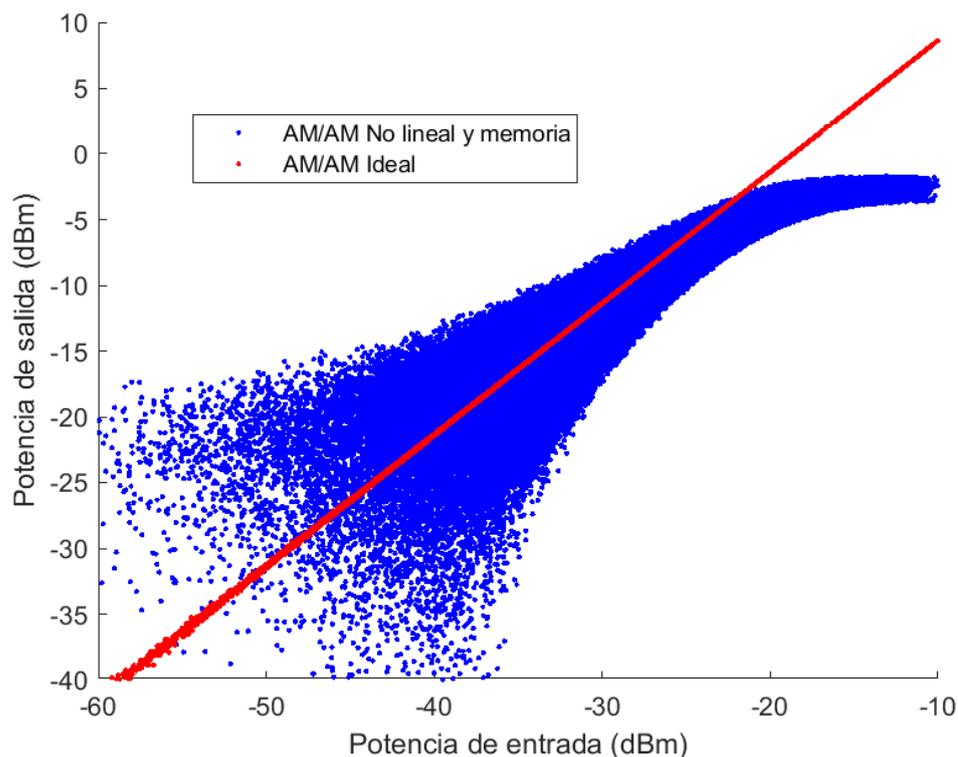


Figura 2.4 Característica AM-AM.

- **Punto de compresión de 1 dB:** Representa la potencia a partir de la cual se considera que los efectos no lineales no pueden ser despreciados. Se calcula buscando el punto en el cual la diferencia entre la ganancia ideal, la pendiente sin saturación extrapolada de la zona lineal, y la ganancia real del dispositivo se diferencian en 1dB.
- **Relación de potencia con el canal adyacente (ACPR):** Mide la potencia que se filtra del canal principal a los adyacentes en el espectro. Este efecto es producido principalmente por las no linealidades, fruto de la saturación que crea componentes fuera del ancho de banda de la señal. En el caso de señales de banda ancha, como es OFDM, a estas emisiones se les llama recrecimiento espectral, uno de los parámetros más importantes a tener en cuenta sobre todo por las exigentes especificaciones 3GGP para las señales bajo el estándar 5G-NR. Aunque todavía no se ha entrado en la cuestión de la doble banda merece la pena comentar como van a estar definidos los canales en ese caso y su notación. Esto se puede ver en la Figura 2.5, donde el subíndice  $C_L$  o  $C_R$  indica que es referido a la señal de la banda izquierda o derecha, mientras que el  $C_{m1L}$  indica el desplazamiento desde la banda indicada,  $L$  o  $R$  a la que se estudia, en este caso  $m1$  significa menos uno, por lo que se refiere al canal que esta adyacente a la izquierda de la señal concurrente de la izquierda.

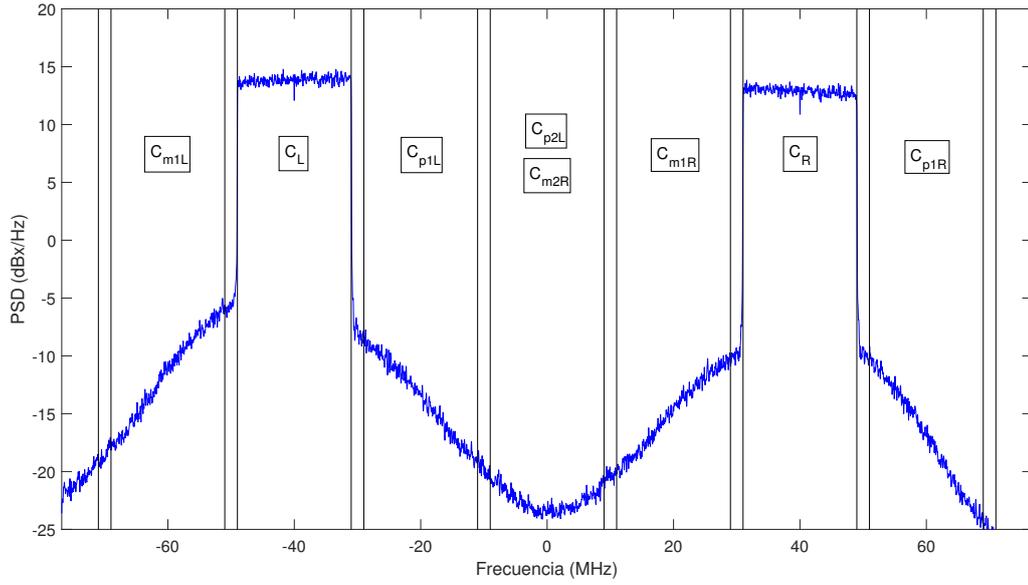


Figura 2.5 Notación de los canales en doble banda concurrente.

$$\text{ACPR}_{nC} = \frac{\int_{F_{off} \pm nC - B/2}^{F_{off} \pm nC + B/2} Y_o(f) df}{\int_{F_{off} - B/2}^{F_{off} + B/2} Y_o(f) df} \quad (2.2)$$

La letra  $Y_o$  denota la salida del sistema en frecuencia, mientras que la integral se realiza en el canal correspondiente. Los límites de las integrales se corresponden a, en el denominador, el ancho de la señal centrada, mientras que en el numerador los límites son los extremos de la señal desplazada al canal  $nC$ .

- **Magnitud del vector de error (EVM):** Muestra la diferencia promedio entre los símbolos de las constelaciones demoduladas que contiene la señal OFDM con los símbolos transmitidos, expresado en porcentaje. Este parámetro está sujeto tanto a efectos no lineales como a la memoria del sistema. Su expresión es

$$\text{EVM} [\%] = \sqrt{\frac{\sum_{l=1}^L |\hat{y}_l - \hat{x}_l|^2}{\sum_{l=1}^L |\hat{x}_l|^2}} \times 100. \quad (2.3)$$

- **Error cuadrático medio normalizado (NMSE):** Valor que típicamente es usado en dB debido a las pequeñas cantidades que se suelen manejar en estos sistemas, relaciona dos señales mediante el parecido de ambas, calculado como una resta de los vectores de ambas señales al cuadrado, al que se le realiza la media. Esta figura de mérito muestra el error total de la señal, tanto en banda como fuera de ella.

$$\text{NMSE} [\text{dB}] = 10 \log \left\{ \frac{\sum_{k=1}^N (y_{ideal}[k] - y_{med}[k])^2}{\sum_{k=1}^N |y_{med}[k]|^2} \right\}. \quad (2.4)$$

## 2.3 Sistemas de predistorsión digital

Conociendo tanto las características de la señal de entrada como las de la fuente de error a estudiar, el amplificador, y habiendo definido sus parámetros más importantes ya están listos los ingredientes principales para la pieza fundamental del trabajo, la predistorsión digital. Esta técnica tiene como objetivo reducir al mínimo los efectos no lineales y de memoria, de esta manera consiguiendo que el NMSE, ACPR y EVM logren los valores más pequeños posibles, reduciendo la probabilidad de error y las fugas de potencia a canales adyacentes.

El pilar principal de esta técnica se basa en la inversión del efecto de la característica tanto estática como dinámica del amplificador. Se supondrá que el amplificador es igual a cierta función, la cual da lugar a dos efectos principales, uno de ellos el deseado: la amplificación de la señal, mientras que el otro serían los no deseados tales como las no linealidades y la memoria. Mediante la acción del bloque DPD [8], en inglés *Digital Pre-Distortion*, colocado antes del amplificador la señal sufrirá una distorsión, la cual se buscará que sea lo más parecida posible a la inversa de la distorsión propia del amplificador. De esta manera, a la salida de todo el grupo se anularán las distorsiones entre ellas al ser iguales pero invertidas dando lugar a una versión de la entrada con la amplitud igual a la entrada original, pero viendo su amplitud incrementada por la ganancia del amplificador.

Un tema que se ha pasado un poco por alto es el ruido, que por supuesto estará presente en todo momento en el sistema. Pero aun así, para estos sistemas no afectará mucho ya que se podrá evitar que este problema afecte a las medidas y por tanto se pueda realizar una buena predistorsión mediante la realización de medidas varias veces con la misma señal para posteriormente hacerles un promediado, de esta manera reduciendo la potencia de ruido.

Este concepto de predistorsión trae consigo dos partes principales, la identificación y el modelado. La primera de estas trata de encontrar la combinación que mejor resultados ofrezca en la minimización de los parámetros básicos como el NMSE. Si este fuera un caso sencillo, es decir, lineal, el proceso sería más simple puesto que se convertiría en un problema de ecualización. Al no ser el caso, será necesario el uso de herramientas más potentes, aquí es donde entra el concepto de modelado de la mano de los regresores y las series de Volterra. Ya que el comportamiento que se pretende invertir es no lineal, no bastará con buscar relaciones entre las componentes de la señal, habrá que crear nuevas componentes sobre las que trabajar y los cuales serán los regresores. Estos serán señales que contendrán información no lineal de la señal en cuestión y a partir de combinaciones ponderadas de estos se modelará la señal que se necesite según el esquema. Las componentes se pueden crear de muchas maneras, pero la base principal para todas serán las series de Volterra.

## 2.4 Series de Volterra y su aplicación en el modelado no lineal

Las series de Volterra [9] han sido históricamente la base de los modelos de regresión del DPD, gracias a su buen ajuste a las distorsiones causadas por el amplificador. La implementación principal de serie de Volterra, en su versión discreta, puede verse como una generalización de las series de Taylor, pero añadiendo efectos de memoria, que sumado a la capacidad de ambas series de adaptarse a los comportamientos no lineales, aporta exactamente lo que requiere el sistema DPD.

### 2.4.1 Serie de Volterra

La forma exacta de la serie de Volterra para señales reales es la que se muestra en la siguiente ecuación,

$$y[k] = h_0 + \sum_{n=1}^{\infty} \sum_{\mathbf{q}_n=0}^{\mathbf{Q}_n} h_n[\mathbf{q}_n] \prod_{r=1}^n x[k - q_r], \quad (2.5)$$

donde  $h_n[\mathbf{q}_n]$  representa el kernel de Volterra, se tienen en cuenta hasta  $n$  órdenes y  $\mathbf{q}_r = 0, 1, \dots, Q_n$  para cada valor de  $r$  donde  $\mathbf{Q}_n$  es el vector donde se almacenan los retrasos.

En cambio, la que se usará en este proyecto debido a que se trabajará con las envolventes complejas es de la forma:

$$y[k] = \sum_{m=0}^{\infty} \sum_{q_{2m+1}=0}^{Q_{2m+1}} h_{2m+1}[q_{2m+1}] \prod_{r=1}^{m+1} x[k - q_r] \prod_{r=m+2}^{2m+1} x^*[k - q_r], \quad (2.6)$$

siendo  $n = 2m + 1$  el orden no lineal. Esto significa que solo los términos impares son tomados en cuenta en esta ecuación [10], pero como se demuestra en [11] pueden ser usados también los pares para mejor rendimiento. De esta otra manera serán implementados los modelos usados en este proyecto.

La longitud del vector  $\mathbf{h}$  depende del orden máximo de los regresores y de la memoria que se le asigne, existiendo un regresor por coeficiente cuya función será escalarlos de manera que la suma de todos ellos sea igual a la señal objetivo. Una elección muy importante para el sistema y que se debe hacer a priori es la elección del modelo, concretamente la elección de estos parámetros, que se realizará de forma experimental hasta encontrar el que de mejor rendimiento.

#### 2.4.2 Modelado con las series

De esta manera, cada combinación concreta de los parámetros anteriormente expuestos dará lugar a un regresor en concreto. El conjunto de todos para los límites establecidos será llamado modelo y tendrá una importancia crucial a la hora del proceso del modelado. Dependiendo del comportamiento a modelar, es decir, el amplificador que se use, serán más útiles unos regresores que otros, por lo que interesa saber cuáles serán las más importantes para la señal a modelar en concreto. Para este cometido se podrá hacer uso de diferentes técnicas de selección, que se verán en profundidad en el Capítulo 5. Otro importante detalle es la cantidad de regresores del modelo, mayor cantidad implicará más variedad y riqueza para modelar, pero debido a que todos se crean a partir de la misma señal con modificaciones, se parecerán bastante, por lo que puede hacer que aparezcan problemas numéricos en ciertas partes del sistema DPD. Este problema se verá más avanzado en el trabajo.

Una forma muy conveniente de ver el modelado es mediante un espacio de señal. Las bases de este serán los regresores creados por el modelo, esto es, tendrá tantas dimensiones como regresores haya. En general, la señal objetivo tendrá muchas más dimensiones, pero no será necesario que sea exactamente igual, sumado a que, al añadir cada vez más regresores, su peso en la suma total es menor, aportando cada vez menos en la reducción del error. Un ejemplo en un caso más reducido, tres regresores, se muestra en la Figura 2.6, donde se modela la señal  $u$ .

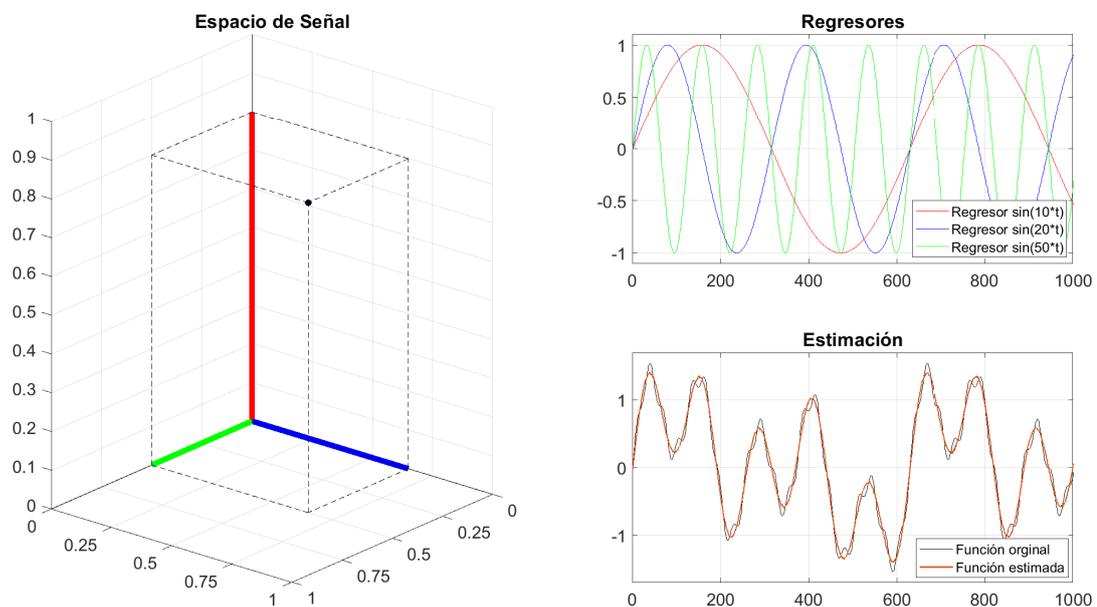


Figura 2.6 Modelado en el espacio de señal.

Este caso es una versión muy simplificada del modelado. En la figura de título Regresores se muestran las señales usadas como base para la regresión, para estimar cierta señal, la cual es en este caso una combinación lineal de cuatro señales sinusoides, bastante más reducido que en el caso más general que se verá en este Trabajo Fin de Grado.

De esta manera, mediante una regresión de mínimos cuadrados se calculará el peso de cada componente en la salida del sistema. En la figura del espacio de señal se muestra en cada eje un regresor (de forma ortogonal,

aunque esto no es así, pero facilita la representación) y en colores la importancia de cada regresor para el peso final. El punto es la combinación concreta de pesos que define a la señal en este subespacio creado por los regresores. Mediante este método se consigue modelar la señal, pero solo las componentes de esta cuya información se tenía a priori en los regresores, y dejando un ligero error debido a la señal sinusoidal desconocida. Este suceso será similar a lo que se encontrará con los regresores en un PA, se podrá modelar gran parte de la señal pero habrá ciertas componentes que no estarán presentes debido a la gran cantidad de coeficientes que serían necesarios para esto.



## 3 Sistemas de predistorsión digital

---

*No quería ser ya el más grande, el más fuerte o el más inteligente. Todo eso lo había superado. Deseaba ser querido como era, bueno o malo, hermoso o feo, listo o tonto, con todos sus defectos... o precisamente por ellos.*

LA HISTORIA INTERMINABLE, 1979

Habiendo revisado ya varios conceptos sobre el proceso general es momento de presentar el núcleo del trabajo que es el bloque de predistorsión digital. Como se ha ido comentando anteriormente, este sistema precederá al amplificador y aplicará cierta distorsión en la señal para anular los efectos negativos que cause. Las técnicas de predistorsión son muy variadas y han sufrido una constante evolución desde los años 30s cuando empezaron a aparecer los primeros sistemas [12, 13] y asentaron bases de lo que sería. El primer tema que tratar es como se conforma el esquema del DPD, que elementos lo constituyen, donde interviene el concepto de arquitectura y como está estructurado el modelo.

### 3.1 Arquitecturas de DPD

Actualmente existen dos arquitecturas principales bien diferenciadas para los sistemas de predistorsión: de aprendizaje directo e indirecto. Ambas toman las bases principales del DPD, como es el modelado y la inversión de la característica no lineal, pero las interpretan de maneras muy diferentes, como se expondrá a continuación. Los bloques constituyentes serán explicados en profundidad en la sección siguiente.

#### 3.1.1 Arquitectura de aprendizaje indirecto

Esta arquitectura se traduce en un algoritmo secuencial en el que se pretende modelar el comportamiento del amplificador de una pasada. De esta manera la señal de entrada del amplificador será la de información pasada por la función estimada inversa del amplificador, tal que

$$\mathbf{u} = \mathbf{X}\mathbf{h}, \quad (3.1)$$

siendo  $\mathbf{X}$  la matriz de regresores creada a partir de la señal de información,  $\mathbf{h}$  el vector de pesos y  $\mathbf{u}$  la entrada del amplificador de potencia. Mientras que el cálculo del vector de pesos se realiza de la forma

$$\mathbf{h} = \mathbf{Y}^\dagger \mathbf{u}, \quad (3.2)$$

donde  $\mathbf{u}$  y  $\mathbf{h}$  son los mismos vectores que en la ecuación (3.1), mientras que  $\mathbf{Y}$  es la matriz de regresores pero creada a partir de la salida [14]. La operación se realiza así, modelar la entrada con la salida, para que los valores de los coeficientes obtenidos sean la inversa del amplificador.

Una de las diferencias fundamentales entre los dos tipos de aprendizajes es el hecho de que es secuencial, al contrario que el directo, el cual es iterativo. Esta diferencia aportará ciertas ventajas, a costa, como siempre, de otros detalles negativos. También es notable destacar la notación utilizada, la cual es diferente para las dos arquitecturas ya que representan conceptos distintos. La notación del esquema indirecto es más clásica, por ejemplo, al vector de coeficientes  $\mathbf{h}$  se le conoce de esa manera por que representa el kernel de Volterra,

llamándose igual. De hecho, al propio vector en el esquema indirecto es más común denominarlo como vector de coeficientes, mientras que en el otro esquema es vector de pesos.

El esquema indirecto es mostrado en la Figura 3.1, donde se hace hincapié en el hecho de la secuencialidad, la no existencia de bucles.

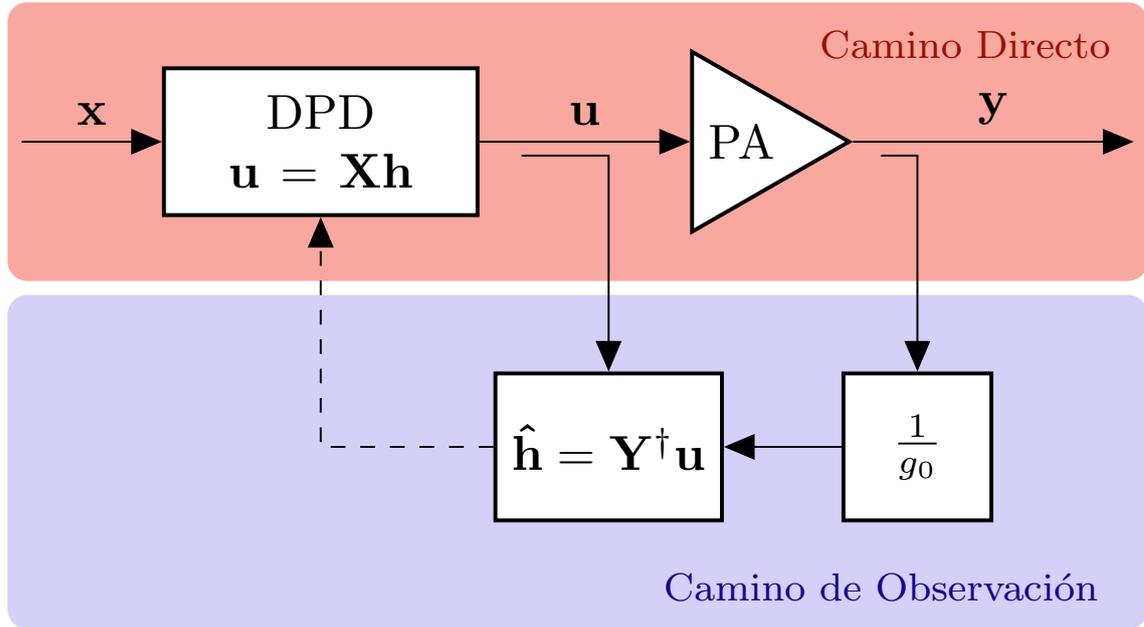


Figura 3.1 Esquema aprendizaje indirecto.

Se puede observar cómo es un camino de una sola dirección, que comienza con la entrada aplicándole un DPD cuyos coeficientes están a cero, excepto el de la señal original sin modificaciones, para ser medida posteriormente y de ahí sacar la inversa del amplificador en un único cálculo.

### 3.1.2 Arquitectura de aprendizaje directo

La arquitectura directa es un proceso iterativo [15, 16]. Esta consta de un bucle principal donde se calculará la señal de entrada existiendo dos configuraciones diferentes, para después volver a pasar por el PA y calcular la actualización de los pesos.

De esta manera, las dos posibles ecuaciones que gobiernan la generación de la entrada del amplificador,  $\mathbf{x}$ , son:

$$\mathbf{x} = \mathbf{U}\mathbf{w}, \quad (3.3)$$

$$\mathbf{x} = \mathbf{u} - \mathbf{U}\mathbf{w}, \quad (3.4)$$

donde  $\mathbf{u}$  representa la señal de entrada para ambos,  $\mathbf{U}$  es la matriz de regresores en ambos casos y  $\mathbf{w}$  el vector de pesos, donde se almacena la información del DPD.

El esquema de las dos configuraciones se puede observar en la Figura 3.2, donde en el bloque de DPD se muestran ambas opciones, aunque normalmente es más utilizada la mostrada en la ecuación (3.4), debido a varias razones, entre ellas la mayor libertad a la hora de calcular los coeficientes o la mejor diferenciación entre los conceptos de aprendizaje indirecto y directo. Este último detalle se debe a la suma de un nuevo concepto en el DPD, la llamada señal de distorsión, denotada por  $-\mathbf{U}\mathbf{w}$ , y donde se volcará todo el cálculo del algoritmo.

Como se puede observar, hay varias diferencias con el aprendizaje indirecto, donde lo más reseñable es el diseño en bucle. Esta estructura está más inspirada en los algoritmos típicos de *machine learning*, la existencia de un bucle de entrenamiento, un parámetro  $\mu$  para controlar la velocidad de aprendizaje o el uso de la técnica del descenso del gradiente para el cálculo del error. Una de las influencias de este hecho se puede observar en la notación, específicamente en el vector de coeficientes que en vez de llamarse  $\mathbf{h}$  como en el indirecto es  $\mathbf{w}$  de *weights*, en inglés pesos.

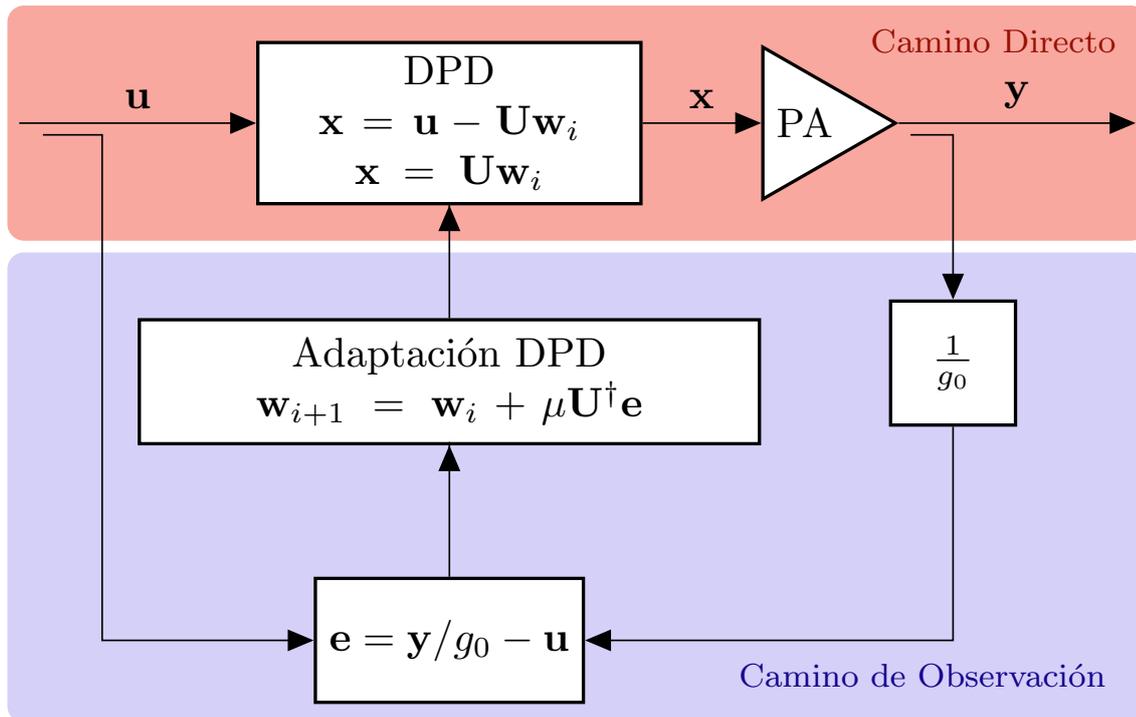


Figura 3.2 Esquema aprendizaje directo.

### 3.1.3 Diferencias entre arquitecturas

Al presentar las dos arquitecturas típicas se han comentado por encima algunos detalles que las hacen diferentes, como el tema de la secuencialidad o iteración. Estas características son muy interesantes y pueden aportar información sobre por cual decantarse en determinadas ocasiones en las que los requisitos del sistema son muy ajustados.

Primero, como no puede ser de otra forma, ¿qué ventajas e inconvenientes trae el hecho de que el aprendizaje directo sea iterativo con respecto al indirecto? En general, la arquitectura de aprendizaje directo da lugar a mayor número de bloques y complejidad, además de que, debido a tener que recorrer varias veces el bucle hasta dar por válida la solución, el tiempo de ejecución se incrementa bastante. Por su parte, el aprendizaje indirecto teniendo en comparación menos bloques y menor tiempo de ejecución, es menos preciso que su contraparte directa, ofreciendo peores resultados en cuanto a linealización, es decir, valores de NMSE, EVM, ACPR y demás. Otro punto a favor del esquema directo es que gracias a ser iterativo y tener un parámetro para controlar el aprendizaje, puede manejarse mejor la divergencia del algoritmo, reduciéndolo en un caso de alta divergencia, cosa que no se puede hacer en el indirecto ya que su adaptación es similar a una sola iteración del aprendizaje directo con  $\mu$  tomando el valor 1 [17].

Por dichas razones, teniendo en consideración que en general tanto los algoritmos de DPD como las técnicas que se usan en conjunto a ellos son *greedy*, esto es, computacionalmente costosas, la elección para este trabajo ha sido la arquitectura directa, ya que se busca la mejor linealización aunque sea muy exigente. Esto ha sido debido a los mejores resultados que ofrece y siendo que el hecho de que sea más costosa no marcará un gran impacto comparado con las demás técnicas, que tomarán prácticamente todo el tiempo de ejecución del algoritmo.

## 3.2 Bloques constitutivos del sistema DPD

En esta sección se comentarán los bloques que componen el sistema DPD de aprendizaje directo en profundidad, puesto que el funcionamiento de ellos es muy importante para el proceso general y tienen muchos detalles por separado. Estos son el bloque de DPD, amplificador, cálculo del error, adaptación del DPD e inicialización.

### 3.2.1 Bloque de generación de la señal predistorsionada

Esta parte del sistema recibe la señal de entrada original junto con el vector de pesos, que junto con a la matriz de regresores generará la señal de entrada, predistorsionada. El funcionamiento del bloque es muy directo, lo cual será útil para explicar conceptos básicos.

Primero su función, la predistorsión, que se rige por la ecuación (3.4), ya vista en la sección anterior, donde  $\mathbf{x}$  será la señal predistorsionada, entrada del amplificador y que se espera que reduzca los efectos no lineales. La matriz  $\mathbf{U}$  estará conformada por los regresores elegidos según el modelo correspondiente, los cuales tendrán la misma longitud que la señal de entrada, aunque se pueden aplicar técnicas de selección de índices para en el momento del cálculo no necesitar la señal completa y de esta manera facilitar los cálculos, a costa normalmente de algo de deterioro en el NMSE. El vector  $\mathbf{w}$  contiene los valores que escalarán los regresores para sumarlos todos y generar la señal de distorsión, la cual se restará a la entrada para obtener la señal final que pasará por el amplificador.

### 3.2.2 Amplificador

No hay mucho mas que decir de él a parte de lo ya comentado. Dispositivo activo y fuente de efectos no lineales, memoria y ruido. Aumentará el nivel de la señal según su ganancia y su eficiencia será mayor mientras mayor sea la potencia de entrada en este.

### 3.2.3 Normalización de ganancia y cálculo del error

Estos pasos son breves y contribuyen ambos a generar la señal que se usará para el cálculo de los pesos. En esta parte lo primero será normalizar la ganancia, para posteriormente restar esta señal a la de entrada. Estos pasos se pueden resumir con la ecuación (3.5):

$$\mathbf{e} = \frac{\mathbf{y}}{g_0} - \mathbf{u}, \quad (3.5)$$

operación que es bastante directa. El error es represetado por  $\mathbf{e}$ ,  $\mathbf{y}$  es la salida del amplificador y  $\mathbf{u}$ , la señal original de entrada. En cambio, el parámetro  $g_0$ , que corresponde a la objetivo considerada para el amplificador si necesita de una operación previa. Una de las cuestiones es cuál es la forma correcta de calcularlo, donde por lo general se considera el valor de la ganancia promedio según la ecuación (3.6):

$$g_0 = \frac{\|\mathbf{y}\|_2}{\|\mathbf{u}\|_2}. \quad (3.6)$$

Una alternativa puede ser, en vez de considerar la ganancia como un parámetro estático, inicializado una vez al comienzo del bucle y que no cambie más, o tomarla como uno dinámico que sea modificado a cada iteración, recalculándose. Por lo general, se optará por la primera opción al ser más estable [18, 19], propiedad que en este esquema tan propenso a la divergencia es muy deseable.

### 3.2.4 Cálculo del vector de pesos, pseudoinversa de Moore-Penrose

Tras el cálculo del error hace falta calcular cual será la actualización de los pesos, la cual vendrá determina por la propia señal error, el parámetro de aprendizaje  $\mu$  y la pseudoinversa de Moore-Penrose [20, 21], de la manera que se puede ver en la ecuación (3.7):

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu \mathbf{U}^\dagger \mathbf{e}, \quad (3.7)$$

donde  $\mathbf{U}^\dagger$  representa la pseudoinversa de Moore-Penrose. Esta pseudoinversa es la generalización de la matriz inversa y se usa para realizar el cómputo de los mínimos cuadrados, donde sería necesario invertir la matriz de regresores  $\mathbf{U}$ , pero al no ser cuadrada no se puede realizar con el método tradicional. La pseudoinversa se define en general como:

$$\mathbf{A}^\dagger = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H, \quad (3.8)$$

donde  $H$  indica hérmitica. Este procedimiento juega un papel fundamental a la hora del cálculo de los coeficientes. Se puede observar en la ecuación (3.7) de actualización de pesos, que lo que se esta usando es, en esencia, un algoritmo de descenso del gradiente [22], minimizando el error  $\mathbf{e}$ . Es de notar que este es el cálculo que causa la divergencia en estos sistemas, ya que al invertir la matriz  $(\mathbf{A}^H \mathbf{A})$  debido a la alta

correlación de los regresores el número de condición será muy alto y esto causará problemas numéricos. Para resolver este problema será usada una técnica de regularización que reducirá el número de condición de tal matriz para obtener mejores resultados, llamada regularización de Ridge y que será explicada en el Capítulo 5.

### 3.2.5 Bloque de inicialización

Para acabar, debe ser nombrado el bloque inicial, la parte del sistema donde se preparan todas las variables y procesos previos que se realizan antes de entrar en el bucle. Este bloque es normalmente olvidado a la hora de describir el funcionamiento general del proceso, pero contiene bastantes detalles interesantes a la hora de su correcto funcionamiento.

Por supuesto, todas las variables son creadas y se les asigna un valor inicial, como por ejemplo los vectores de pesos, que de forma general comienzan valiendo cero, pudiéndose aplicar técnicas para tener un *hot start*, como en el caso visto en [23], o el parámetro de aprendizaje a el cual le es asignado un valor inicial pero que con otras técnicas también puede ser modificado dinámicamente, lo que es llamado paso adaptativo. Adicionalmente, otros parámetros como el número de iteraciones del bucle o las medidas por iteración también serán decididas a priori.

Puesto que en la arquitectura directa la pseudoinversa,  $A^\dagger$  ecuación (3.8), es igual en todas las iteraciones también podrá ser calculada previamente y solamente acceder a su valor en memoria cuando se necesite.

También en este bloque son generadas las señales de entrada del sistema y los regresores según el modelo elegido. De esta manera, las técnicas de selección de regresores, regularización o selección del número óptimo de regresores también serán aplicadas en este bloque.

## 3.3 Elección de los parámetros óptimos del DPD

Conociendo la estructura del DPD falta conseguir elegir en los valores correctos que hagan que el sistema pueda desarrollar todo su potencial. Los parámetros principales cuya configuración es libre son el *learning rate* y el número de iteraciones y de medidas, mientras que otro gran factor importante que se debe configurar es el modelo, cuál usar y cuántos y que regresores contendrá.

### 3.3.1 Ritmo de aprendizaje

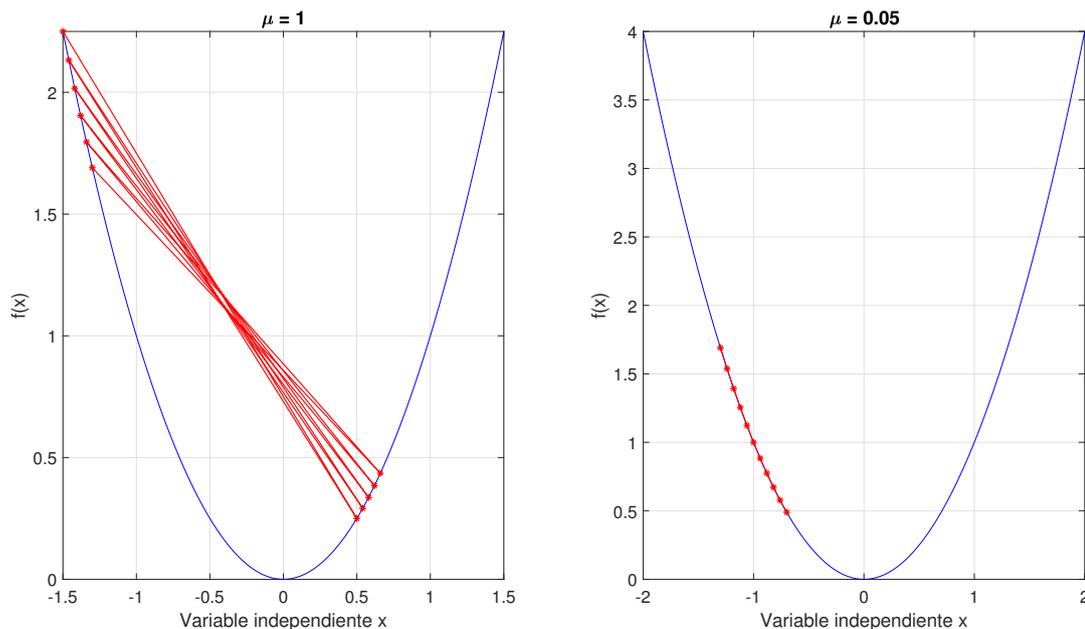
El ritmo de aprendizaje, también llamado tasa o parámetro es una ponderación que se aplica a la actualización de los pesos para controlarla. Esto afectará a la velocidad de convergencia del algoritmo, y en algunos casos provocará su divergencia.

La elección de este coeficiente es un problema típico del *machine learning* [24], que es de suma importancia para cualquier tipo de proceso y puede tomar cualquier valor de 0 a 1. Los problemas que puede acarrear la mala elección de este parámetro pueden ser, si se es muy conservador y se elige un valor muy pequeño necesitar muchas iteraciones para que el algoritmo llegue al punto donde es mínimo el error. Por el contrario, si se selecciona demasiado grande puede ser que nunca se encuentre el mínimo puesto que los *steps*, pasos entre cada iteración, salten alrededor del mínimo y se obvie el mínimo. En la Figura 3.3 se pueden observar ambos efectos en un descenso de gradiente bidimensional, el caso que ocupa en este trabajo es mucho mayor, pudiendo llegar en algunos casos a las quinientas dimensiones. En la gráfica de la izquierda el valor del parámetro es muy grande, lo cual causa el problema de no poder acercarse lo suficiente al mínimo local, mientras que en la gráfica de la derecha es al contrario, podría acercarse con mucha precisión pero debido a la lenta convergencia necesitaría el doble de iteraciones. Estos casos empiezan a sugerir la utilidad de un  $\mu$  dinámico con valores altos al principio y que decaigan con el número de iteraciones.

De todas formas, la elección de este parámetro se realiza de forma arbitraria si no se emplean algoritmos de búsqueda dinámica o adaptativa, por lo que es necesario tener aún más cuidado en el momento de ser usado. Afortunadamente, si el sistema DPD se usa junto con técnicas de regularización o selección de número óptimo de regresores, reduciendo el número de estos y así la divergencia se puede aumentar más el valor de  $\mu$  sin temor a diverger de la solución.

### 3.3.2 Número de iteraciones

Otro parámetro importante al que no se le había prestado atención hasta ahora es el número de iteraciones, este controlará cuántas veces será ejecutado el bucle principal [25].



**Figura 3.3** Efectos de valores de  $\mu$  extremos.

Teniendo un significado tan directo, su elección también lo es: simplemente se basa en seleccionar un número arbitrario para evaluar qué tal se ha desenvuelto el sistema y darle mayor valor si es insuficiente. De manera similar a la tasa de aprendizaje, un número muy alto puede causar divergencia y el uso de ciertas técnicas que eviten este suceso hace que pueda emplearse un valor muy alto, pero generalmente no será necesario pues el sistema llegará normalmente a un punto en el cual no pueda mejorar más, y donde es innecesario seguir iterando. Este punto no se conoce a priori por lo que se pueden aplicar criterios de parada sencillos en vez de simplemente ejecutar el bucle el número de iteraciones definido, pasando este valor a ser un máximo por si no se alcanzan los valores deseados. Los criterios de parada comúnmente tienen como principal parámetro que valorar el NMSE, aunque también se podrían usar otros como el ACPR o EVM si se busca especialmente mejorar una característica en concreto. Con ellos en mente se marcará cierto umbral y en el momento que se cumple se parará el algoritmo.

### 3.3.3 Modelo del sistema

La elección del modelo del sistema y sus regresores es un tema bastante más complejo que los dos anteriores, y cuenta con bastante profundidad en todas sus partes. Para empezar, se necesita saber que modelo usar, generalmente basados en series de Volterra, el número de regresores y cuáles son los mejores.

En cuanto al modelo, como se ha dicho, suelen ser modelos de Volterra, y entre ellos uno de los más conocidos es el GMP, *generalized memory polynomial*, el cual será usado en este trabajo en su versión bidimensional. Adicionalmente, existen muchas modificaciones de tal modelo para aumentar su rendimiento, pero el modelo GMP será suficiente para este caso. Sabiendo qué modelo se va a usar, lo siguiente es escoger qué regresores incluir. Esto dependerá del amplificador en cuestión, puesto que según el tipo cambiarán ciertas características como el orden de los regresores o la *memory depth*, incluso entre PAs del mismo tipo. Esta elección de regresores se puede realizar a priori, añadiendo la memoria y orden necesarios según anteriores casos experimentales o usar selección de regresores en ejecución para quedarse con los que más aporten, lo cual también ayudará a elegir cuántos regresores incluir, aplicando criterios como el bayesiano, que se suelen usar en conjunto con las técnicas de selección, para obtener el mejor número de regresores que añadir a la matriz de regresión.

## 3.4 Posibles problemas del DPD

En esta última sección del capítulo sobre sistemas DPD se verán algunas de las posibles y más comunes incidencias que le pueden suceder al algoritmo, las cuales acaban en resultados negativos como divergencia

o alto NMSE.

### 3.4.1 No reduce el NMSE, ACPR o EVM

Estos son los principales atributos que se buscan minimizar en el DPD, por lo que su reducción debería estar prácticamente garantizada dando por hecho una correcta implementación, aunque puede ser que se den ciertos sucesos que causen que el sistema no reduzca lo suficiente, lo cual se puede dar por varias razones.

Una de estas es que el parámetro  $\mu$  escogido sea muy pequeño, por lo que no le da tiempo en las iteraciones dadas a alcanzar el mínimo. Esto se vería si el valor de estas figuras de mérito es muy estático y su solución sería aumentar el valor de la tasa de aprendizaje. Otra causa puede ser la incapacidad del modelo para adaptarse a los efectos del amplificador debido a que contiene insuficientes regresores para su modelo, por lo que sería necesario añadir más, y a ser posible combinándolos con técnicas de selección.

También hay veces, aunque no es nada habitual, que la característica no lineal no sea posible de modelar con las series de Volterra en ciertos casos. Un ejemplo de esto es la caracterización en doble banda de un amplificador sintético basado en una tangente hiperbólica, donde se podrá reducir el NMSE hasta cierto punto para no bajar más.

### 3.4.2 Alta PAPR

Otro caso bastante común es el hecho de que al ir avanzando de iteraciones la PAPR vaya aumentando descontroladamente. Esto ocurre ya que el sistema es configurado de manera que reduzca el error, sin tener en cuenta la PAPR, que debido a la suma de regresores similares tiende a incrementarse mucho.

Una solución muy sencilla y comúnmente aplicada es emplear *clipping* [26], técnica de CFR basada en reducir la amplitud de los picos de señal que sobrepasan cierto umbral, antes de la entrada del PA, distorsionando la señal, pero quitando la PAPR indeseada. Este procedimiento es necesario en ausencia de otros métodos de control de PAPR ya que un alto factor de cresta lleva irremediablemente a divergencia, al aumentar mucho los efectos no lineales y hacer que estos no se puedan corregir. Otras técnicas empleadas para controlar la PAPR con el paso de las iteraciones es añadir en el cálculo de la actualización de los pesos un término que penalice las PAPRs altas, de esta manera el algoritmo de reducción de error también lo sería de PAPR y evitaría este suceso aunque perdiendo algo de capacidad de llegar al mínimo [27, 28].

### 3.4.3 Divergencia

Este es el problema principal que combatir, relacionado íntimamente con todos los demás efectos y parámetros, siendo causa de ellos o su origen y una mala elección de casi cualquier parámetro puede causar esta situación.

En vez de nombrar qué parámetros pueden dar lugar a este problema, mejor hablar de en qué casos la provocan, al darse en casi todos. Un valor de  $\mu$  muy alto provocará divergencia muy fácilmente, un gran número de iteraciones si no se tienen muy bien controlados los demás parámetros puede causarlo también, un número alto de regresores también, al aumentar el número de condición de la pseudoinversa. Debido a que los amplificadores y la situación a tratar en este Trabajo Fin de Grado es algo complicada, sumado a los extensos modelos bidimensionales, el alto número de regresores está garantizado, por lo que el número de iteraciones y parámetro de aprendizaje deberán ser cuidados para no causar problemas en la ejecución.

Aparte, en el momento de diverger el algoritmo, todos los parámetros que se vigilan para comprobar el correcto funcionamiento de este son seriamente perjudicados, aumentando mucho sus valores y haciéndolo cada vez más si continúan las iteraciones. Una vez el sistema empieza a divergir, es muy complicado que recupere el estado original de búsqueda de convergencia [29].

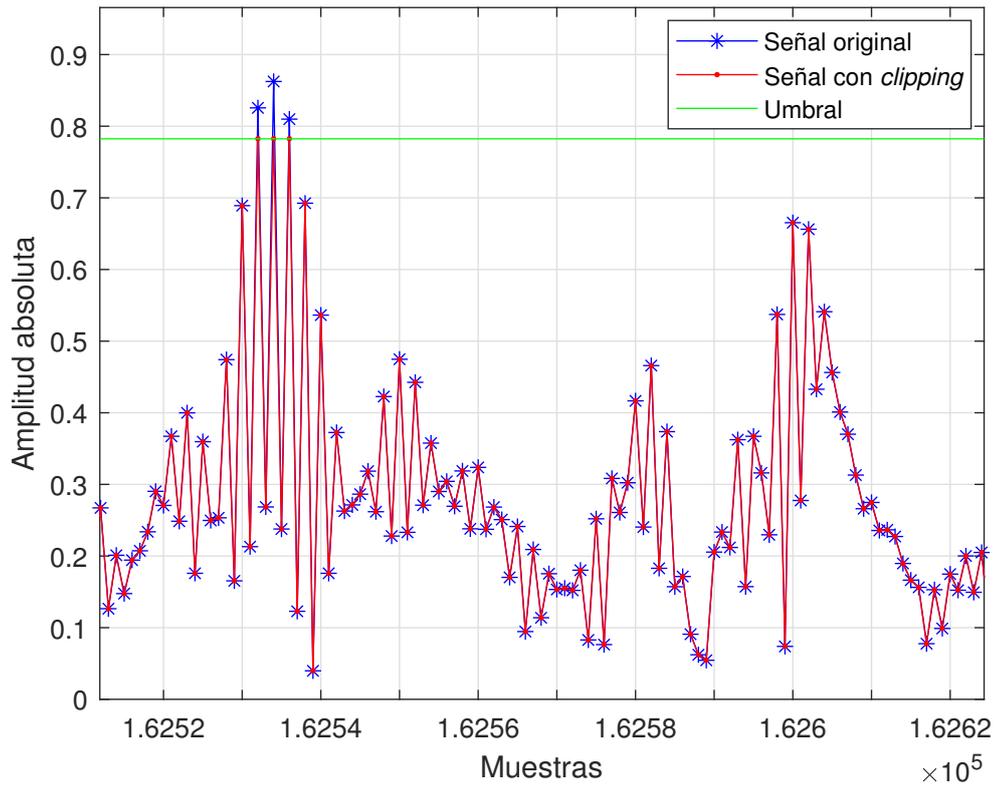


Figura 3.4 Uso de *clipping* sobre señal de alta PAPR.

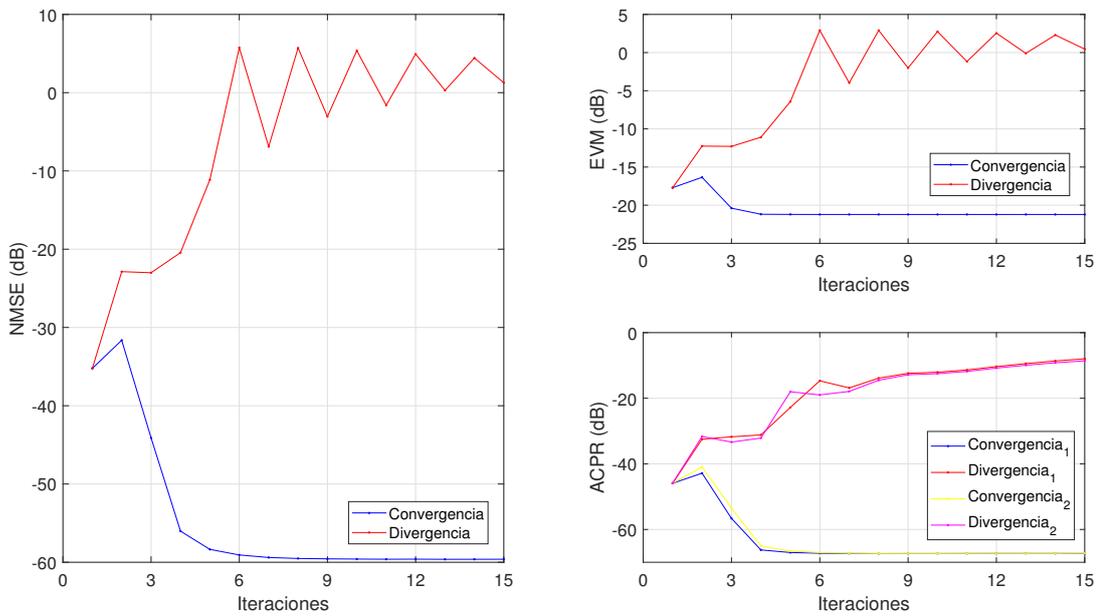


Figura 3.5 Divergencia de los parámetros principales NMSE, EVM y ACPR.

## 4 DPD de doble banda

---

*No se puede enseñar nada a un hombre, sólo se le puede ayudar a descubrirse a sí mismo.*

GALILEO GALILEI

En este capítulo se revisarán los detalles de la implementación del esquema de predistorsión elegido, el directo, en el caso de la doble banda concurrente (2D-DPD). Este añadido producirá cambios radicales en todo el proceso de linealización, y es ahí donde se centrará el capítulo. Primero se expondrá el porqué es importante el tener un esquema dual para las comunicaciones actuales, aun con las desventajas que eso acarrea, las cuales se contarán en la siguiente sección. También, debido a su importancia, se hablará del modelo usado en este proyecto para el 2D-DPD el cual es bastante especial al incluir nuevos términos para adaptarse a la existencia de otra banda pasando por el mismo amplificador. Para acabar se hará un resumen de todo lo expuesto hasta ahora en el trabajo, mostrando el esquema final, comentando algunos detalles y dando pie al siguiente capítulo, donde se verán técnicas avanzadas de mejora del sistema.

### 4.1 DPD en banda dual, qué es y qué aporta

El concepto de 2D-DPD nace de la necesidad de aplicar los conceptos de predistorsión digital a los sistemas de comunicaciones más actuales, los cuales requieren capacidades aún más grandes de manejo del espectro. En ese caso, esta gestión espectral se realiza enviando dos señales OFDM en diferentes bandas, pudiendo separar la información en dos generadores de señal. Por supuesto, este es un gran reto técnico y también complica el diseño del DPD, sobre todo en materia de modelado, necesitando mas coeficientes y modelos más extensos con propiedades especiales para manejar sus particularidades.

#### 4.1.1 ¿Por qué separar en dos bandas?

Como se acaba de comentar, el separar las bandas aportará, entre otras cosas, el poder manejar el espectro de manera más eficiente a la vez que poder incrementar las tasas máximas. Esta tecnología usada en la cuarta y quinta generación, incluso en la tercera con HSPA+, es llamada *carrier aggregation* [30] y basa su funcionamiento en el uso de varias bandas de datos a la vez. Las ventajas de usar esta técnica son obvias, puesto que se obtendrá la capacidad de usar más recursos en el espectro, lo que conlleva mayores tasas, o separar la transmisión entre ellas para ajustarse a los huecos en este. Pero al aumentar estos parámetros también lo hace la complejidad, al tener que gestionar las dos bandas a la vez.

Y este problema se extiende al DPD, el hecho de tener un esquema dual no solo causa una duplicación en todos los elementos del sistema si no que la presencia de una señal extra al crear productos de intermodulación afectará a la otra, los cuales no pueden ser modelados por los regresores habituales. Este hecho dará lugar a un conjunto de modelos teniendo este hecho en cuenta y que se verán en una sección posterior.

Otro interesante beneficio que otorga el trabajar con dos bandas independientes es que se pueden utilizar diferentes elementos para su generación. Esto es especialmente útil en los generadores de señal, que tienen un límite en la frecuencia de muestreo, y al crear una onda que ya contenga las dos señales concurrentes causa que el factor de sobremuestreo no sea tan grande como se quisiese debido a esta limitación en la frecuencia de muestreo. De esta manera, se puede generar una señal individual por dos salidas diferentes,

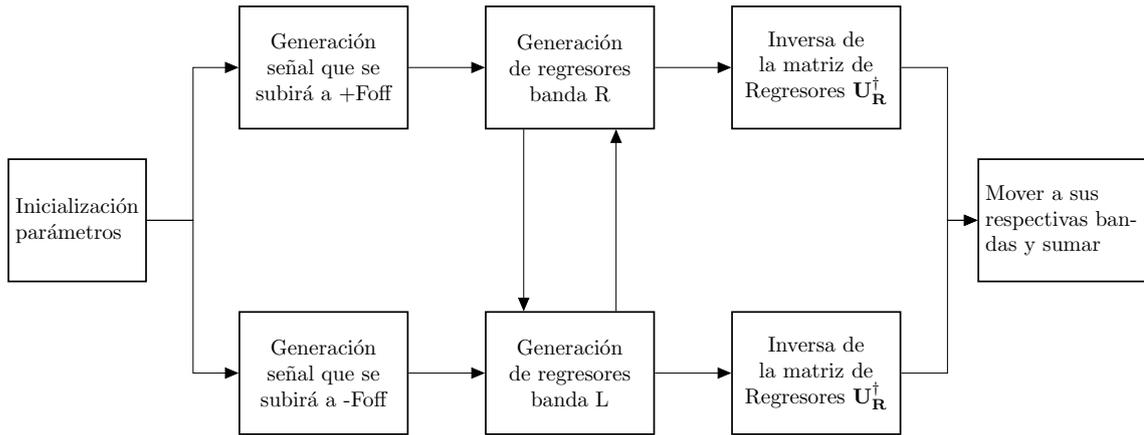


Figura 4.1 Esquema de la inicialización de los elementos del DPD en doble banda.

aprovechando el máximo factor de sobremuestreo disponible y gracias a esto tener más precisión en las medidas. Adicionalmente, el hecho de separar las bandas otorga la capacidad de tener las bandas separadas bastante entre ellas, lo cual es un caso común en la asignación del espectro.

Pero como siempre, se consigue un beneficio, a un coste. Esta implementación creará varios retos extras, los cuales se ven a continuación.

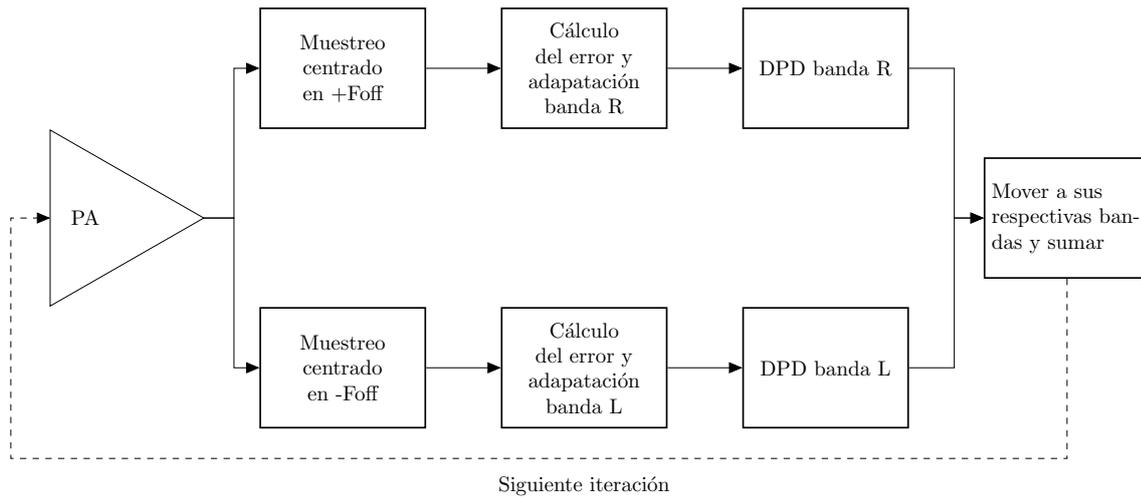
En la Figura 4.2 se observan los pasos previos al bucle en el esquema directo dual, donde partiendo de unos parámetros iniciales como son el ancho de banda de las señales, el número de iteraciones o la frecuencia de *offset*, denotada por *Foff*, que es donde se moverán las señales OFDM una vez realizados los procesos iniciales, y posteriormente generarlas por separado para cada banda. Es de notar que las señales se generan centradas, al estar trabajando con la envolvente compleja, y será en la etapa final donde la frecuencia central será modificada para que cada una acabe en su posición correcta. También es importante comentar la notación usada, en general se empleará el subíndice L para nombrar la banda de la izquierda, de *left*, al igual que R para *right*, siendo la banda de la izquierda la que se centrará en  $-Foff$  y la de la derecha en  $+Foff$ . Los caminos de ambas bandas se tocarán en el momento de la generación del modelo, este hecho se comentará en profundidad en la sección 4.3. Para acabar, y justo antes de combinar las señales de cada banda, se generará la matriz pseudoinversa para el descenso del gradiente.

También merece la pena comentar cómo cambia el proceso dentro del bucle principal, al ser importante también en el caso donde no se usan puertos distintos. La salida del amplificador devolverá en todo caso una señal en la cual estarán multiplexadas en frecuencia ambas bandas. Las envolventes complejas serán muestreadas con las frecuencias centrales de *offset* definidas previamente habiéndolas filtrado paso de banda antes, recuperando la salida en el mismo formato que se generó la entrada, centrada en frecuencia y con la misma tasa de muestreo. De esta manera se calculará el error de forma normal, se realizará el paso de descenso de gradiente y predistorionará la señal para la siguiente iteración. De forma similar a la vista en la Figura 4.2, al tener las señales listas se moverán en frecuencia y se sumarán para dar la señal de doble banda concurrente.

## 4.2 Dificultades añadidas del modelo dual

Los problemas causados por la doble dimensionalidad del DPD son varios, pero casi todos son provocados por la duplicación de los elementos que lo constituyen. En general, y ya que la predistorción es una técnica demandante, más aún si sobre esta se aplican técnicas de selección o regularización, disparan el tiempo de cómputo. Una opción podría ser no usarlas, pero estas técnicas mejoran los resultados en conjunto y previenen de problemas tan negativos como la divergencia, de forma que es difícil obviarlas. Además, el objetivo es obtener la mejor linealización posible, lo que se consigue con los métodos más exhaustivos. Pero también existen ciertas ocasiones en las que se podrán emplear versiones de complejidad reducida de ciertos algoritmos con mínimas diferencias en rendimiento.

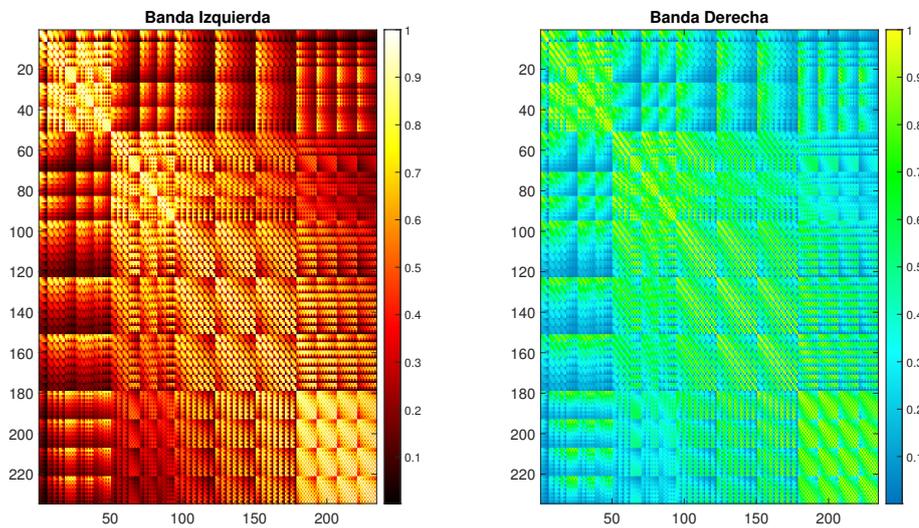
Por esto, la duplicación de elementos supondrá un aumento en las necesidades de cómputo y memoria del sistema, el cual es habitualmente una FPGA (matriz de puertas lógicas programable en campo, *field-programmable gate array*), de manera que a mayor requerimiento de potencia más coste de la placa [31]. Por ejemplo, el factor de sobremuestreo es necesario para ofrecer más resolución a la vez que poder apreciar el



**Figura 4.2** Esquema de los elementos del bucle del DPD en doble banda.

efecto de las señales en los canales adyacentes (es decir, determinar el ACPR). Este sobremuestreo causará el aumento de la memoria necesaria para manejar los vectores en tantas veces como valor tenga, pero es indispensable.

Este problema se extiende a todos los elementos, causando la ocupación de más memoria (regresores y señales) y tardando más en el cálculo (calcular la derivada del error, todos los regresores y la unión/separación de las bandas). Pero de todo, el mayor problema es el que aparece al manejar matrices con el gran número de parámetros que hay, causados por la necesidad de tener modelos más grandes que se adapten mejor a las intermodulaciones. Este fenómeno lleva a la llamada “*curse of dimensionality*” [32], la cual ocurre al tener demasiados regresores (en este caso serían las dimensiones) y que causarán problemas numéricos debido a la alta correlación entre muchos de ellos al ser generados a partir de la misma señal.



**Figura 4.3** Matrices de correlación de los regresores de las bandas derecha e izquierda. Un valor más alto indica mayor correlación, alcanzando el máximo en 1.

En la Figura 4.3 se muestran las matrices de correlación de los regresores de ambas bandas. Estas matrices de correlación se generan mediante las ecuaciones siguientes:

$$C_L = U_L^T U_L, \tag{4.1}$$

$$\mathbf{C}_R = \mathbf{U}_R^T \mathbf{U}_R, \quad (4.2)$$

donde  $\mathbf{C}_L$  y  $\mathbf{C}_R$  representan las matrices de correlación, mientras que  $U_L$  y  $U_R$  son las matrices de regresores de las bandas izquierda y derecha. Se puede observar como la diagonal tiene valor máximo e igual a uno siempre, puesto que son las autocorrelaciones de los regresores. También se puede observar el problema de la alta correlación entre las componentes, siendo verdad que no todos los valores están cercanos al máximo pero en general son muy altos y causarán problemas numéricos, hecho que se puede ver reflejado también en el número de condición de la matriz, que es del orden de  $10^{17}$  para ambas bandas [33]. Para acabar con el análisis de la figura, se puede ver cierta distribución en cuadrados, la cual es un efecto debido a la generación de diferentes tipos de regresores y que se verá en la siguiente Sección.

El problema de la masiva cantidad de elementos y los problemas numéricos sugieren la necesidad de aplicar técnicas complementarias al sistema, las cuales se tratarán en el siguiente Capítulo, debido a su gran importancia. Tanta es, que en determinadas situaciones el uso de la predistorsión en el esquema dual es imposible si no se hace uso de al menos una.

### 4.3 Modelos de doble banda

Sabiendo el problema que va a causar el hecho de tener dos bandas activas a la vez pasando por el PA por culpa de la intermodulación, es necesario buscar una solución, la cual se encontrará dentro de la configuración del modelo elegido.

#### 4.3.1 Memory Polynomial (MP)

Los modelos de una sola banda, basados en series de Volterra, se limitan a usar la señal de entrada, ya que es única, para generar el conjunto de regresores, que en ese caso es suficiente. El modelo clásico dual va a un paso más allá, realizando combinaciones de la señal de la banda junto con versiones en valor absoluto y elevadas a cierto valor de la señal concurrente, con el objetivo de eliminar esas intermodulaciones. Este modelo se puede considerar un 2D-MP, ya que amplía el concepto de *memory polynomial* [33] (MP, ecuación (4.3)) al espacio dual.

$$y_{MP}[n] = \sum_{k=1}^K \sum_{q=0}^Q a_{kq} x[n-q] |x[n-q]|^{k-1}, \quad (4.3)$$

en esta ecuación se muestra el *memory polynomial* en su forma unidimensional, pudiéndose observar que su objetivo es el combinar versiones retrasadas de la señal original con ellas mismas pero elevadas a cierto valor, de manera que puedan adaptarse a las no linealidades.

#### 4.3.2 Generalized Memory Polynomial (GMP)

Habiendo visto el modelo MP, el siguiente paso es su evolución, el modelo llamado *generalized memory polynomial* [34], GMP. Este modelo, que se muestra en la ecuación (4.4), es esencialmente el MP añadiéndole dos términos más, de manera que haya combinaciones de la señal con cierto retraso con las envolventes pero con diferentes adelantos y retrasos con respecto a ella. Uno de los detalles más relevantes es lo mucho que aumenta el tamaño del sistema pasando de un MP a GMP, hecho importante, entre otras cosas por que en este trabajo el modelo usado será un GMP bidimensional, 2D-GMP, lo que causará los problemas ya comentados debidos a modelos grandes.

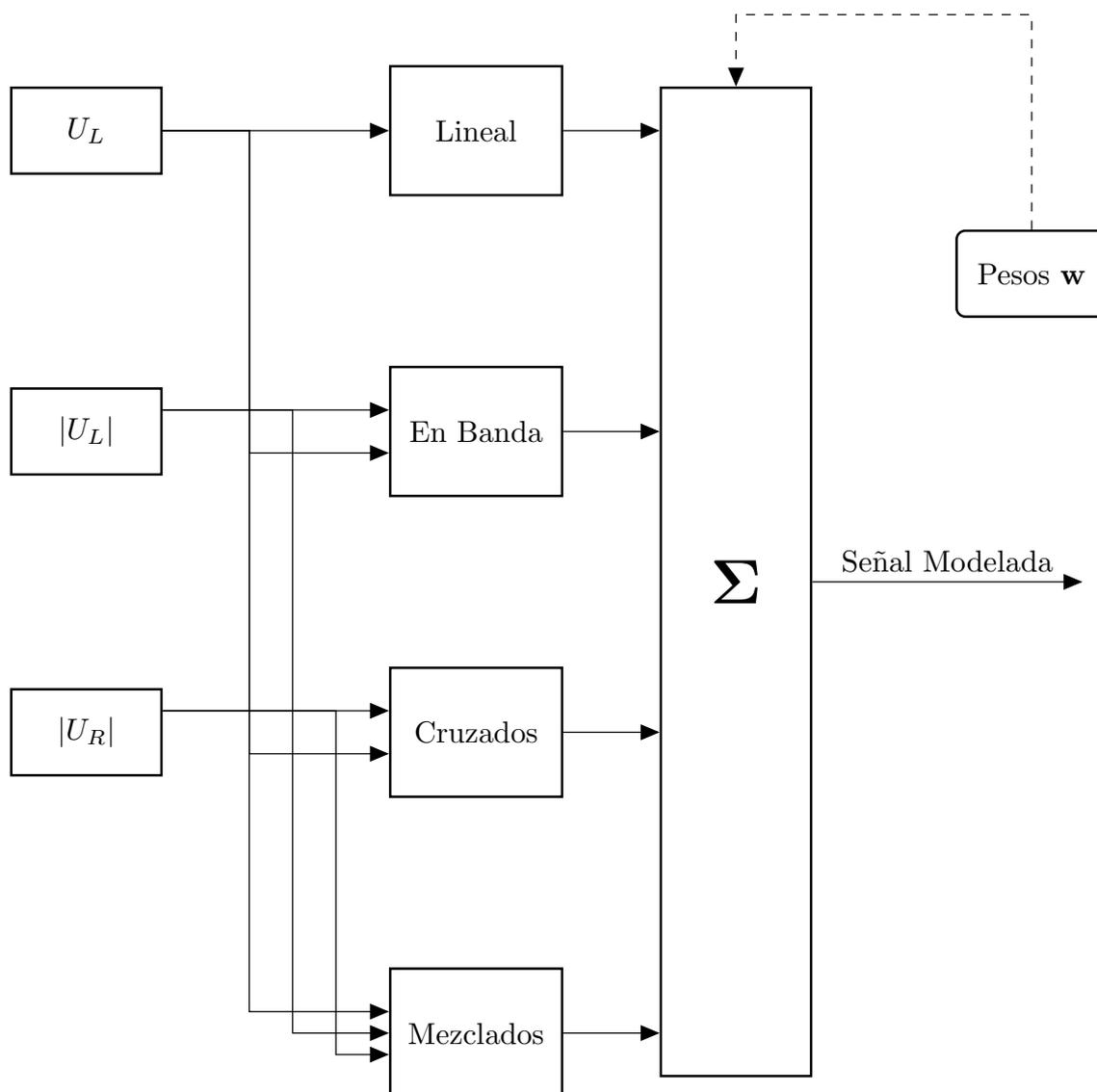
$$\begin{aligned} y_{GMP}[n] = & \sum_{k \in K_a} \sum_{l \in L_a} a_{kl} x[n-l] |x[n-l]|^k \\ & + \sum_{k \in K_b} \sum_{l \in L_b} \sum_{m \in M_b} b_{klm} x[n-l] |x[n-l-m]|^k \\ & + \sum_{k \in K_c} \sum_{l \in L_c} \sum_{m \in M_c} c_{klm} x[n-l] |x[n-l+m]|^k. \end{aligned} \quad (4.4)$$

Así pues, el modelo, presentado por primera vez por Enzinger en la *student competition* de 2018 [23], es la evolución del GMP al territorio dual. Desde luego, el uso de este modelo es un éxito: consigue adaptarse a las intermodulaciones y modelar con un nivel de error que no es, de lejos, alcanzable con los otros modelos de una banda. A cambio aparecen los problemas ya nombrados, se generan muchos regresores (también porque

el problema es más exigente) de alta correlación, lo cual causa errores numéricos que a su vez provocan la divergencia del algoritmo, problema que se debe evitar a toda costa puesto que al llegar a este estado la potencia de entrada y PAPR puede aumentar demasiado, incluso llegando a causar daños en el PA.

#### 4.3.3 2-Dimensional Generalized Memory Polynomial (2D-GMP)

Hablando más en profundidad del modelo de Erzinger, en su estructura son incluidos varios tipos de regresores para darle ese potencial contra la intermodulación. Existen dos tipos de regresores clásicos usados en el problema de una dimensión, uno de ellos el lineal, el cual solo contiene la señal original con retrasos para los efectos de memoria. El otro regresor es llamado “en banda” y se genera a partir de la señal con cierto retraso y su versión en valor absoluto. Los dos regresores especiales que añade Erzinger (siguiendo el modelo 2D-MP [35] propuesto por Bassam) se llaman “cruzados”, donde intervienen la señal original y la concurrente en valor absoluto, y “mezclado” donde toman parte la señal original tanto normal retrasada como en valor absoluto y la concurrente solo en valor absoluto. Un esquema de la generación de los diferentes tipos de regresores se muestra en la Figura 4.4, de manera que sea más simple ver las implicaciones de cada señal en los tipos.



**Figura 4.4** Tipos de regresores del modelo 2D-GMP de Erzinger, aplicados a la banda izquierda.

A continuación se mostrarán las expresiones de los regresores del 2D-GMP [23], empezando por los lineales. Todas las expresiones manejadas están en tiempo discreto, lo cual justifica el uso de corchetes en

estas en vez de paréntesis. El regresor más básico, el lineal, tendrá la forma:

$$x_{L,Lineal}[n] = \sum_{m=0}^M w_m u_L[n-m], \quad (4.5)$$

En cambio, los regresores en banda son algo más complejos, como se muestra en (4.6). En este caso sí que hay algunas diferencias con el modelo normal del GMP, entre ellos la exclusión de la parte lineal ya que está incluida en el bloque anterior, la reformulación del tercer bloque para excluir términos no causales y el añadido de índices en los límites del sumatorio para no introducir regresores propios de otros bloques. Además, esta formulación incluye también a los regresores mezclados, si la señal en valor absoluto cambia de ser  $u_L$  a  $u_R$ .

$$\begin{aligned} x_{L,EnBanda}[n] = & \sum_{m=0}^{M_1} \sum_{p=1}^{P_1[m]} w_{mp}^{(1)} u_L[n-m] |u_L[n-m]|^p \\ & + \sum_{k=1}^{K_2} \sum_{m=0}^{M_2[k]} \sum_{p=1}^{P_2[k,m]} w_{kmp}^{(2)} u_L[n-m] |u_L[n-m-k]|^p \\ & + \sum_{k=1}^{K_3} \sum_{m=0}^{M_3[k]} \sum_{p=1}^{P_3[k,m]} w_{kmp}^{(3)} u_L[n-m-k] |u_L[n-m]|^p. \end{aligned} \quad (4.6)$$

Para acabar, los regresores cruzados se forman como expone la ecuación (4.7):

$$x_{L,Mezclado}[n] = \sum_{m=0}^{M_1} \sum_{p=1}^{P_1[m]} \sum_{p_1=1}^{P_2[m,p_1]} w_{mp_1 p_2} u_L[n-m] |u_L[n-m]|^{p_1} |u_R[n-m]|^{p_2}. \quad (4.7)$$

Después de ver todas las ecuaciones que generan los regresores del modelo usado puede entenderse de verdad tanto su tamaño como el parecido que tendrán las señales entre ellas, al ser muy parecidos todos los métodos. Con esta sección queda cerrado el asunto de los regresores, por lo que a continuación se mostrará el esquema final empleado.

## 4.4 Esquema usado

Para cerrar el Capítulo, se resumirán las generalidades del sistema final, que se muestra en la Figura 4.5.

Para empezar, el bloque del DPD se aplicará tal cual es expuesto, en paralelo se calcularán las señales predistorsionadas  $x_L$  y  $x_R$  para la etapa de síntesis. En este bloque se moverán a las frecuencias de *offset*, como ya ha sido comentado, y a su salida se obtendrá la suma de ambas, para después convertirlas de su versión de envolvente compleja a señal de radio frecuencia para ser amplificada. Se supone que la potencia de las señales será, en general, igual en ambas bandas por lo que la salida tendrá una potencia 3 dB superior a sus componentes. A la salida del amplificador se normalizará la ganancia del amplificador en la señal para obtener una versión de misma potencia que la entrada  $\mathbf{u}$  para poder compararlas. En el bloque de análisis se separarán las señales y puesto que las dos tienen la misma potencia se devolverán ambas con 3 dB menos, la mitad de potencia a cada una, y estarán centradas. En este bloque también se convertirán las señales de radio frecuencia a envolventes complejas, al contrario que en la síntesis. Posteriormente se realiza el cálculo del error y se calcula el paso del DPD para actualizar los pesos, mediante la inversa de Moore-Penrose.

Aquí finaliza el capítulo dedicado a hablar del modelo dual y su implementación, en el siguiente capítulo se tratarán diferentes técnicas de mejora y optimización del algoritmo, las cuales son muy importantes para este tema y en el posterior a este los resultados experimentales obtenidos por el algoritmo.

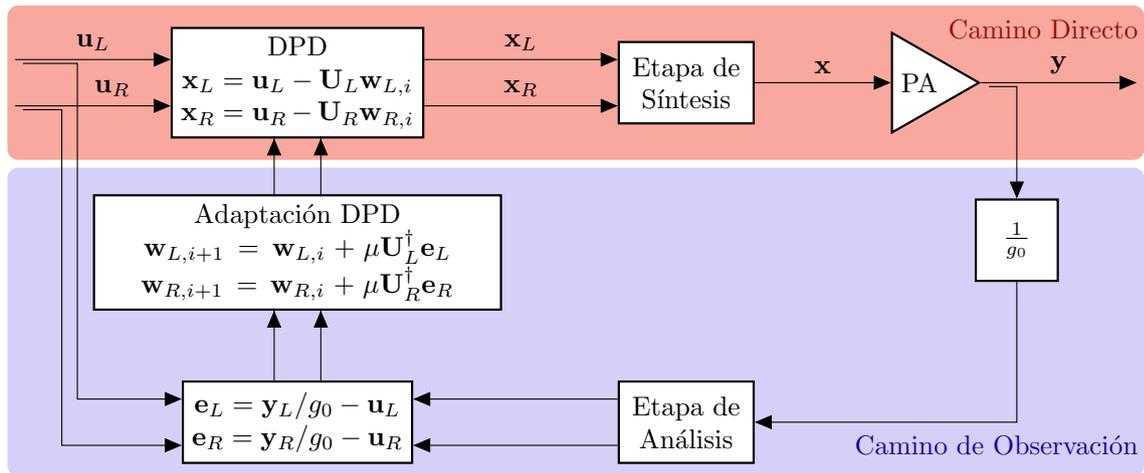


Figura 4.5 Esquema final usado, modelo 2D-GMP.



## 5 Técnicas complementarias del DPD

---

*Nosotros escogemos a quién dejamos entrar en nuestro mundo.*

SEAN MGUIRE EN EL INDOMABLE WILL HUNTING

Habiendo visto ya las dificultades del esquema dual para converger, en ciertas situaciones, debido a los problemas numéricos que ocurren, resulta necesario añadir ciertos complementos al algoritmo para mejorarlo. En general, estas técnicas no cambiarán las bases de cómo se compone el sistema, solo modificarán sus parámetros o añadirán ciertos valores para conseguir mejoras concretas. Estos métodos se explicarán para el caso dual, pero por supuesto pueden ser usados en el caso de una sola banda, donde la divergencia no es un problema tan recurrente, pero sí que conseguirán una gran mejora en rendimiento.

En este Trabajo Fin de Grado se hablará principalmente de tres grupos de técnicas, las que controlan el valor del coeficiente de aprendizaje, las técnicas de selección de coeficientes y las técnicas de regularización, con ejemplos concretos.

### 5.1 Técnicas de coeficiente de aprendizaje dinámico

Este grupo de procedimientos tiene como objetivo modificar el parámetro de aprendizaje en cada iteración del algoritmo, de manera que, a ser posible, este se adapte a las necesidades de cada paso.

Como ya ha sido comentado, es habitual que el sistema DPD funcione mejor con unos valores de  $\mu$  altos en las primeras iteraciones para ir decayendo rápidamente, ya que las primeras iteraciones tienen la capacidad de reducir mucho el error, pero se va reduciendo, causando que al llegar a cierto punto es necesario reducir el tamaño del paso o la divergencia es muy probable (ver la sección (3.3.2)). Este concepto es denominado normalmente *cooling* o enfriamiento.

Lo idóneo sería tener un algoritmo que, en base a ciertas circunstancias de la ejecución, como puede ser la pendiente del descenso del gradiente, ajustara el valor para decidir en el caso de que la caída sea muy pronunciada avanzar menos y viceversa, lo cual se conoce como tasa de aprendizaje adaptativa [36].

En este proyecto, debido al conocido comportamiento, la implementación de este tipo de métodos ha sido realizada escogiendo previo a ejecución un vector con los valores de tal parámetro, empezando con el valor alto en la primera iteración y reduciéndolo poco a poco. Esta es una técnica muy simple pero tremendamente efectiva, y permite que, en situaciones de clara divergencia, impídala, y consiguiendo muy buenos resultados en cuanto NMSE y otras figuras de mérito. Una técnica sencilla que se puede implementar aprovechando esto es generar varias funciones con longitud igual a las iteraciones, para así ir probando cual es la más adecuada. Varias opciones se muestran en la Figura 5.1. En la figura se puede comparar el rendimiento de estas funciones, siendo especialmente eficaz las basadas en funciones exponencialmente decrecientes. También merece una mención especial la llamada personalizada, que en este caso concreto no parece ser muy útil pero ha arrojado muy buenos resultados experimentales.

Sobre todo, es reseñable que una técnica cuya complejidad es tan reducida entre técnicas tan costosas como suelen ser, como por ejemplo las siguientes, logre tan buenos resultados además de poder ser mejorable de muchas formas, y también fue importante para la resolución de este trabajo debido a los problemas iniciales de divergencia que ocurrían tomando medidas en el laboratorio de la Universidad de Chalmers.

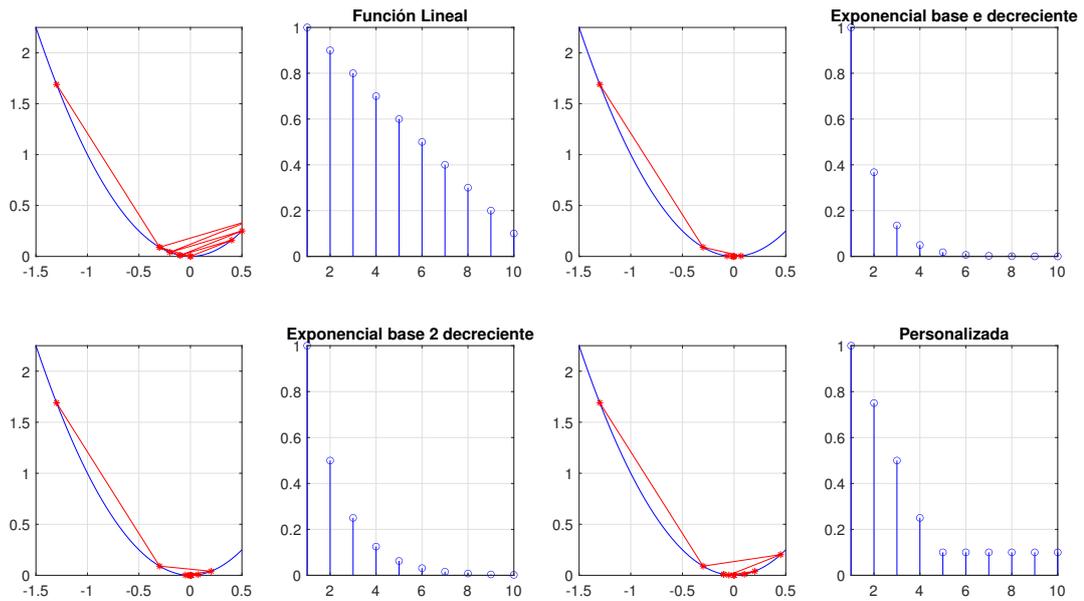


Figura 5.1 Diferentes tipos de funciones para el enfriamiento.

## 5.2 Técnicas de selección de coeficientes

Otro método que se puede aplicar a los sistemas DPD es buscar los regresores que tengan el mayor impacto en la señal para quedarse con estos y descartar los de menor peso. Esta evaluación se realizaría en general mediante correlaciones, que es el caso que se va a explicar. Las técnicas más típicas de selección de regresores son los algoritmos *orthogonal matching pursuit* (OMP) y *doubly OMP* (DOMP), mientras que el criterio que se suele usar para seleccionar el número óptimo de coeficientes, ya que estas técnicas lo que hacen es ordenar los regresores, es el *bayesian information criterion* (BIC).

### 5.2.1 Orthogonal Matching Pursuit

Esta técnica es conceptualmente muy simple pero bastante costosa en cuanto a cómputo. Lo que pretende este procedimiento es realizar comparaciones entre los regresores mediante correlaciones con el vector de salida, el cual contiene las no linealidades, para elegir el que tenga la mayor correlación de todas. Posteriormente se le restará a la salida el modelado de ese regresor de la señal para seguir analizando los demás regresores [37].

En la Figura 5.2 se muestra el diagrama del algoritmo. Se puede ver que las entradas que recibe son la matriz de regresores y la salida asociada a esa señal del amplificador. Lo primero será normalizar las columnas de  $U$ , para que todos los regresores tengan el mismo peso, y justo después se entra en el bucle con tantas iteraciones como regresores se quieran ordenar. Mediante la correlación de estos con la salida  $\mathbf{y}_{\text{orth}}$ , que es la salida original a la cual se le han sustraído la proyección de los regresores ya elegidos, se reordenarán y se añadirá el mejor a una matriz auxiliar donde aparecerán ordenados, para con esa misma modelar la salida y restársela a la salida original. Este algoritmo busca ir seleccionando regresores y sustraer su influencia de la salida, de esta manera evitando situaciones en las que haya varios regresores de alta correlación con la salida, pero alta también entre ellos, añadiendo información repetida y posibles problemas numéricos. De aquí viene el concepto de ortogonalidad del algoritmo, el hecho de que se vayan sustrayendo los regresores ya seleccionados de la salida para realizar las correlaciones crea cierta ortogonalidad en la selección de regresores. Para acabar la iteración, se calculará el valor del criterio seleccionado, en este caso BIC, y se guardará para al finalizar el bucle extraer el mínimo.

Habiendo acabado el algoritmo se usarán los valores del mínimo del BIC junto con el orden de los regresores, de esta manera se emplearán no solo menos regresores para el modelado si no que serán mejores que elegidos arbitrariamente.

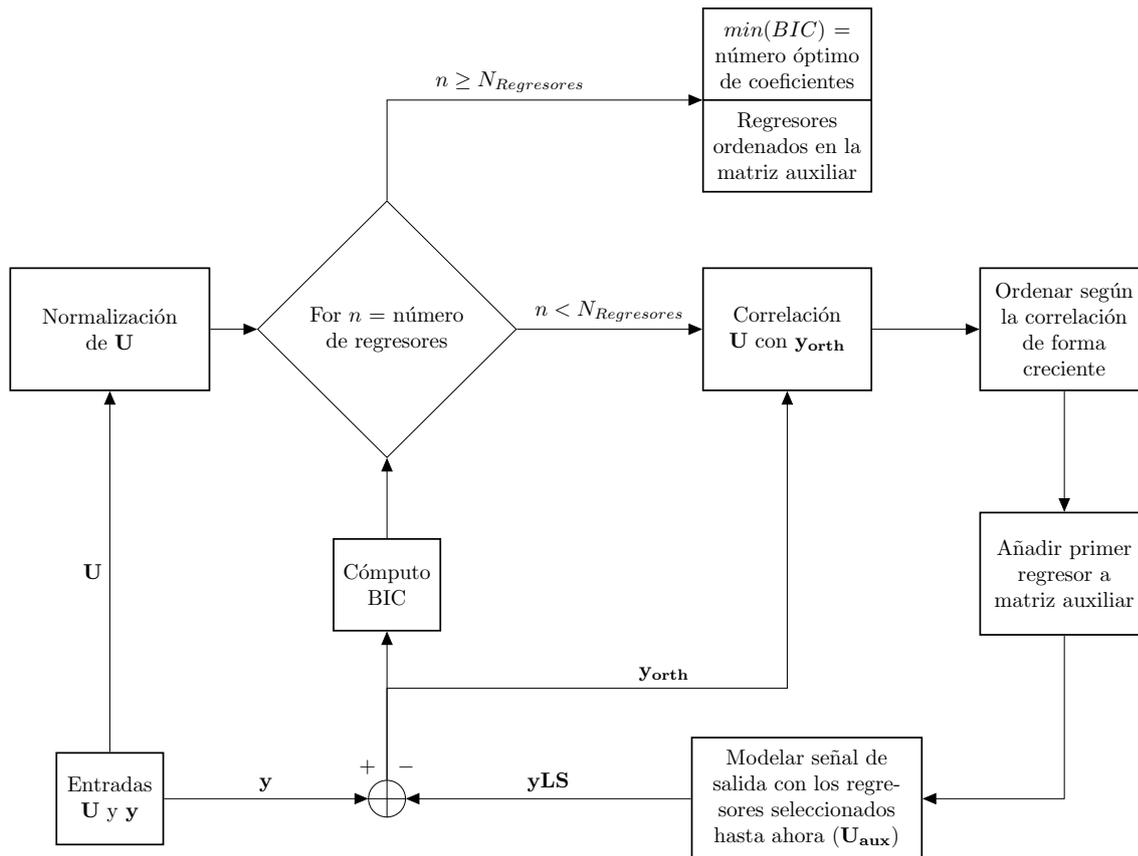


Figura 5.2 Diagrama de flujo del algoritmo *orthogonal matching pursuit*.

### 5.2.2 Doubly OMP

Pero hay un inconveniente con el algoritmo OMP, la alta correlación entre los regresores causará que las elecciones no sean tan buenas como deberían. Al restar un regresor a la salida es posible perder la información que iba a seleccionar a otros con información importante pero con poco peso dentro de un vector.

Este último caso es fácil de ver, si acaba de ser seleccionado el regresor  $\mathbf{x}_1$  existiendo el vector  $\mathbf{x}_2 = 0.95\mathbf{x}_1 + 0.05\mathbf{x}_3$ , donde  $\mathbf{x}_3$  es información nueva e importante para la regresión con el OMP es posible que se pierda. Para solucionar los problemas del anterior algoritmo y de esta forma mejorarlo está el DOMP [38], el cual es "doblemente ortogonal", al no solo restar los regresores al vector salida sino que ortogonaliza la propia matriz de regresores mediante el proceso de Gram-Schmidt. De esta manera la información única individual de cada regresor se mantendrá. Continuando con el ejemplo anterior, al ser seleccionado  $\mathbf{x}_1$  y pasar el proceso de Gram-Schmidt el regresor  $\mathbf{x}_2$  pasaría a ser  $\mathbf{x}_2 = \mathbf{x}_3$ , en este caso volviendo a tener posibilidades de ser seleccionado. El algoritmo se muestra en la Figura 5.3, el cual es muy similar al OMP, cambiando un poco la notación, por ejemplo la matriz de regresores pasa a llamarse  $Z$ , al ser ortonormal, y se añade el proceso de Gram-Schmidt para la doble ortogonalidad.

Y también, de la misma manera que en el algoritmo OMP, se le aplicara el criterio bayesiano. El algoritmo DOMP es altamente costoso, ya que el proceso de ortogonalización complica mucho los cálculos, pero los resultados son muy buenos. Para este problema, se puede emplear una variante de complejidad reducida donde se corta el proceso al llegar a cierto número de coeficiente y así no tener que ortogonalizar todo el conjunto de regresores, llamada *reduced complexity DOMP*, RC-DOMP o el SSPI-DOMP [39] donde se evita recalcular toda la pseudoinversa añadiendole columnas.

### 5.2.3 Criterio Bayesiano

El criterio bayesiano, *bayesian information criterion* [40] (BIC), es una regla basada en la función de verosimilitud con la que se calcula el número óptimo de coeficientes de un modelo.

Esta técnica tiene especial importancia en conjunto con las anteriormente mencionadas OMP y DOMP ya que, aunque ordenan los regresores, si no se limita el número de estos a un subconjunto del original se

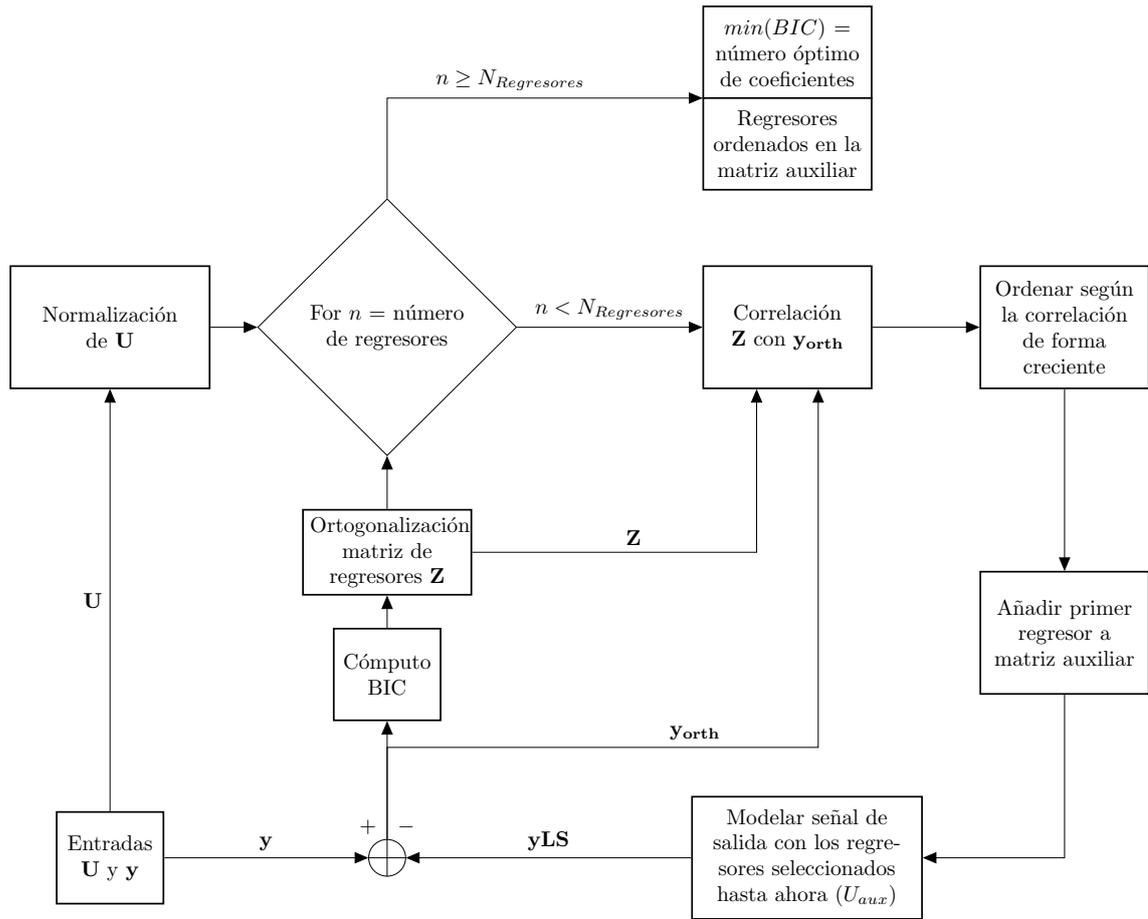


Figura 5.3 Diagrama de flujo del algoritmo *doubly orthogonal matching pursuit*.

mantendrían todos, lo cual es equivalente a no hacer nada. Una vez ordenados el número de regresores podría ser elegido de forma supervisada, pero el BIC ofrecerá el mejor valor haciendo un balance entre calidad y cantidad. La función que representa al BIC puede ser escrita como

$$\text{BIC} = k \ln(n) - 2 \ln(\hat{L}). \quad (5.1)$$

donde la parte de la derecha, denotada por  $L$ , representa la función que se busca mejorar, en este caso el NMSE y la de la izquierda es una penalización que se aplica al número de regresores,  $k$ , y su longitud,  $n$ . En cuanto a la implementación de la técnica en los métodos OMP y DOMP se realiza de la siguiente manera

$$\text{BIC} = 2M \ln \left( \frac{1}{M} \|\mathbf{r}\|_2^2 \right) + 2n \ln(2M), \quad (5.2)$$

donde  $M$  representa la longitud de los vectores y  $n$  el número de regresores. De esta manera el criterio bayesiano pasa a ser reescrito como criterio de parada para ser implementado con las técnicas mencionadas [41].

En la Figura 5.4 se muestran los NMSE para el OMP y DOMP, de donde se pueden sacar varias conclusiones. Una de ellas es lo muy superior que es el DOMP al OMP, puesto que en esta realización alcanza prácticamente el mejor orden. Ese orden óptimo se logra cuando la derivada del error es monótonamente creciente, es decir, la mejora que existe al añadir un regresor más cada vez es más pequeña. Otro dato interesante es el número óptimo de regresores elegido por el BIC en los dos casos, debido a que se selecciona con respecto a la derivada del error y en el caso del DOMP se ordenan muy bien, tiende a elegir un número muy alto, en el OMP ocurre parecido pero sin ser tan pronunciado este efecto.

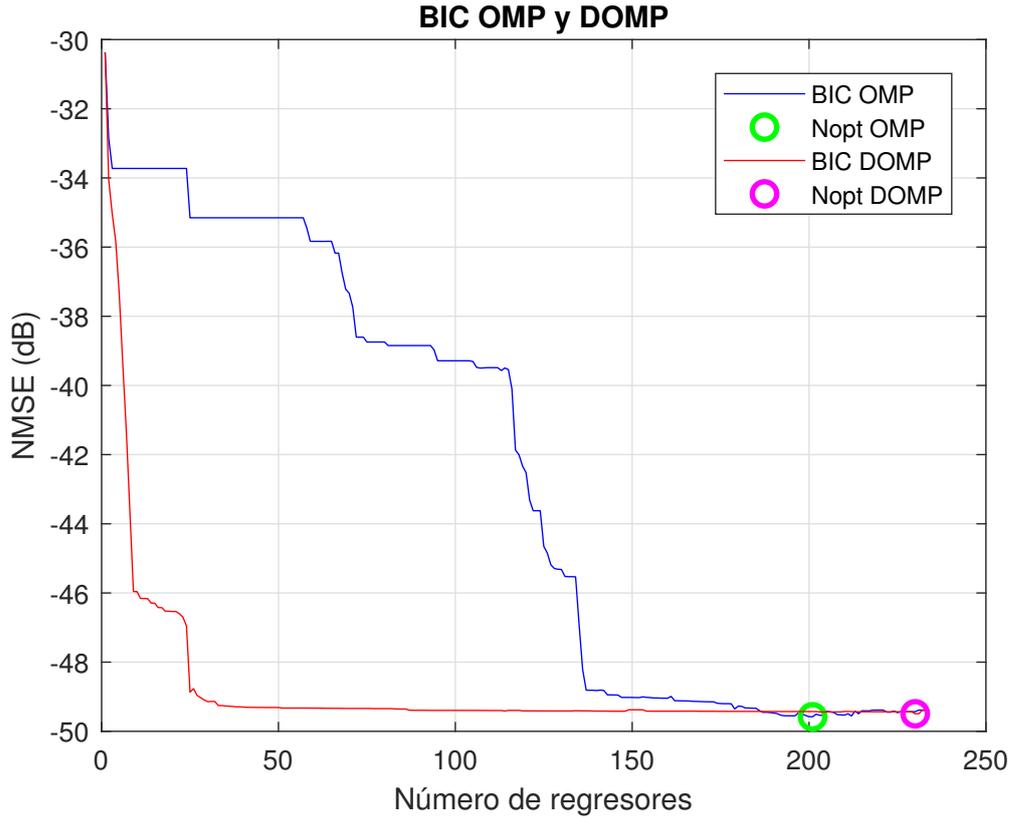


Figura 5.4 NMSE según las iteraciones para el OMP y DOMP, junto con el número óptimo de coeficientes.

### 5.3 Técnicas de regularización

Otra de las alternativas posibles para mejorar el sistema DPD es aplicar regularización. Esta técnica se basa en la introducción de ciertos valores para suavizar un efecto determinado.

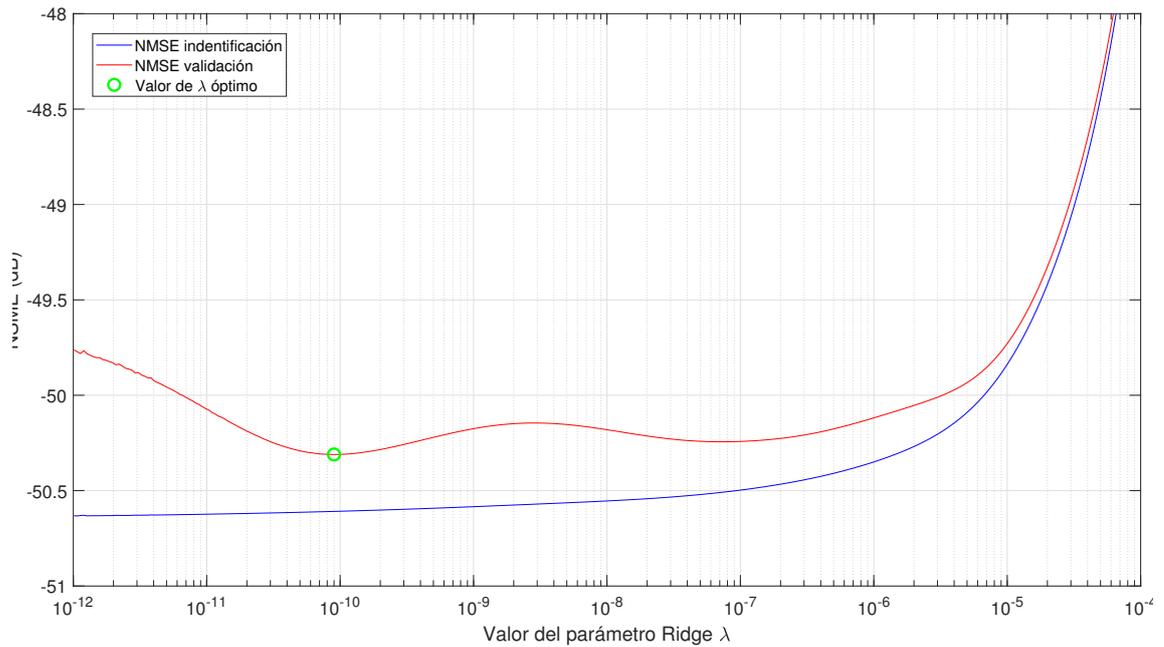
En el caso de este Trabajo Fin de Grado la técnica de regularización escogida ha sido Ridge [42, 43]. Este tipo en concreto afecta a la matriz  $\mathbf{U}^T \mathbf{U}$ , añadiéndole unos términos  $\lambda$  a la diagonal, de esta manera reduciendo la dispersión en los valores absolutos de las filas de la matriz, que son las que causan problemas numéricos al intentar invertirla. La formulación de la regularización de Ridge cambia la adaptación del DPD a

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu (\mathbf{U}^T \mathbf{U} + \lambda \mathbf{I})^{-1} \mathbf{U}^T \mathbf{e}, \quad (5.3)$$

siendo  $\lambda$  el parámetro Ridge e  $\mathbf{I}$  la matriz identidad. Para la selección del valor óptimo del coeficiente de Ridge se pueden usar varias técnicas, entre ellas un barrido o una búsqueda de mínimo local, que compararán validación con identificación, donde teóricamente se debería encontrar un mínimo.

Al añadir el término en la matriz pseudoinversa se modificarán los resultados de NMSE del modelado de la señal de salida, así pues se puede generar una gráfica como la de la Figura 5.5, donde se pueden ver la gráfica de identificación, que sería el mínimo teórico de error, y validación, donde se contrastarán los resultados y se elegirá el  $\lambda$  donde esta función sea menor. Estas curvas se generan siguiendo un esquema indirecto, es decir, modelando la señal de la manera  $\mathbf{y} = \mathbf{U}\mathbf{h}$ , haciéndose de esta manera gracias a que solo con el par matriz de regresores y salida basta para realizar la regresión, mientras que en un esquema directo se necesitaría iterar y realizar medidas con el PA.

El objetivo de la técnica de regularización de Ridge es la mejora de la precisión a costa de empeorar el sesgo, mejorando así los resultados. Además, la inclusión del parámetro en la matriz que ocasiona problemas numéricos tendrá un efecto regulador que reducirá bastante el número de condición y contrarestará muy bien la divergencia.



**Figura 5.5** Gráfica de los valores de NMSE en validación e identificación para ciertos valores de  $\lambda$ .

Esta técnica ha sido presentada en un documento para el Simposio Nacional de la Unión Científica Internacional de Radio (URSI) que tendrá lugar en septiembre de 2021 [42], en la cual se utiliza este esquema de doble banda con aprendizaje directo junto con regularización de Ridge consiguiendo muy buenos resultados. El documento se adjunta en el Apéndice D.

## 6 Pruebas experimentales

---

*Con la voluntad se puede hacer de todo. Yo soy lo que yo quiero.*

FERRUCCIO PAPINI EN LA VIDA ES BELLA

Este penúltimo capítulo se encargará de mostrar los resultados obtenidos por las configuraciones nombradas anteriormente, todas de sistemas DPD dual, que se diferenciarán en varias características, como el número de regresores usados, la técnica de mejora, como puede ser DOMP, regularización de Ridge o *Least Squares*, o el laboratorio donde se han tomado esas medidas, habiendo probado los esquemas en los laboratorios *online* de la Universidad de Chalmers y en el laboratorio de Sistemas de Radiocomunicación de la Universidad de Sevilla.

### 6.1 Banco de medida experimental

Antes de empezar con los resultados de los experimentos se hablarán un par de detalles sobre la implementación de los sistemas que aún no han sido mencionados.

Primero de todo, los modelos. Como ya se ha comentado son modelos 2D-GMP basados en [23], de los cuales se han elegido dos configuraciones de regresores para probar, una más ligera con un total de 234 regresores y otra bastante más grande, de 534, para conseguir mejores valores de NMSE a la vez que acercarse más a la divergencia para usar diferentes métodos. Se añaden ambas configuraciones de modelos en el Apéndice B.

Otro tema importante es la PAPR. Para no dañar los equipos físicos, se usa un limitador en el control remoto de las medidas de manera que el software detecta si se cumplen ciertos requisitos de PAPR relacionados también con la potencia de entrada de la señal, en el caso de no ser cumplidos se interrumpe la medida. Por esto, y puesto que los valores típicos de PAPR en estas señales es de 10 a 11 dB se ha elegido usar *clipping*, con un umbral de 12 dB para evitar problemas a la hora de tomar medidas. Aunque es verdad que el límite impuesto por el laboratorio de la Universidad de Chalmers era de entorno a 17 dB, muy lejos del que deberían tener las señales que se manejen en estos experimentos, es buena práctica limitar los valores de PAPR, entre otras cosas puesto que pueden ser causantes de problemas en el modelado.

La potencia de entrada de las señales se fijará entre -24 dBm y -22.5 dBm, intentando apurar lo máximo posible para tener la mayor eficiencia posible del amplificador de potencia. Adicionalmente, las señales se colocarán a una frecuencia de *offset* diferente dependiendo del ancho de banda de la señal, condicionado por la posibilidad que ofrezca el equipo en cuanto a *oversampling*, en el caso de Chalmers un valor de 5, lo que ofrecerá unos anchos de banda de 20 MHz, mientras que en el laboratorio de la universidad de Sevilla el sobremuestreo tomará valor 3, para 10 MHz de ancho de banda. Puede parecer que las prestaciones del laboratorio de Chalmers son mejores, pero a la hora de cálculos y resultados, estos serán mejores en el laboratorio de Sistemas de Radiocomunicación gracias a la mayor precisión de los equipos usados (y además mucho más rápidos).

Para terminar, las señales OFDM que se usarán para aplicar al amplificador tendrán 25 o 50 bloques de recursos y llevarán modulaciones 16-QAM en sus subportadoras, conforme al enlace descendente de 5G-NR.

## 6.2 Laboratorio de la Universidad de Chalmers

Dicho esto, lo siguiente es entrar a discutir los resultados experimentales obtenidos, empezando por el laboratorio de la Universidad de Chalmers, anticipando desde ya, que serán diferentes comparados con el equipo disponible en la Universidad de Sevilla.

### 6.2.1 Montaje de Chalmers

Chalmers es una universidad sueca que ofrece desde el año 2015, por la competición de estudiantes de DPD, acceso gratuito para cualquier usuario a su equipo de medidas para predistorsión digital, enviando la señal por internet vía *MatLab*, realizando la medida allí y reenviando la salida del amplificador. Esta herramienta *online* ha sido de gran ayuda en la realización de este Trabajo Fin de Grado gracias a la posibilidad de realizar cálculos sobre un dispositivo real, ya que en una primera instancia las implementaciones fueron comprobadas con un amplificador sintético simulado por software el cual busca imitar un comportamiento no lineal que no llegó a ser del todo compatible con el esquema dual.

En cuanto al montaje físico, se usa una cascada de preamplificador, amplificador y acoplador, con una carga adaptada al final. Este es el típico esquema en el que se montan este tipo de sistemas.

El laboratorio de Chalmers es accedido de forma *online* mediante su plataforma, *dpdcompetition.com*. Un esquema de su montaje se puede ver en la Figura 6.1, junto con los dispositivos físicos que intervienen.

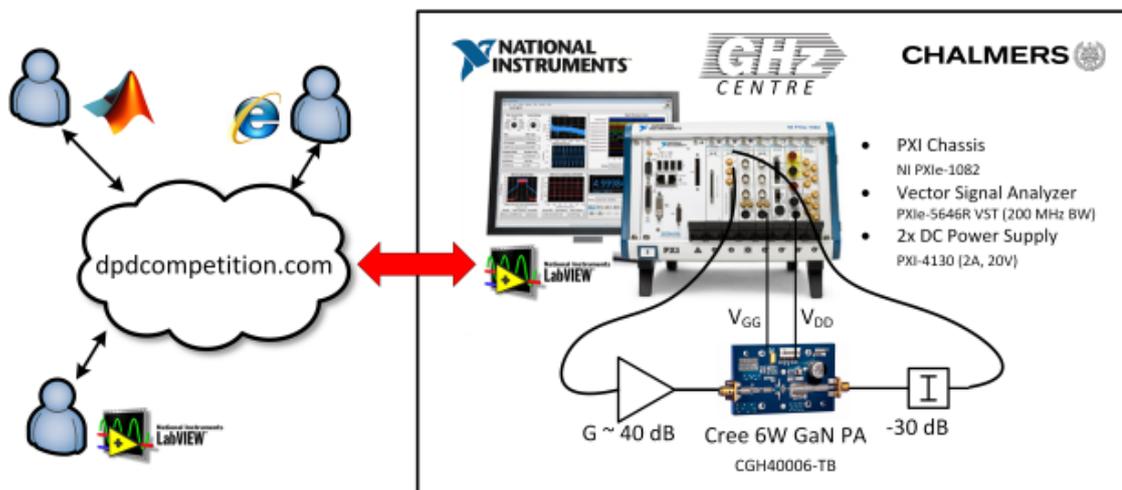


Figura 6.1 Montaje del laboratorio de la Universidad de Chalmers.

### 6.2.2 Medidas experimentales

Antes de empezar con los resultados en NMSE y demás figuras de mérito, conviene observar una de las gráficas más importantes para el seguimiento del rendimiento del DPD en ejecución, el espectro de potencia, donde se ve reflejada la reducción tanto de NMSE y ACPR con el paso de las iteraciones. Esta gráfica se encuentra en la Figura 6.2, donde se representan el paso de diez iteraciones del sistema DPD en el amplificador de Chalmers sin usar técnicas de mejora.

Otra forma de ver como mejora el error al aplicar el DPD es con la constelación, ya que sin aplicarlo se suele ver muy dispersa y según ocurren las iteraciones se van juntando los puntos dando una situación con mucha menos separación. Este efecto también se ve reflejado en el EVM que mide esencialmente esta dispersión en la constelación, y se mostrará en la tabla de resultados. La evolución de la constelación gracias al sistema DPD directo se observa en la Figura 6.3.

Para el laboratorio de Chalmers se exponen los datos con el modelo de 234 regresores puesto que el modelo de 534 es demasiado grande para la precisión que ofrecen las medidas, por lo que se reservará para su uso posterior en el laboratorio de Sevilla. La elección para el parámetro  $\mu$  será de 0.25. También se mostrará el caso de la combinación entre las técnicas de regularización de Ridge y de *cooling*, ya que, aún siendo verdad que todas estas técnicas pueden ser usadas en combinación, hay algunas que no funcionan tan bien. Esto es debido a que la regularización de Ridge hace que desaparezca prácticamente la posibilidad de divergencia, por lo que el hecho de ir modificando el parámetro  $\mu$  no es tan efectivo, de hecho es contraproducente,

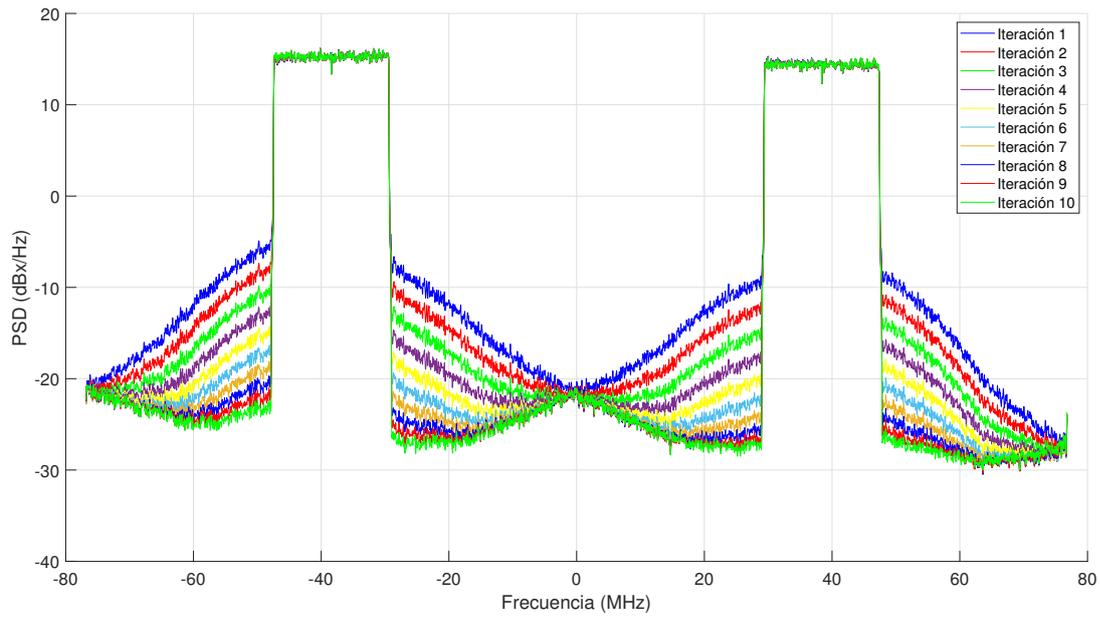


Figura 6.2 Espectro de la señal dual con las iteraciones.

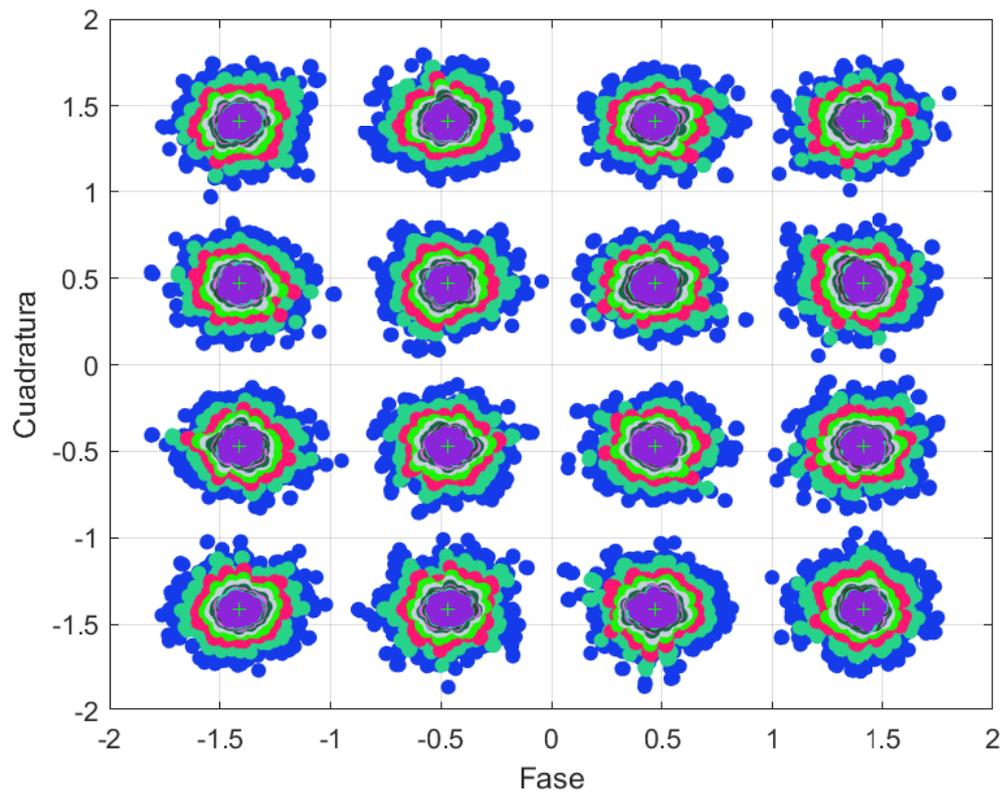


Figura 6.3 Constelación de las portadoras de la señal OFDM según pasan las iteraciones.

**Tabla 6.1** Indicadores de linealización para diferentes casos de DPD-DLA en Chalmers.

Caso	$NMSE_L$	$NMSE_R$	$ACPR_{m1L}$	$ACPR_{p1L}$	$ACPR_{m1R}$	$ACPR_{p1R}$	$EVM_L$	$EVM_R$	Pout	PAE
	(dB)		(dBc)				(%)		(dBm)	(%)
Sin DPD	-17.9	-19.6	-25.1	-27.4	-27.5	-28.3	10.4	8.5	29.6	77.6
Sin técnica	-33.1	-34.5	-39.3	-40.5	-41.1	-42.5	2.1	1.8	29.0	77.5
DOMP	-33.3	-34.8	-36.9	-41.2	-37.9	-43.3	2.1	1.8	28.1	76.5
Ridge	-33.5	-34.3	-39.7	-42.2	-41.5	-43.2	2.0	1.8	28.1	75.2
$\mu$ dinámico	-33.5	-35.1	-39.8	-41.9	-41.5	-43.2	2.0	1.7	28.5	72.7
$\mu$ + Ridge	-31.1	-32.6	-36.7	-39.2	-38.9	-41.1	2.5	2.1	28.1	75.2

siendo la mejor opción darle el valor máximo, 1, en la mayoría de las iteraciones, con unos valores bajos al final para el ajuste fino.

Los resultados obtenidos para cada caso se muestran en la Tabla 6.1. Es sobre todo interesante observar como la mejor técnica en este caso es el  $\mu$  dinámico, seguido de la regularización de Ridge, y muy de cerca del algoritmo DOMP, el cual es usado seleccionando los primeros 50 regresores y no el criterio Bayesiano, lo cual limita su funcionamiento. Muy de cerca está el DPD sin aplicar técnicas adicionales y es que el laboratorio de Chalmers ya sea por las ruidosas medidas o el tipo de amplificador, no necesita tanto de una técnica para lograr resultados cercanos a lo mejor que se puede obtener con estas, incluso funcionando mejor que la regularización de Ridge con *cooling*, al no conseguir converger en el punto óptimo. El tipo de función para el  $\mu$  dinámico usado fue una función exponencial  $\mu = 2^{-n}$ .

### 6.3 Laboratorio de Sistemas de Radiocomunicación de la Universidad de Sevilla

Las medidas tomadas de este laboratorio fueron un tanto diferentes a las de Chalmers. De forma contraria al laboratorio virtual de Chalmers que se accede de forma online, los aparatos del laboratorio de Sistemas de Radiocomunicación están conectados de forma local a los equipos allí presentes, y la forma de realizar las medidas es conectarse remotamente a los equipos de tal sala.

#### 6.3.1 Montaje de Sevilla

El hecho de no poder enviar las medidas traía el inconveniente de ser incapaz de usar un ordenador personal para todo el procedimiento, inconveniente que no lo era tanto tomando en cuenta que el ordenador de sobremesa del laboratorio es bastante más potente que cualquier ordenador portátil habitual. Otras ventajas eran que al generarse las señales en el equipo que está muy próximo a los aparatos de medida es que al no tener que enviar esas largas señales de datos se acortaban mucho los tiempos para medir, haciendo el proceso mucho más ágil.

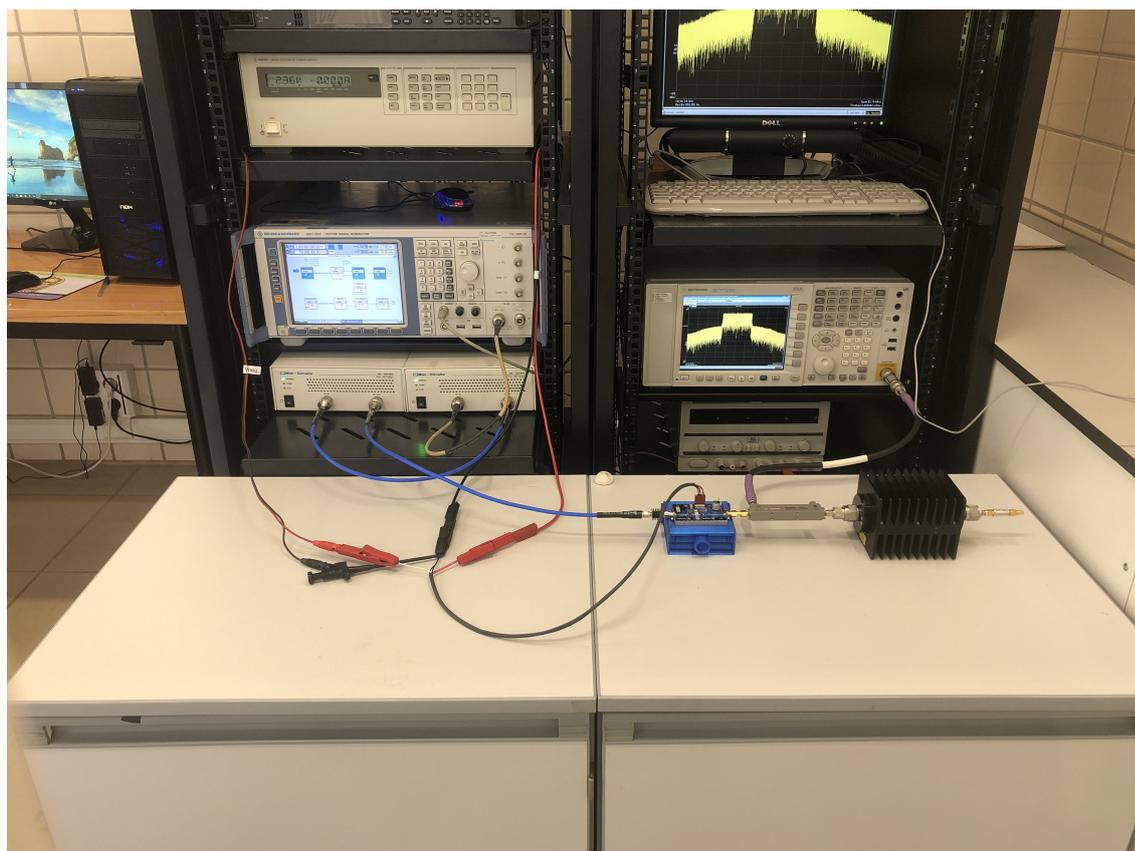
Similar al laboratorio de Chalmers, el montaje de Sevilla consiste en un conjunto de preamplificadores, seguido del dispositivo bajo prueba (DUT), para colocar después un acoplador conectando su salida a un atenuador terminado en una carga adaptada. La salida acoplada se conecta al analizador vectorial de señal para obtener las medidas de las señales de salida. En la Figura 6.4 se pueden ver una imagen del laboratorio con los equipos usados.

Como se ha comentado previamente, los parámetros de las señales OFDM 5G-NR usadas aquí varían ligeramente. El número de bloques de recursos se reduce a la mitad, 25, mientras que las frecuencias de *offset* y anchos de banda también, para  $\pm 20$  MHz y 10 MHz respectivamente. Aparte, el factor de sobremuestreo toma el valor de 3.

Para las medidas, los modelos usados en este caso serán los de 534 regresores, mientras que los valores de  $\mu$  se fijarán a  $1/3$ . Con estos cambios y ofreciendo mejor calidad a la hora de medir, se conseguirán muy buenos resultados de linealización.

#### 6.3.2 Medidas experimentales

En la Tabla 6.2 se pueden ver los resultados obtenidos en el laboratorio de Sevilla. Como se podría esperar, los valores conseguidos con DOMP y regularización de Ridge son muy superiores al resto, estando la regularización de Ridge mínimamente por encima, otra vez por la limitación en el número de regresores. En el caso del DPD sin aplicar técnicas adicionales se observa una clara divergencia en la banda izquierda,



**Figura 6.4** Montaje del laboratorio de Sistemas de Radiocomunicación de la Universidad de Sevilla.

**Tabla 6.2** Indicadores de linealización para diferentes casos de DPD-DLA en Sevilla.

Caso	$NMSE_L$	$NMSE_R$	$ACPR_{m1L}$	$ACPR_{p1L}$	$ACPR_{m1R}$	$ACPR_{p1R}$	$EVM_L$	$EVM_R$	Pout	PAE
	(dB)		(dBc)				(%)		(dBm)	(%)
Sin DPD	-30.5	-30.0	-39.3	-39.1	-38.7	-38.7	2.6	2.7	30.1	17.9
Sin técnica	-35.7	-44.0	-44.0	-44.3	-55.0	-54.9	1.5	1.1	29.7	16.8
DOMP	-49.6	-50.3	-62.5	-63.1	-62.9	-62.0	1.1	1.0	29.7	16.7
Ridge	-50.1	-50.7	-63.6	-64.1	-63.4	-62.6	1.1	1.0	30.1	18.1

la cual provoca la parada del algoritmo al incumplir los requisitos de PAPR. Se puede también ver la gran mejora de los sistemas DPD de doble banda concurrente en estas señales, obteniendo hasta 20 dB de mejora en el NMSE en cada banda. Para acabar, en la Figura 6.5 se puede ver una captura del analizador vectorial mostrando las señales que se están enviando junto con parámetros tales como el NMSE o ACPR. En las ventanas superiores se puede observar la señal en frecuencia, en la inferior izquierda en tiempo mientras que la derecha muestra ciertos valores de la señal.

## 6.4 Uso de técnicas de mejora del DPD dual

Para acabar con los resultados experimentales resta comentar un par de detalles sobre la divergencia, la tasa de aprendizaje y las técnicas usadas.

El problema de la divergencia es muy común en este tipo de sistemas, aunque no lo parezca por los resultados mostrados ya que han sido seleccionados los que se han tomado usando técnicas varias para contrarrestar este hecho, y por lo cual han sido elegidos tasas de aprendizaje  $\mu$  bajas, de 1/4 o 1/3. Una muestra de los efectos de la divergencia se puede ver en la Figura 6.6 donde se observa como la banda derecha esta comenzando a empeorar hasta que se activa el limitador del amplificador y no se puede medir más.

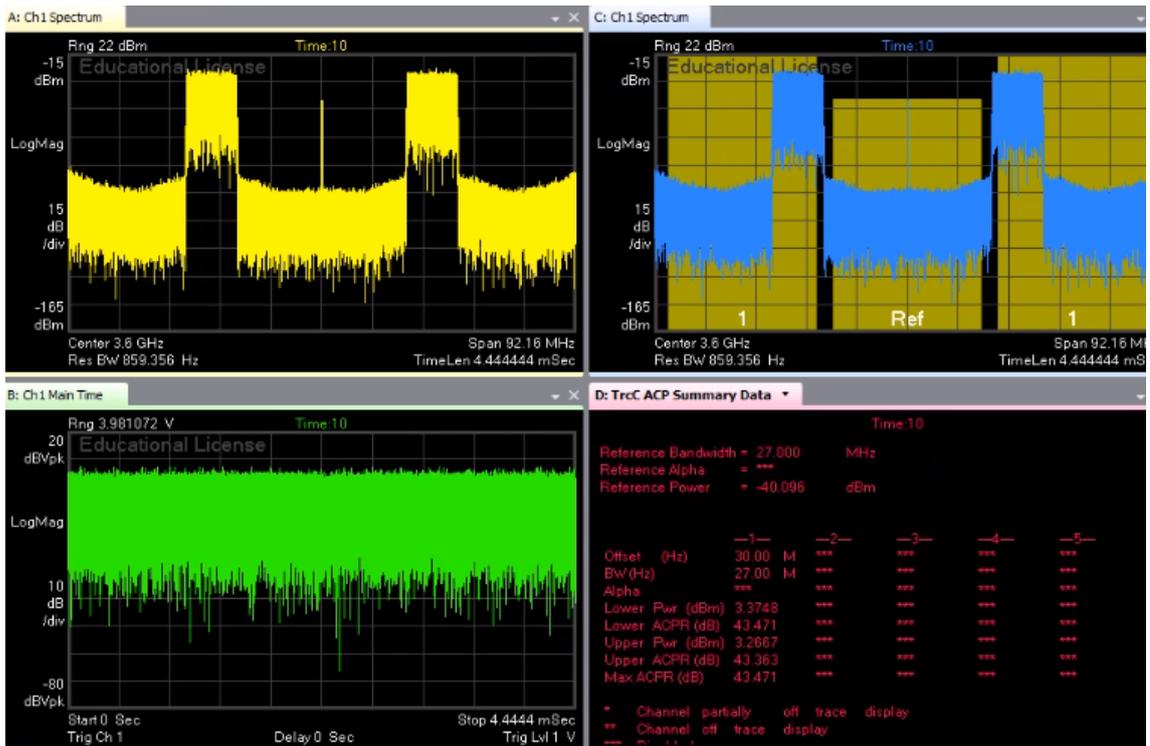


Figura 6.5 Muestra de la señal en el analizador de espectro.

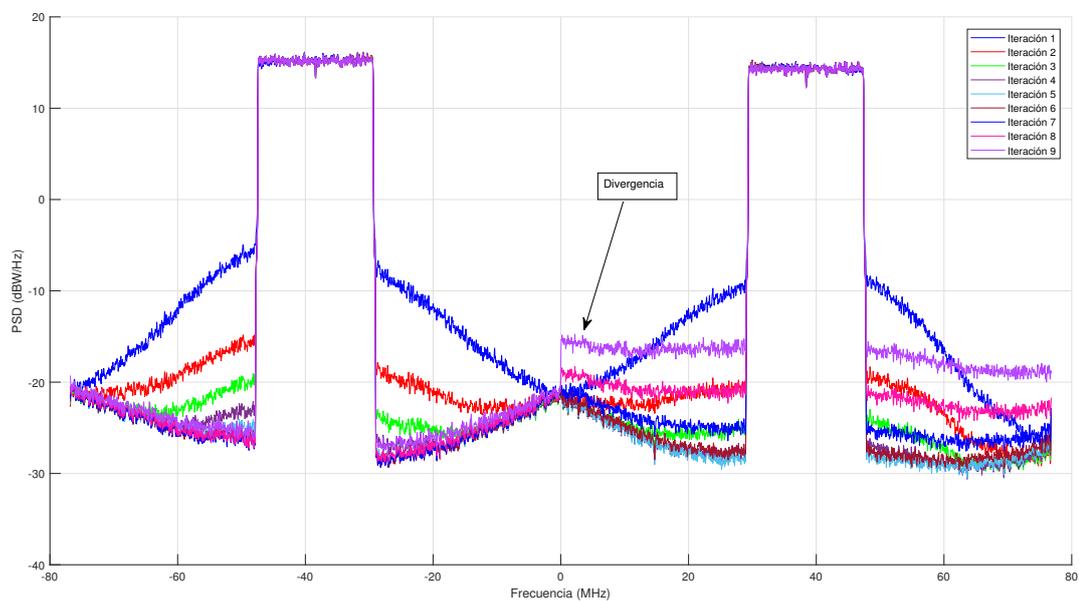
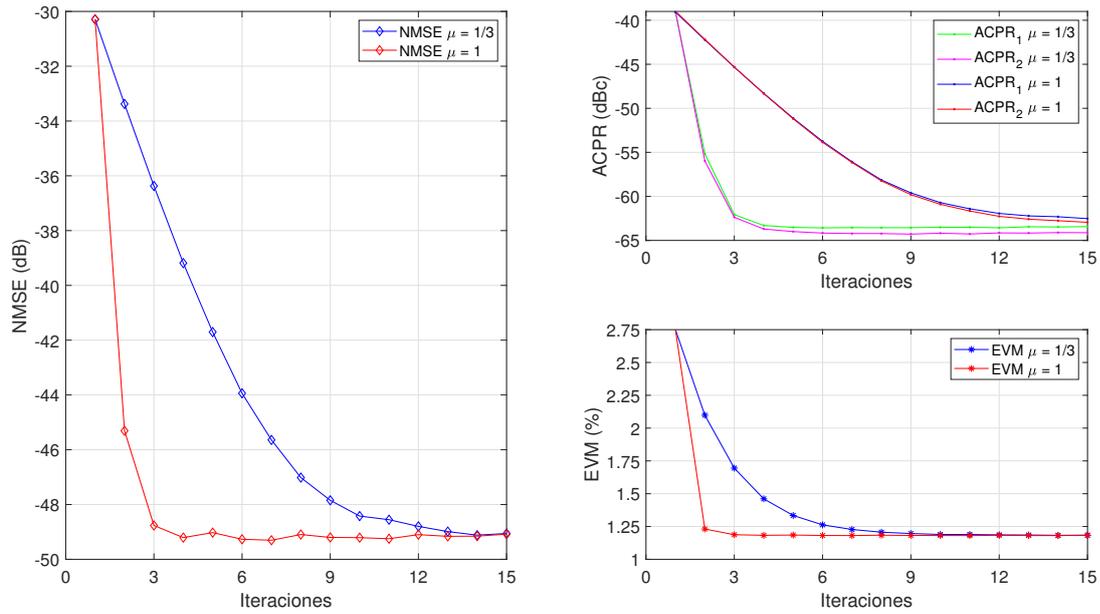


Figura 6.6 Efecto de la divergencia en el espectro.

Pero, al usar DOMP y regularización de Ridge la divergencia ya no es un problema por lo que es buena idea maximizar el valor de  $\mu$  en las primeras etapas para acercarse mucho al mínimo y acabar con algunas etapas de  $\mu$  bajo y así atinar más en el mínimo. Es por esta razón que usar un  $\mu$  adaptativo es muy buena idea en un caso sin divergencia, como en Chalmers y usando un modelo pequeño, mientras que usarlo con modelos más grandes como en el caso de Sevilla puede traer mucha divergencia.

Este comentario se debe a los no mostrados valores de las figuras de mérito a lo largo de las iteraciones,



**Figura 6.7** Comparación de la evolución de NMSE, ACPR y EVM con  $\mu$  igual a  $1/3$  y  $1/2$ .

puendiéndose ver el ritmo que toman al reducirse y poder comparar. Por ello, en la Figura 6.7 se muestran las figuras de mérito tomadas con diferentes parámetros de aprendizaje.



## 7 Conclusiones

---

*Like the legend of the phoenix  
Our ends were beginnings*

DAFT PUNK, GET LUCKY

En este Trabajo Fin de Grado se ha pretendido realizar una introducción lo más general posible a los sistemas de predistorsión, pasando desde los bloques más básicos a la vez que fundamentales, como pueden ser el tipo de señal, OFDM, o los amplificadores de potencia, hasta las técnicas de predistorsión más avanzadas y punteras como la regularización de Ridge.

Por el camino han sido presentados importantes detalles de los sistemas DPD, como las arquitecturas indirecta y directa, escogiendo la directa como base para el trabajo, mientras que la implementación se ha realizado con la situación de la doble banda concurrente en mente, debido a la importancia de este tipo de esquemas en las comunicaciones actuales, cada vez más demandantes de ancho de banda y con necesidad de hacer uso de varias bandas de información a la vez. Este caso, al tener que luchar contra las intermodulaciones y dos señales a la vez, dispara las variables del sistema que se deben manejar, pero aun así logrando muy buenos resultados. Otros importantes bloques comentados a lo largo del trabajo han sido como encontrar los parámetros óptimos para el mejor funcionamiento del sistema, la generalización al modelo dual de cada uno de los bloques constituyentes del DPD directo y varias técnicas muy potentes de apoyo al DPD, como son las técnicas de selección de coeficientes, encabezadas por el algoritmo DOMP, las técnicas de ajuste de la tasa de aprendizaje o las de regularización, donde se trata en profundidad la de Ridge.

En cuanto a los resultados obtenidos, han sido muy satisfactorios. La implementación del esquema dual concurrente directo ha sido un éxito, tanto en las medidas realizadas en los laboratorios de Chalmers University of Technology como en el del Grupo de Sistemas de Radiocomunicación de la Universidad de Sevilla. En general, casi todas las configuraciones han conseguido lograr buenos resultados por si mismas, haciendo una mención especial a la regularización de Ridge, técnica que ha sido parte de una implementación conjunta de este Trabajo Fin de Grado junto con mi compañero en el grupo de investigación, cuyo tema era este tipo de regularización, combinando ambos proyectos y de esta manera realizando un artículo entre ambos con ayuda de nuestros tutores para el XXXVI Simposio Nacional de la Unión Científica Internacional de Radio que se celebrará en septiembre de 2021.

Finalmente, las líneas futuras son muy diversas e interesantes. Entre ellas están el estudio de posibles técnicas de tasa de aprendizaje dinámica, en las cuales el parámetro  $\mu$  sea calculado en cada iteración. Otra línea interesante sería la investigación de alternativas a la pseudo inversa de Moore-Penrose para el cálculo de los coeficientes. También, otro paso puede ser la adaptación del sistema dual al uso de dos generadores para disponer de toda la tasa de muestreo posible. Para acabar, la implementación de técnicas de *deep learning* para cualquiera de las partes de la predistorsión digital sería muy atractivo, ya sea en una de las principales como la generación de regresores y cálculo de coeficientes, la regresión en general o de otras más pequeñas pero que pueden arrojar una gran diferencia como la selección de coeficientes. Viendo el avance de las tecnologías de aprendizaje profundo, sobre todo en el análisis de secuencias de mano de las redes neuronales de tipo *transformers* y la posible generalización de los sistemas con *transfer learning*, incluso podría ser realizable el cálculo directo de la señal predistorsionada mediante el único *input* de la entrada del sistema a la red.



# Apéndice A

## Códigos MatLab

Preparación inicial.....	1
Definición de funciones inline.....	1
Carga de modelos.....	1
Configuración parámetros principales.....	1
Generacion señal OFDM .....	2
Primera medida e inicialización .....	2
Cálculo de búfers de memoria máximos (Qpmax) y (Qnmax), 2D-GMP.....	3
Generación de las matrices de regresores.....	3
Inicialización de vectores de pesos .....	4
Cálculo de lambda_opt .....	4
Cálculo de matriz inversa .....	4
Bucle DLA.....	4
Validación.....	6
Guardado de datos.....	7

### Preparación inicial

```
clearvars,clc,close all;  
addpath('toolbox'); addpath('toolbox_Chalmers');
```

### Definición de funciones inline

```
nmse = @(ymod,y) 20*log10(norm(ymod-y)/norm(y));  
dBm = @(x) 10*log10(rms(x).^2/100)+30;  
scale_dBm = @(x,P) x*10^((P-dBm(x))/20);  
PAPR = @(x) 20*log10(max(abs(x))/rms(x));  
dBminst = @(x) 10*log10(abs(x).^2/100)+30;
```

### Carga de modelos

```
cd('confset'); configuration; cd ..; %234 reg  
% cd('confset'); configurationB; cd ..; %534 reg
```

### Configuración parámetros principales

```
UsarRidge = false; % Bandera para controlar el uso de Ridge  
UsarDomp = true; % Bandera para controlar el uso de DOMP  
RMSin = -22.5; % Potencia de entrada  
  
Nmeas = 5; % Número de medidas  
Niter = 10; % Número de iteraciones  
Nval = 3; % Número de validaciones  
Nave = 10; % Muestras para promediado  
mu = 0.25*ones(1,Niter); % Coeficiente mu estático  
% mu = 2.^(-1:-10); % Función mu dinámico 2^-n  
% mu = 1:-0.1:0; % Función mu lineal -0.1*n+1
```

```

ovs = 5; % OverSampling
fsgen = 1024*30e3*ovs; % Frecuencia de generación
fsmed = fsgen; % Frecuencia de medida
NRB = 50; % Número de bloques de recursos

```

```
% 50 NRB -> 20 MHz. 25 NRB -> 10 MHz
```

## Generación señal OFDM

Foff = [-20e6,20e6]; % Varias configuraciones de Foff para evitar Foff = [-22e6,22e6]; % el aliasing.

```

Foff = [-fsgen/4,fsgen/4];

% Generación OFDM para Sevilla (1) y Chalmers (2)
% [u, us, ~, datau] = genera_5GNR_multicarrier_v4([NRB,NRB], [30e3,30e3],[10e6,10e6],
Foff, fsgen, [8,8], [2,1]);
[u, us, ~, datau] = genera_5GNR_multicarrier_v4([NRB,NRB], [30e3,30e3],[20e6,20e6],
Foff, fsgen, [8,8], [2,1]);
usL = us(:,1);
usR = us(:,2);
fsint = fsgen/2; % Frecuencia a la que se bajarán las bandas

imed = 1;
medida_out = [];

% Fechas para guardado de datos
filenamedate = datestr(clock,30);
filename = ['measurements' filesep 'exp_' filenamedate];
figfilename = ['results' filesep 'exp_' filenamedate];

```

## Primera medida e inicialización

```

u = scale_dBm(u,RMSin); % Ajuste de potencia de la señal

% Medida con PA, para Sevilla (1) y Chalmers (2)
% [y, x, data] = myPA(u, dBm(u), Nmeas, Nave);
[y, x, data] = myPACHalmers160(u, RMSin, Nmeas);

% Separación señales en bandas Izq y Der
[uLS, uRS] = separate_bands(u,fsmed,fsint,Foff);
[yLS, yRS] = separate_bands(y,fsmed,fsint,Foff);
Pout = dBm(y);
Pin = dBm(u);

% Cálculo de la ganancia estática
G = norm(y)/norm(u);

% Cálculo figuras de mérito
analiza_medidas_OFDM_B_MIG(uLS,yLS,datau,0,fsint,fsgen,99);
analiza_medidas_OFDM_B_MIG(uRS,yRS,datau,0,fsint,fsgen,101);

```

## Cálculo de búfers de memoria máximos (Qpmax) y (Qnmax), 2D-GMP

```
Qa1 = max([max(x1_limits.intra.La), max(x1_limits.cross.La), max(x1_limits.mixed.La)]);
Qa2 = max([max(x2_limits.intra.La), max(x2_limits.cross.La), max(x2_limits.mixed.La)]);
Qb1 = max([max(x1_limits.intra.Lb) + max(x1_limits.intra.Mb), ...
          max(x1_limits.cross.Lb) + max(x1_limits.cross.Mb), ...
          max(x1_limits.mixed.Lb) + max(x1_limits.mixed.Mb)]);
Qb2 = max([max(x2_limits.intra.Lb) + max(x2_limits.intra.Mb), ...
          max(x2_limits.cross.Lb) + max(x2_limits.cross.Mb), ...
          max(x2_limits.mixed.Lb) + max(x2_limits.mixed.Mb)]);
Qc1 = max([max(x1_limits.intra.Lc) - min(x1_limits.intra.Mc), ...
          max(x1_limits.cross.Lc) - min(x1_limits.cross.Mc), ...
          max(x1_limits.mixed.Lc) - min(x1_limits.mixed.Mc)]);
Qc2 = max([max(x2_limits.intra.Lc) - min(x2_limits.intra.Mc), ...
          max(x2_limits.cross.Lc) - min(x2_limits.cross.Mc), ...
          max(x2_limits.mixed.Lc) - min(x2_limits.mixed.Mc)]);

Qpmax1 = max([Qa1, Qb1, Qc1]);
Qpmax2 = max([Qa2, Qb2, Qc2]);

Qnmax1 = max([x1_limits.intra.Mc, x1_limits.cross.Mc, 0]);
Qnmax2 = max([x2_limits.intra.Mc, x2_limits.cross.Mc, 0]);

N = length(yLs);
```

## Generación de las matrices de regresores

```
[UL, RmatL] = buildw_2GMP(uLs, uRs, [N-Qpmax1+1:N, 1:N, 1:Qnmax1], Qpmax1, Qnmax1,
x1_limits);
[UR, RmatR] = buildw_2GMP(uRs, uLs, [N-Qpmax2+1:N, 1:N, 1:Qnmax2], Qpmax2, Qnmax2,
x2_limits);

if UsarDomp == true

    % Selección de índices para reducción del tiempo de cómputo de
    % DOMP y RC-DOMP
    % Uso de DOMP o RC-DOMP dependiendo del objetivo del sistema
    indicesL = sel_indices(yLs,yLs,.1);
    % [~, sL, noptL, ~, ~] = RCDOMP(UL(indices,:), yLs(indices), RmatL, 51); %Me ahorro
    hacer los 500 porque solo me quedo con 50
    indicesR = sel_indices(yRs,yRs,.1);
    % [~, sR, noptR, ~, ~] = RCDOMP(UR(indices,:), yRs(indices), RmatR, 51);
    [~, sL, noptL, ~, ~, ~] = omp_domp(UL(indicesL,:), yLs(indicesL), RmatL, 'DOMP');
    [~, sR, noptR, ~, ~, ~] = omp_domp(UR(indicesR,:), yRs(indicesR), RmatR, 'DOMP');

    CL = noptL;
    CR = noptR;
    SL = sL(1:CL).';
    SR = sR(1:CR).';

else
    [~,cL]=size(UL);
    [~,cR]=size(UR);
    SL = 1:cL;
    SR = 1:cR;
end
```

## Inicialización de vectores de pesos

```
wL = zeros(cL,1);
dwL = zeros(cL,1);
wR = zeros(cR,1);
dwR = zeros(cR,1);
```

## Cálculo de lambda\_opt

Cálculo por medio de la función de búsqueda de mínimo fminsearch

```
if UsarRidge == true
    options = optimset('Display','iter','PlotFcns',@optimplotfval);
    [lambda_opt_L, NMSE_L_opt] = fminsearch(@(lambda)
    ridge_function(UL(:,sL),yLs,length(RmatL),lambda),10^-6,options);
    [lambda_opt_R, NMSE_R_opt] = fminsearch(@(lambda)
    ridge_function(UR(:,sR),yRs,length(RmatR),lambda),10^-6,options);
    % Seguridad por si el algoritmo da un lambda_opt negativo
    if lambda_opt_L < 0
        lambda_opt_L = 10^-8;
    elseif lambda_opt_R < 0
        lambda_opt_R = 10^-8;
    end
end
```

## Cálculo de matriz inversa

Uso de la matriz normalizada solo si se usa Ridge, aunque se puede usar en el caso sin Ridge para un aumento en estabilidad.

```
fprintf('\nCálculo matriz inversa \n');
if UsarRidge == true
    uL_pot_col = vecnorm(UL(:,sL)); %Potencia de cada colum
    UL_norm = UL(:,sL)*diag(uL_pot_col.^(-1));
    uR_pot_col = vecnorm(UR(:,sR)); %Potencia de cada colum
    UR_norm = UR(:,sR)*diag(uR_pot_col.^(-1));

    UL_inv = (inv(UL_norm'*UL_norm+eye(cL)*lambda_opt_L)*UL_norm');
    UR_inv = (inv(UR_norm'*UR_norm+eye(cR)*lambda_opt_R)*UR_norm');
else
    UL_norm = UL(:,sL);
    UR_norm = UR(:,sR);

    UL_inv = (inv(UL_norm'*UL_norm)*UL_norm');
    UR_inv = (inv(UR_norm'*UR_norm)*UR_norm');
end
```

## Bucle DLA

```
figure(30),title('y');

nmseL = [];
nmseR = [];

for iter=1:Niter
```

```

fprintf(['Iter: ' num2str(iter) '\n']);

% Generación señales predistorcionadas
xLS = uLS-UL(:,sL)*wL;
xRS = uRS-UR(:,sR)*wR;

% Unión bandas para la medida del PA
x = unite_bands(xLS,xRS,fsint,fsmed,Foff);

% Clipping
PAPRx = 20*log10(max(abs(x))/rms(x));
disp(sprintf('PAPR antes de clipping: %4.2f dB\n', PAPRx));
PAPRd = 12;
clip = (10^((PAPRd-PAPRx)/20))*max(abs(x));
x(abs(x)>clip) = clip*exp(1i*angle(x(abs(x)>clip)));
disp(sprintf('PAPR después de clipping: %4.2f dB\n', PAPR(x)));

% Medida PA
[y, x, data] = myPA(x, dBm(x), Nmeas, Nave);
[y, x, data_med] = myPACHalmers160(x, RMSin, Nmeas);

% Cálculo y representación espectro de potencia
espectro(y - mean(y),fsgen,30);

%Separación de bandas
[xLS, xRS] = separate_bands(x,fsmed,fsint,Foff);
[yLS, yRS] = separate_bands(y,fsmed,fsint,Foff);

% Normalización de ganancia y cálculo de error
eL = yLS/GL-uLS;
eR = yRS/GR-uRS;

% Si se normalizaron los regresores, desnormalización
if UsarRidge == true
    dwL = diag(uL_pot_col.^(-1))*UL_inv*eL;
    dwR = diag(uR_pot_col.^(-1))*UR_inv*eR;
else
    dwL = UL_inv*eL;
    dwR = UR_inv*eR;
end

% Adaptación DPD
wL = wL + mu(iter) * dwL;
wR = wR + mu(iter) * dwR;

% Gráfica AM-AM
figure(10); plot(dBminst(u), dBminst(y)- dBminst(u),'.');
hold on;

% Cálculo figuras de mérito
[PerformanceL] = analiza_medidas_OFDM_B_MIG(uLS,yLS,datau,0,fsint,fsgen,99);
[PerformanceR] = analiza_medidas_OFDM_B_MIG(uRS,yRS,datau,0,fsint,fsgen,101);
medida.descr = ['Itera : ' num2str(iter)];
medida.PerformanceL = PerformanceL;
medida.PerformanceR = PerformanceR;
medida.x = x;
medida.y = y;

```

```

medida.data = data;
medida_out = [medida_out medida];
medida2figureJA(medida_out)
end

```

## Validación

```

for iter = 1:Nval

    % Generación señales de validación
    [uval, ~, ~, datau] = genera_5GNR_multicarrier_v4([NRB,NRB],
[30e3,30e3],[10e6,10e6], Foff, fsgen, [8,8], [2*iter+101,2*iter+100]);
    uval = scale_dBm(uval,RMSin);

    % Separación de bandas
    [uLs, uRs] = separate_bands(uval,fsmed,fsint,Foff);

    % Matrices de regresores
    [UL, RmatL] = buildw_2GMP(uLs, uRs, [N-Qpmax1+1:N , 1:N , 1:Qnmax1], Qpmax1, Qnmax1,
x1_limits);
    [UR, RmatR] = buildw_2GMP(uRs, uLs, [N-Qpmax2+1:N , 1:N , 1:Qnmax2], Qpmax2, Qnmax2,
x2_limits);

    % Predistorsión
    xLS = uLs-UL(:,sL)*wL;
    xRS = uRs-UR(:,sR)*wR;

    % Unión de bandas
    x = unite_bands(xLS,xRS,fsint,fsmed,Foff);

    % Clipping
    PAPRx = 20*log10(max(abs(x))/rms(x));
    disp(sprintf('PAPR antes de clipping: %4.2f dB\n', PAPRx));
    PAPRd = 12;
    clip = (10^((PAPRd-PAPRx)/20))*max(abs(x));
    x(abs(x)>clip) = clip*exp(1i*angle(x(abs(x)>clip)));
    disp(sprintf('PAPR después de clipping: %4.2f dB\n', PAPR(x)));

%    [y, x, data] = myPA(x, dBm(x), Nmeas, Nave);
    [y, x, data_med] = myPACHalmers160(x, RMSin, Nmeas);

    % Cálculo y representación espectro de potencia
    espectro(y - mean(y),fsgen,31);

    % Separación de bandas
    [yLs, yRs] = separate_bands(y,fsmed,fsint,Foff);

    % Gráfica AM-AM
    figure(11); plot(dBminst(uval), dBminst(y)- dBminst(uval),'.');
    hold on;

    % Cálculo figuras de mérito
    [PerformanceL] = analiza_medidas_OFDM_B_MIG(uLs,yLs,datau,0,fsint,fsgen,99);
    [PerformanceR] = analiza_medidas_OFDM_B_MIG(uRs,yRs,datau,0,fsint,fsgen,101);
    medida.descr = ['Itera_val : ' num2str(iter)];
    medida.PerformanceL = PerformanceL;
    medida.PerformanceR = PerformanceR;

```

```
medida.x = x;  
medida.y = y;  
medida.data = data;  
medida_out = [medida_out medida];  
medida2figureJA(medida_out)
```

end

Guardado de datos

```
save(filename,'medida_out','RMSin','model_id','sL','sR','mu','lambda_opt_L','lambda_opt_R');
```

*Published with MATLAB® R2021a*



# Apéndice B

## Modelos 2D-GMP Usados

Modelo de 234 coeficientes.....	1
Modelo de 534 coeficientes.....	2

### Modelo de 234 coeficientes

```
model_id.tipo = '2D-GMP';
model_id.pe = 1;
model_id.grafica = 0;
model_id.h1 = [];
model_id.h2 = [];

%Model coeff
x1_limits.intra.Ka = [0 1 2 3 4 5 6 7];
x1_limits.intra.La = [5 3 3 3 1 1 1 1];
x1_limits.intra.Kb = [1 2 3 4 5];
x1_limits.intra.Lb = [2 2 1 1 1];
x1_limits.intra.Mb = [1 1 1 1 1];
x1_limits.intra.Kc = [1 2 3 4 5];
x1_limits.intra.Lc = [2 2 1 1 1];
x1_limits.intra.Mc = [1 1 1 1 1];

x1_limits.cross.Ka = [1 2 3 4 5 6 7]; %Ka-cross sin orden 0!
x1_limits.cross.La = [3 3 3 1 1 1 1];
x1_limits.cross.Kb = [1 2 3 4 5];
x1_limits.cross.Lb = [2 2 1 1 1];
x1_limits.cross.Mb = [1 1 1 1 1];
x1_limits.cross.Kc = [1 2 3 4 5];
x1_limits.cross.Lc = [2 2 1 1 1];
x1_limits.cross.Mc = [1 1 1 1 1];

x2_limits.intra.Ka = [0 1 2 3 4 5 6 7];
x2_limits.intra.La = [5 3 3 3 1 1 1 1];
x2_limits.intra.Kb = [1 2 3 4 5];
x2_limits.intra.Lb = [2 2 1 1 1];
x2_limits.intra.Mb = [1 1 1 1 1];
x2_limits.intra.Kc = [1 2 3 4 5];
x2_limits.intra.Lc = [2 2 1 1 1];
x2_limits.intra.Mc = [1 1 1 1 1];

x2_limits.cross.Ka = [1 2 3 4 5 6 7]; %Ka-cross sin orden 0!
x2_limits.cross.La = [3 3 3 1 1 1 1];
x2_limits.cross.Kb = [1 2 3 4 5];
x2_limits.cross.Lb = [2 2 1 1 1];
x2_limits.cross.Mb = [1 1 1 1 1];
x2_limits.cross.Kc = [1 2 3 4 5];
x2_limits.cross.Lc = [2 2 1 1 1];
x2_limits.cross.Mc = [1 1 1 1 1];

x1_limits.mixed.Ka = [1 2 3 4 5 6 7 ]; %NOT USED for extended
x1_limits.mixed.La = [3 3 3 1 1 1 1 ];
x1_limits.mixed.Kb = [2 3 4 5 6 7 8 9]; %Not the exponents but the regressor order (min
value is 2)
x1_limits.mixed.Lb = [2 2 2 2 2 2 2 2];
x1_limits.mixed.Mb = [1 1 1 1 1 1 1 1];
x1_limits.mixed.Kc = [1 2 3 4 5 6 7 8 9];
x1_limits.mixed.Lc = [2 2 2 2 2 2 2 2];
x1_limits.mixed.Mc = [1 1 1 1 1 1 1 1]; 53
```

```

x2_limits.mixed.Ka = [1 2 3 4 5 6 7]; %NOT USED for extended
x2_limits.mixed.La = [3 3 3 1 1 1 1];
x2_limits.mixed.Kb = [2 3 4 5 6 7 8 9]; %Not the exponents but the regressor order (min
value is 2)
x2_limits.mixed.Lb = [2 2 2 2 2 2 2];
x2_limits.mixed.Mb = [1 1 1 1 1 1 1];
x2_limits.mixed.Kc = [1 2 3 4 5 6 7 8 9];
x2_limits.mixed.Lc = [2 2 2 2 2 2 2];
x2_limits.mixed.Mc = [1 1 1 1 1 1 1];

model_id.x1_limits = x1_limits;
model_id.x2_limits = x2_limits;

```

## Modelo de 534 coeficientes

```

model_id.tipo = '2D-GMP';
model_id.pe = 1;
model_id.grafica = 0;
model_id.h1 = [];
model_id.h2 = [];

%Model coeff
x1_limits.intra.Ka = [0 1 2 3 4 5 6 7 8 9 10 11 12 13];
x1_limits.intra.La = [5 3 3 3 1 1 1 1 1 1 1 1 1 1];
x1_limits.intra.Kb = [1 2 3 4 5];
x1_limits.intra.Lb = [2 2 1 1 1];
x1_limits.intra.Mb = [1 1 1 1 1];
x1_limits.intra.Kc = [1 2 3 4 5];
x1_limits.intra.Lc = [2 2 1 1 1];
x1_limits.intra.Mc = [1 1 1 1 1];

x1_limits.cross.Ka = [1 2 3 4 5 6 7 8 9 10 11 12 13]; %Ka-cross sin orden 0!
x1_limits.cross.La = [3 3 3 1 1 1 1 1 1 1 1 1 1];
x1_limits.cross.Kb = [1 2 3 4 5];
x1_limits.cross.Lb = [2 2 1 1 1];
x1_limits.cross.Mb = [1 1 1 1 1];
x1_limits.cross.Kc = [1 2 3 4 5];
x1_limits.cross.Lc = [2 2 1 1 1];
x1_limits.cross.Mc = [1 1 1 1 1];

x2_limits.intra.Ka = [0 1 2 3 4 5 6 7 8 9 10 11 12 13];
x2_limits.intra.La = [5 3 3 3 1 1 1 1 1 1 1 1 1];
x2_limits.intra.Kb = [1 2 3 4 5];
x2_limits.intra.Lb = [2 2 1 1 1];
x2_limits.intra.Mb = [1 1 1 1 1];
x2_limits.intra.Kc = [1 2 3 4 5];
x2_limits.intra.Lc = [2 2 1 1 1];
x2_limits.intra.Mc = [1 1 1 1 1];

x2_limits.cross.Ka = [1 2 3 4 5 6 7 8 9 10 11 12 13]; %Ka-cross sin orden 0!
x2_limits.cross.La = [3 3 3 1 1 1 1 1 1 1 1 1 1];
x2_limits.cross.Kb = [1 2 3 4 5];
x2_limits.cross.Lb = [2 2 1 1 1];
x2_limits.cross.Mb = [1 1 1 1 1];
x2_limits.cross.Kc = [1 2 3 4 5];
x2_limits.cross.Lc = [2 2 1 1 1];

```

```

x2_limits.cross.Mc = [1 1 1 1 1];

x1_limits.mixed.Ka = [1 2 3 4 5 6 7 8 9 10 11 12 13]; %NOT USED for extended
x1_limits.mixed.La = [3 3 3 1 1 1 1 1 1 1 1 1 1];
x1_limits.mixed.Kb = [2 3 4 5 6 7 8 9]; %Not the exponents but the regressor order (min
value is 2)
x1_limits.mixed.Lb = [2 2 2 2 2 2 2 2];
x1_limits.mixed.Mb = [1 1 1 1 1 1 1 1];
x1_limits.mixed.Kc = [1 2 3 4 5 6 7 8 9];
x1_limits.mixed.Lc = [2 2 2 2 2 2 2 2];
x1_limits.mixed.Mc = [1 1 1 1 1 1 1 1];

x2_limits.mixed.Ka = [1 2 3 4 5 6 7 8 9 10 11 12 13]; %NOT USED for extended
x2_limits.mixed.La = [3 3 3 1 1 1 1 1 1 1 1 1 1];
x2_limits.mixed.Kb = [2 3 4 5 6 7 8 9]; %Not the exponents but the regressor order (min
value is 2)
x2_limits.mixed.Lb = [2 2 2 2 2 2 2 2];
x2_limits.mixed.Mb = [1 1 1 1 1 1 1 1];
x2_limits.mixed.Kc = [1 2 3 4 5 6 7 8 9];
x2_limits.mixed.Lc = [2 2 2 2 2 2 2 2];
x2_limits.mixed.Mc = [1 1 1 1 1 1 1 1];

model_id.x1_limits = x1_limits;
model_id.x2_limits = x2_limits;

```

*Published with MATLAB® R2021a*



# Apéndice C

## Artículo Técnicas de CFR

### Técnicas de reducción de factor de cresta aplicadas a señales OFDM

Miguel Nogales González-Regueral  
Universidad De Sevilla  
Seville, Spain  
mignoggon@alum.us.es

**Abstract**— Las técnicas de reducción de factor de cresta (CFR) son ampliamente usadas en conjunto con las señales OFDM debido a su alta PAPR, de manera que la eficiencia del amplificador sea maximizada. En este artículo se revisarán algunos de los métodos más populares de CFR, divididos en técnicas de *clipping*, probabilísticas y de codificación.

**Keywords**— *cfr, ofdm, amplificador de potencia, papr*

#### I. INTRODUCCIÓN

Las técnicas de reducción de factor de cresta (CFR) son ampliamente usadas en conjunto con las señales OFDM debido a su alto ratio de potencia entre pico y media (PAPR) inherente en este tipo de ondas. Esta alta PAPR es muy negativa ya que, operando en la zona eficiente de un amplificador puede causar daños en el equipo además de entrar en la zona no lineal, causando efectos no deseados como emisiones fuera de banda. Una de las posibles soluciones podría ser realizar *backoff*, trabajar en un punto de operación de potencia menor, pero esto disminuiría la eficiencia del amplificador, lo cual no es deseable. Las técnicas de CFR aportarán distintas maneras de reducir la PAPR y poder operar en la zona deseada, afectando a la señal de diferentes maneras. Estas técnicas, que serán revisadas en el artículo, son: *Clipping*, *Peak Cancellation*, *Error Waveform Subtraction* y *Peak Windowing*, dentro del grupo de las técnicas de *Clipping and filtering*, *Partial Transmit Sequence*, *Selective Mapping*, *Tone Reservation*, y *Constellation Extension* dentro de las probabilísticas y *Coding*, donde se tratarán algunos tipos de codificación.

#### II. TÉCNICAS

##### A. Clipping

El *Clipping* (recorte) es una técnica de CFR que trata de suprimir los picos de la señal que pasen de cierto umbral, de manera que se pueda definir una PAPR deseada, y así asegurar que las señales OFDM a las que se les aplique esta técnica tengan como máximo ese valor. Esto se realiza de forma que toda muestra de la onda que supere en valor absoluto al umbral impuesto vea reducida su amplitud a la misma del umbral, asegurando así que la nueva PAPR valga exactamente lo mismo que la definida previamente. La gran ventaja del *Clipping* es su sencillez a la hora de la implementación, un detalle muy importante y puede llegar a ser el punto clave para muchos sistemas. En cuanto a las desventajas, aunque no escasas, tampoco son muy graves, entre ellas están las emisiones fuera de banda, solucionables con sistemas DPD o filtros, los cuales suelen ir acompañando a este tipo de señales, la necesidad de un *oversampling* mayor a cuatro y el deterioro del *error vector magnitude* (EVM), lo cual puede llegar a afectar mucho a señales con un gran ratio de bits por símbolo, por lo que hay que tener especial cuidado al usar modulaciones como la 64-QAM o

similares de órdenes superiores ya que los efectos se pueden volver problemáticos.

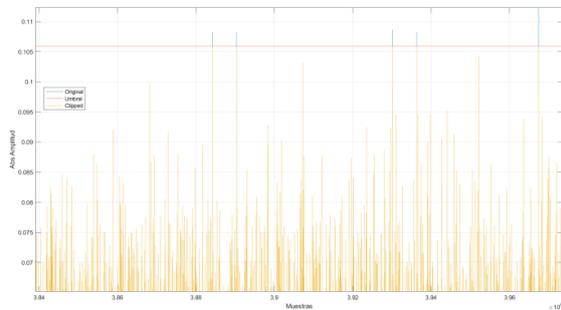


Figura 1: Efectos del *Clipping* sobre segmento de la señal OFDM

##### B. Peak Cancellation

La estrategia llamada *Peak Cancellation* trabaja con dos ondas, una de ellas es la propia señal OFDM y la otra será una señal complementaria de error, la cual estará compuesta de picos (deltas de Dirac) en las posiciones donde el umbral de PAPR deseada es sobrepasado. Esta señal de picos será escalada para que tenga amplitud equivalente a la diferencia entre la original y el valor deseado, de manera que al restar original con error se obtenga una versión de la señal con reducida PAPR al haber eliminado los picos indeseados. Esta técnica agrega bastantes elementos al concepto de *Clipping*, y, al igual que las siguientes técnicas, añade cierto grado de complejidad acompañado de mejoras en rendimiento.

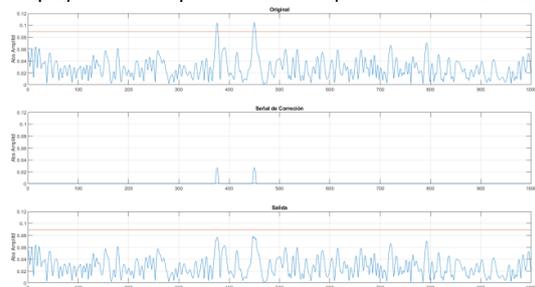


Figura 2: Muestras de las señales que intervienen en la técnica *Peak Cancellation*

Como pequeño punto negativo, este procedimiento no es tan eficiente si hay muchos picos seguidos, aunque la ocurrencia de estos es proporcional a la exigencia en la PAPR deseada, a mayor necesidad de reducirla más picos encontrará el algoritmo, por lo que es importante realizar una correcta elección del umbral. Como añadidos, este procedimiento se puede usar iterativamente además de poder crear múltiples señales de error, cada una con un filtro diferente, al no ser necesario usar el del canal, y en el

momento de la transmisión valorar cual da mejores resultados para que sea elegida como salida del sistema.

### C. Error Waveform Substraction

La siguiente técnica de *clipping and filtering* es la *Error Waveform Substraction* donde se busca generar una señal de corrección de error que cumpla ciertos requisitos para que al restarla se genere una salida con una PAPR menor. Esta forma de onda se crea a partir de la diferencia entre la señal habiéndole realizado anteriormente *Clipping* junto con la original, el resultado será procesado mediante un filtrado, el cual, como en *Peak Cancellation*, puede ser realizado por cualquiera entre muchos tipos de pulsos a elección del diseñador. Tras esto, se le aplica cierta ganancia que se debe buscar de manera que minimice la PAPR y así se le restará a la entrada del sistema para obtener la versión de reducido factor de cresta. Este procedimiento goza de las ventajas del anterior método, se puede iterar para lograr mejores resultados y usar un set de filtros para después escoger el mejor resultado entre todos, aunque también tiene ciertas desventajas a la hora de ser usado en señales OFDM por los exigentes requisitos en las emisiones fuera de banda, problema común en *clipping and filtering*. Otro problema es la incompatibilidad de esta técnica con la tecnología MIMO, al causar interferencias destructivas entre la señal de error y los flujos de datos cercanos, pero, para solucionar este problema en concreto se le puede realizar una modificación la cual distribuye el procesamiento directamente a las portadoras, mejorando el rendimiento y resolviendo el problema, pero aumentando mucho la complejidad.

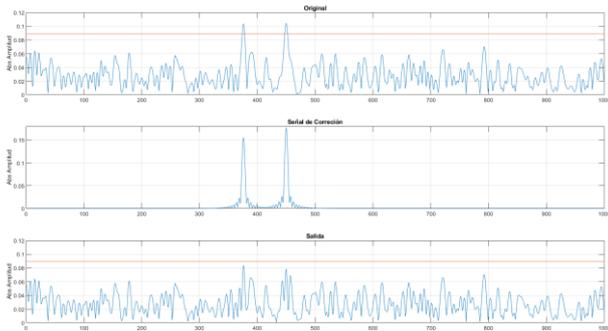


Figura 3: Comparativa de las diferentes señales que intervienen en la técnica *Error Waveform Substraction*.

### D. Peak Windowing

Otra alternativa a las anteriormente propuestas técnicas en las cuales se reducía la amplitud en los picos restando cierto valor constante o una señal de error es el *Peak Windowing*. En este método se compone una señal de atenuación, la cual al ser multiplicada por la señal original reducirá el nivel en las zonas problemáticas, dejando el resto intacto. En concreto, esta onda será generada detectando los picos, escalándolos y filtrándolos. Al haber realizado este procesamiento, le restaremos esta señal a un vector de unos, resultando una onda constante excepto en las zonas donde los picos hubiesen superado el umbral que se definió previamente donde tendrá un valor menor a uno inversamente proporcional a cuanto excedió el límite la señal original. Después solo faltará multiplicar esta señal

por la original y la PAPR se reducirá. Un detalle importante de todas estas técnicas que requieren de un filtrado es la posibilidad de elegir el pulso que más nos convenga entre los muchos que hay como pueden ser Gaussiano, Hann, Hamming o Kaiser, pero teniendo en cuenta que cada uno tendrá unas prestaciones diferentes en cuanto a emisión fuera de banda, efecto que también podemos controlar mediante la longitud del pulso. Generalmente, mientras mayor sea la ventana en el tiempo, menores sus efectos en frecuencia, en este caso emisiones fuera de banda, pero también será menor la energía transmitida lo cual, si la ventana es demasiado grande, acabaría por parecerse a realizar un *backoff*, lo cual no es deseable pues la eficiencia del amplificador decaería, y es justo lo que se debe evitar.

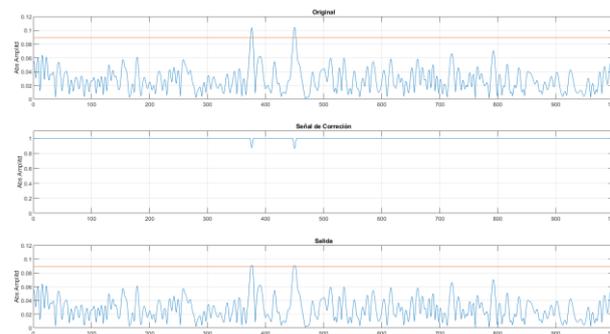


Figura 5: *Peak Windowing* sobre señal OFDM

### E. PTS y SLM

PTS (*Partial Transmit Sequence*) y SLM (*Selective Mapping*) son técnicas en las se aprovecha la libertad de realizar cambios en la fase de las portadoras para prevenir que se formen picos debidos a varias de estas coincidiendo en sus máximos, razón por la cual ocurre que la PAPR sea tan alta en este tipo de señales. En ambos métodos se multiplicará la señal OFDM por cierto vector conteniendo la información de las fases, en SLM antes de realizar la IFFT y en PTS después. Este vector de fases es importante ya que contiene toda la complejidad del algoritmo, tanto en implementación como en cómputo debido a que la búsqueda de la combinación óptima es exhaustiva. Para resolver este problema existen muchas variantes que intentan ofrecer una solución, como los métodos de búsqueda *adjacent* o *pseudorandom* [1], cuyo funcionamiento se basa en la prueba de un subconjunto de combinaciones creadas a priori, en la primera con conjuntos de bloques con el mismo cambio de fase y en el segundo se prueban combinaciones aleatorias. También existen otros algoritmos específicos que intentan sacar el máximo provecho de la técnica, basados en estrategias de optimización como el *Chicken Swarm* [2] o el uso de códigos convolucionales [3]. Otra alternativa es aplicar métodos de optimización no exclusivos de PTS y SLM, por ejemplo, algoritmos clásicos de búsqueda del óptimo o algoritmos genéticos, por poner algunos ejemplos, dependiendo la elección de la complejidad del problema (número de parámetros) y de la reducción de PAPR deseada.

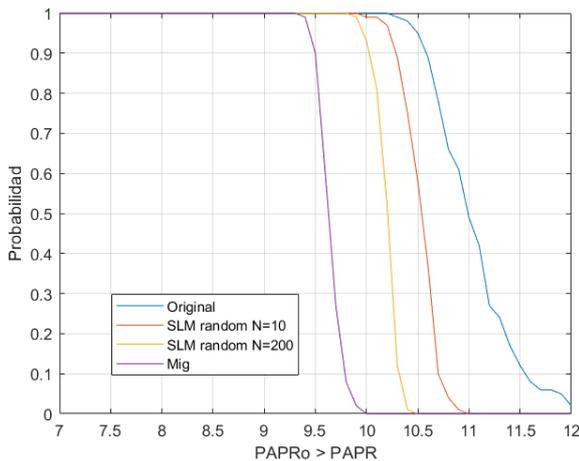


Figura 4: CCDF de varias técnicas de SLM

También se puede facilitar la búsqueda poniendo como meta no el vector óptimo si no el que cumpla cierto requerimiento de PAPR, parando la búsqueda en el momento que se haya satisfecho, lo cual ayuda mucho en velocidad de convergencia a estos métodos, y puede ser aplicado a todas las técnicas probabilísticas.

En general, estas técnicas son muy potentes, aunque conllevan una desventaja a tener en cuenta (a parte de la búsqueda del vector óptimo) y es que al realizar la implementación real se perderá algo de tasa de datos, esto es debido a que el vector de fases debe ser enviado junto a la señal para la demodulación, por ejemplo, en una portadora piloto, pero realmente es poco sacrificio para las buenas prestaciones que ofrecen ambas técnicas en lo referente a emisiones fuera de banda, EVM y reducción de PAPR.

#### F. Tone Reservation (TR)

Esta clase de método de CFR hace uso de un subconjunto de las portadoras de la señal OFDM para que sean usadas exclusivamente para disminuir la PAPR. Primero se deben seleccionar cuales portadoras se van a emplear para datos y cuales para la reducción de PAPR, donde se diferencian dos esquemas: localizado y distribuido, donde el primero de ellos concentrará todas las portadoras CFR en una zona en concreto y la otra las dispersará entre todas las opciones disponibles de manera que no formen bloques. Esta elección puede ser realizada cada vez que se genere una señal OFDM, ya que habrá una combinación concreta que ofrezca mejores resultados, o fijar las portadoras de reducción de PAPR para que sean las mismas en cada transmisión. El problema de la búsqueda de las portadoras óptimas combinado con las otras opciones que nos aportará esta técnica, dará un espacio de búsqueda demasiado grande, como se verá a continuación, por lo que es habitual usar un set de portadoras CFR fijo. La elección de la amplitud de las portadoras depende de cómo se vayan a usar, donde existen tres alternativas: dejarlas a cero, darles una amplitud constante y unas fases concretas para reducir la PAPR o que esas portadoras envíen sendas señales de corrección de error ya sea empleando técnicas ya revisadas como la *Error Waveform Substracion*, *Peak Cancellation* u otros tipos. La primera de estas opciones es la menos ambiciosa, pero es eficaz y sobre todo computacionalmente asequible. La

segunda opción es muy costosa, pero ofrece muy buenos resultados, esto es debido a que el espacio de búsqueda contiene la doble tarea de encontrar las portadoras y fases óptimas, lo cual requiere muchas operaciones. Para acabar está la tercera opción, que ofrece también muy buenos resultados y además supone mucho menos coste computacional, es la combinación más beneficiosa. Como último e importante detalle, una de las grandes ventajas de esta técnica es su sinergia con la tecnología MIMO, siendo una de las técnicas que mejor aprovecha sus características, al poder separar flujos de datos y portadoras.

#### G. Constellation extension (CE)

El método *Constellation Extension* consiste en ceder un bit de la modulación de algunas de las portadoras para poder tener más libertad a la hora de mandar los símbolos, y así provocar que la información de cada uno de ellos ya no sea única, pudiendo transmitir la misma combinación de bits mediante dos símbolos diferentes. La forma de implementarla típicamente es haciendo que la información en cada símbolo sea igual a la del signo contrario o a la del complejo conjugado, de manera que a la hora de la transmisión se buscará la combinación óptima de símbolos que minimicen la PAPR. La pérdida de tasa será siempre de un bit (podría hacerse de más, pero no ofrece mejor rendimiento que usando uno) por lo que afecta de forma diferente a cada modulación, en el caso de una QPSK se perdería la mitad de la tasa, pero en una 64QAM solo un sexto. Afortunadamente, no es necesario aplicar CE a todas las portadoras, basta con la mitad de estas ya que a partir de ahí no mejoraría la PAPR más. Aun con la pérdida de tasa es una muy buena técnica ya que aparte de no empeorar ni el EVM ni el espectro es fácilmente combinable con otras tales como la reserva de tonos para mejores resultados en conjunto. Esta, como las otras técnicas probabilísticas tales como *Tone Reservation* o *PTS*, goza de la posibilidad de ser usada en combinación con algoritmos de búsqueda del óptimo y algoritmos de parada, e imponer un valor de PAPR deseado donde parar la búsqueda, de manera que se vuelva un problema de optimización más sencillo.

#### H. Coding

Las técnicas de codificación son muy numerosas, y existen varias formas de aplicarlas consiguiendo diferentes resultados y obteniendo rendimientos variados. En estas se pierde algo de tasa a cambio de no degradar la EVM, ya que se busca hacer uso de información redundante de manera que escogiéndola correctamente reduzca la PAPR. Algunos ejemplos de estos métodos son, por ejemplo, la codificación en bloque, que solo se puede usar con modulaciones de modulo constante [5], los códigos de Walsh para CDMA donde se aplican recursivamente hasta obtener PAPR de 0 dB (aunque luego el propio filtro de la modulación lo aumentará) y la codificación SC-FDMA (*Single Carrier FDMA*) donde se mapean las portadoras realizando una FFT antes de la IFFT propia de la señal OFDM.

### III. BALANCE DE RENDIMIENTO

Todas las técnicas presentadas en este artículo son perfectamente implementables y ofrecerán buenos resultados en cuanto a reducción de PAPR, pues ese es su objetivo, pero con diferentes resultados, en forma de rendimiento e inconvenientes. En general, el primer grupo de técnicas consistente en *Clipping*, *Peak Cancellation*, *Error Waveform Substraction* y *Peak Windowing* son opciones sencillas de implementar, especialmente el *Clipping*, y que de forma directa reducirán la PAPR, pero a costa de producir emisiones fuera de banda, por lo cual son buenas si no son importantes estas emisiones para el sistema o si se puede realizar un buen filtrado tras hacer uso de ellas. Mientras tanto, entre las técnicas probabilísticas brillan especialmente el PTS y SLM, las cuales son muy potentes y se ven beneficiadas por los avances en el campo de la optimización. También tiene un puesto importante entre las técnicas de CFR el *Tone Reservation*, que ofrece la posibilidad de combinar varios métodos de CFR de forma sencilla y sinérgica, siendo muy importante para la tecnología MIMO, que está muy presente en las comunicaciones actuales.

### IV. CONCLUSIONES

La técnica OFDM es actualmente la elección más popular en comunicaciones inalámbricas, aunque tiene un gran inconveniente, su alta PAPR, para lo cual son usadas las técnicas de CFR. En este artículo son revisadas las opciones más populares en cuanto a estos métodos, mostrando la amplia variedad que existe entre ellos y sus ventajas e inconvenientes. Con este artículo se busca que el lector

pueda seleccionar la mejor para cada caso, además de conocer sus bases y ciertos detalles de su implementación, ya que no todas encajan con cualquier sistema y ofrecen características variadas.

### REFERENCES

- [1] Y. A. Jawhar et al., "A Review of Partial Transmit Sequence for PAPR Reduction in the OFDM Systems," *IEEE Access*, vol. 7, pp. 18021–18041, 2019, doi: 10.1109/ACCESS.2019.2894527.
- [2] A. H. M. Rafiul Awal, Z. Dou, M. Al Shayokh, and M. I. Zahoor, "Implementation of chicken swarm optimization (CSO) with partial transmit sequences for the reduction of PAPR in OFDM system," 2017 9th IEEE Int. Conf. Commun. Softw. Networks, ICCSN 2017, vol. 2017-January, no. 2, pp. 468–472, 2017, doi: 10.1109/ICCSN.2017.8230156.
- [3] S. Verma, P. Sharma, S. Ahuja, and P. Hajela, "Partial transmit sequence with convolutional codes for reducing the PAPR of the OFDM signal," ICECT 2011 - 2011 3rd Int. Conf. Electron. Comput. Technol., vol. 3, pp. 70–73, 2011, doi: 10.1109/ICECTECH.2011.5941803.
- [4] H. Enzinger, K. Freiberger, and C. Vogel, "Competitive Linearity for Envelope Tracking," *IEEE Microw. Mag.*, vol. 19, no. 1, pp. 69–77, 2018, doi: 10.1109/MMM.2017.2759618.
- [5] T. Wada, T. Yamazoto, M. Katayama, and A. Ogawa, "A constant amplitude coding for orthogonal multi-code CDMA systems," *IEICE Trans. Fundamentals*, Vol. E80-A, No. 12, pp. 2477–2484, Dec. 1997.
- [6] L. Wang and C. Tellambura, "An overview of peak-to-average power ratio reduction techniques for OFDM systems," *Sixth IEEE Int. Symp. Signal Process. Inf. Technol. ISSPIT*, no. 2, pp. 840–845, 2006, doi: 10.1109/ISSPIT.2006.270915.
- [7] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, 2008, doi: 10.1109/TBC.2008.915770.
- [8] J. Wood, "Behavioral modeling and linearization of RF power amplifiers." Boston: Artech House, 2014. ISBN 1-60807-121-9.

# Apéndice D

## Artículo URSI DPD Dual con Ridge

### Arquitectura de aprendizaje directo en DPD de doble banda con regularización de Ridge

Miguel Nogales, Elías Marqués-Valderrama, Juan A. Becerra, María J. Madero-Ayora, Carlos Crespo-Cadenas  
mignoggon@alum.us.es, elimarval@alum.us.es, jabcerra@us.es, mjmadero@us.es, ccrespo@us.es  
Dpto. de Teoría de la Señal y Comunicaciones. Universidad de Sevilla. Camino de los descubrimientos s/n. 41092. Sevilla.

**Resumen**—In this paper, a direct learning architecture is employed for the digital predistortion of a power amplifier with a concurrent dual-band input signal. Conventional dual-band behavioral models need to cope with a large number of regressors, being prone to algebraic issues that arise in the computation of the inverse of the covariance matrix. Ridge regularization is applied to the modelling equation in order to improve the stability of the learning process and the precision of the estimated coefficients in validation. Experimental results for the linearization performance achieved with the proposed approach are provided for a class AB power amplifier driven by two concurrent 10-MHz 5G New Radio signals.

#### I. INTRODUCCIÓN

A medida que las generaciones de comunicaciones móviles han evolucionado se han establecido requerimientos de cada vez mayores tasas de datos. Con este fin, el uso de esquemas de modulación eficientes como OFDM permiten un aprovechamiento óptimo del espectro radioeléctrico. La señales basadas en este esquema de modulación se caracterizan por tener una alta relación potencia promedio a potencia de pico (PAPR), lo que obliga al amplificador de potencia a trabajar en régimen no lineal. Técnicas como la predistorsión digital (DPD) permiten mitigar la distorsión generada por la no linealidad, evitando así emisiones fuera de banda no deseadas [1].

Otra estrategia ampliamente utilizada en arquitectura de comunicaciones inalámbricas es la multiplexación en el dominio de la frecuencia. La combinación de esta técnica con el acceso múltiple por división en frecuencia FDMA permite una gran flexibilidad de asignación de recursos, como ocurre en la capa física de la quinta generación de comunicaciones móviles (5G-NR). El uso de DPD con este tipo de esquemas multibanda presenta un gran reto, ya que el número de coeficientes del modelo se dispara y aparecen problemas numéricos que traen inestabilidad a la regresión.

En el caso de modelos multibanda, los regresores de Volterra tienen la forma de productos de los equivalentes paso bajo de cada banda y sus envolventes con distintas combinaciones de retrasos [2], [3]. Dicha combinatoria es muy rica, y además de generar un número ingente de coeficientes crea un conjunto de regresores del modelo muy parecidos entre sí, lo que lleva a un mal condicionamiento de la matriz de medida en la regresión [4].

El enfoque principal a la hora de abordar una regresión con una matriz mal condicionada consiste en la regularización, que se basa en modificar ligeramente la función de coste que se pretende minimizar para que la dependencia de los coeficientes con las señales de entrada disminuya. La regularización de Ridge es una de las técnicas clásicas de este tipo [5], [6].

En este trabajo exploramos el uso de regularización de Ridge en un esquema de aprendizaje directo (DLA) multibanda. Para ello, en la Sección II se introducen los conceptos básicos de la DLA y cómo se extienden en un contexto multibanda. A continuación, en la Sección III se establece el marco de trabajo de regresión de Ridge. La Sección IV presenta resultados experimentales de predistorsión de un amplificador de potencia comercial. Finalmente, la Sección V cubre las conclusiones y finaliza el presente trabajo.

#### II. ARQUITECTURA DE APRENDIZAJE DIRECTO EN DOBLE BANDA

El esquema tradicional de DLA consta principalmente de un bucle cuya tarea es predistorsionar cierta señal de entrada  $\mathbf{u}$ ,

$$\mathbf{x} = \mathbf{u} - \mathbf{d}, \quad (1)$$

de modo que se le restará la señal llamada de distorsión,  $\mathbf{d}$ , y dará como resultado la envolvente compleja de la señal que se empleará como entrada al amplificador,  $\mathbf{x}$ . La señal  $\mathbf{d}$  se define como

$$\mathbf{d} = \mathbf{U}\mathbf{w}, \quad (2)$$

donde la información adquirida por el sistema será almacenada en el vector de pesos  $\mathbf{w}$  mientras que la matriz  $\mathbf{U}$  contendrá los regresores, que serán generados siguiendo la estructura del modelo adoptado, entre los que el modelo *generalized memory polynomial* (GMP) [7] es uno de los más ampliamente utilizados. En cuanto al cálculo de los coeficientes  $\mathbf{w}$  en cada iteración,

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu \mathbf{U}^\dagger \mathbf{e}, \quad (3)$$

este será realizado mediante la pseudoinversa de Moore-Penrose

$$\mathbf{U}^\dagger = (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H = \mathbf{R}^{-1} \mathbf{U}^H, \quad (4)$$

donde  $\mathbf{R}$  es la matriz de correlación y la actualización estará regulada por un coeficiente que determinará la velocidad del aprendizaje,  $\mu$ , cuyo valor pertenecerá al intervalo  $(0, 1)$ . El lazo se cierra obteniendo el error,

$$\mathbf{e} = \frac{\mathbf{y}}{g_0} - \mathbf{u}, \quad (5)$$

siendo y la envolvente compleja a la salida del amplificador de potencia en esa iteración, mientras que  $g_0$  es la ganancia estática, obtenida del cociente entre las normas de  $\mathbf{y}$  y  $\mathbf{u}$ .

Una vez definido el sistema DLA de una sola banda, el sistema dual surge del primero duplicando sus bloques de procesamiento de señal, dedicando uno a cada banda. Esta separación de los sistemas en bandas da lugar a varias implicaciones, como la necesidad de trabajar con el doble de variables

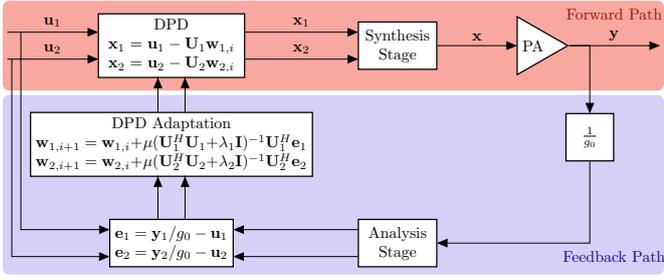


Fig. 1: Esquema de linealización con aprendizaje directo y regularización de Ridge para una señal de doble banda concurrente.

o generar los regresores con modelos de dos bandas, que de forma general contemplan términos mezclados y cruzados. Un ejemplo de este tipo de modelos es el GMP bidimensional (2D-GMP) [8]. Este cambio aportará importantes mejoras en rendimiento para el sistema debido a las interacciones entre las dos bandas.

Un ejemplo del esquema final se muestra en la Fig. 1, donde se puede apreciar la similitud con el sistema adaptado a una sola banda, pero con los añadidos propios del modelo dual. Así pues, siguiendo dicho esquema, se denota con el subíndice 1 a la banda con frecuencia central inferior y con el subíndice 2 a la banda con la frecuencia central superior.

### III. REGULARIZACIÓN DE RIDGE

En esta sección se introduce la técnica de regularización con carácter general, para ser aplicada posteriormente en la arquitectura DLA. Un problema de modelado tiene la siguiente forma:

$$\mathbf{y} = \mathbf{X} \mathbf{h} + \mathbf{e}, \quad (6)$$

donde  $\mathbf{X}$  es la matriz de regresores,  $\mathbf{h}$  los coeficientes del modelo y  $\mathbf{e}$  el error. El objetivo de este problema será encontrar los coeficientes que hagan que el error sea nulo. Dado que se trata de una estimación es aconsejable emplear la notación  $\hat{\mathbf{y}} = \mathbf{X} \hat{\mathbf{h}}$ , donde  $\hat{\mathbf{h}}$  denota estimación.

El gran número de regresores que se requieren para poder abordar un modelo de doble banda, puede provocar que el estimador de mínimos cuadrados (LS) no sea único. Además la existencia de regresores muy correlados conduce a que la matriz de correlación generalmente sea singular o cercana a serlo, es decir, presenta un número de condición muy elevado. Esto quiere decir que, en el caso en el que sea invertible, ligeros cambios en  $\mathbf{X}$  desembocarían en grandes cambios en  $(\mathbf{R})^{-1}$ . Esto se traduce en que puede obtenerse una buena estimación de los coeficientes del modelo en test pero malos para validación.

Así pues, la solución que se presenta es la regularización de Ridge, que si bien abandona la idea de conseguir un estimador insesgado, es capaz de llegar a un compromiso entre el sesgo y varianza en la estimación. A continuación se muestra la deducción de la expresión del nuevo modelo a partir de una penalización en la suma residual de cuadrados (PRSS),

$$\text{PRSS} = \|\mathbf{y} - \hat{\mathbf{y}}\|_2^2 + \lambda \|\hat{\mathbf{h}}\|_2^2, \quad (7)$$

donde  $\lambda$  es el coeficiente de regularización. Como el objetivo es encontrar el mínimo de la función PRSS, a continuación, se deriva respecto de  $\mathbf{h}$  y se iguala a 0,

$$\frac{d\text{PRSS}}{d\mathbf{h}} = -2\mathbf{X}^H \mathbf{y} + 2\mathbf{X}^H \mathbf{X} \hat{\mathbf{h}} + 2\lambda \hat{\mathbf{h}} = 0, \quad (8)$$

y finalmente se despeja el modelo regularizado  $\hat{\mathbf{h}}$ :

$$\hat{\mathbf{h}} = (\mathbf{X}^H \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^H \mathbf{y}, \quad (9)$$

donde  $\mathbf{I}$  es la matriz identidad.

De manera que, observando la ecuación (7), se extraen varios análisis. El primero de ellos es que la minimización ya no consiste en aumentar los coeficientes  $\hat{h}_j$ , ya que ahora aparece un término de penalización regulado por  $\lambda$  que aumenta con  $\|\hat{\mathbf{h}}\|_2$ . Esto puede interpretarse como una restricción en la norma  $\ell_2$  del vector de coeficientes. El segundo análisis se induce del primero, y es que el hecho de restringir  $\|\hat{\mathbf{h}}\|_2$  está sesgando la estimación, y esto asegura la invertibilidad de la matriz de regresores regularizada.

La implicación que esto tiene en el modelado se comprende sobre el espacio de componentes principales (PC) de  $\mathbf{X}$ . La regresión consiste en proyectar y en el subespacio que contiene a  $\mathbf{X}$ , cuya base en este desarrollo se considera el espacio de PC de  $\mathbf{X}$ . Dado que  $\hat{\mathbf{h}}$  representa las coordenadas de la proyección en dicha base, el hecho de introducir la misma penalización en todas las componentes de la diagonal de la matriz de correlación provoca una restricción mayor en aquellas componentes con menor variabilidad en  $\hat{\mathbf{y}}$ . De modo que además de conseguir mayor estabilidad y fiabilidad en la regresión sirve como método de simplificación en cuanto a su complejidad.

#### A. Aplicación de la regularización de Ridge en DLA

Generalmente la matriz de regresores se denota como  $\mathbf{X}$ , hecho que motiva la notación empleada en este apartado. En cambio la arquitectura DLA representada en la Fig. 1 entiende que la entrada al sistema es  $\mathbf{u}$  en lugar de  $\mathbf{x}$ . Esto es debido a que tradicionalmente se emplea  $\mathbf{x}$  para indicar la entrada en el amplificador. Para no llevar a confusión se ha decidido que  $\mathbf{u}$  represente la señal generada, cuya matriz de regresores es  $\mathbf{U}$ , de modo que  $\mathbf{x}$  es la señal predistorsionada según (1).

### IV. RESULTADOS EXPERIMENTALES

En esta sección se pondrá a prueba experimentalmente la solución teórica propuesta en las Secciones II y III. En primera instancia se hace una breve explicación del experimento realizado para proceder posteriormente a la explicación de los resultados obtenidos.

#### A. Descripción del banco de medidas

Para confirmar la validez de la propuesta presentada en este trabajo, se evaluaron sus prestaciones para linealización mediante un banco de medidas compuesto por un generador de señal SMU200A de Rohde & Schwarz, dos pre-amplificadores TVA-4W-422A+ de Minicircuits en cascada, un analizador de señal PXA-N9030A de Keysight Technologies y una fuente de alimentación. Como amplificador bajo prueba se empleó la placa de evaluación de un amplificador de clase AB basado en el HEMT de GaN CGH40010 de Cree, con una frecuencia de trabajo de 3.6 GHz.

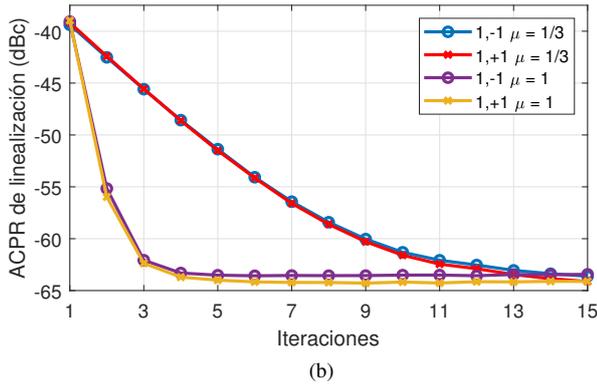
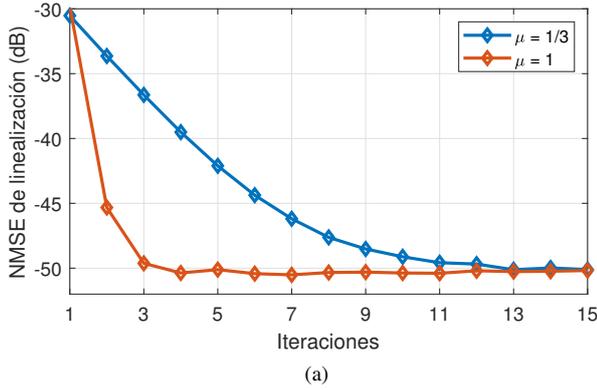


Fig. 2: Evolución del (a) NMSE de linealización y (b) ACPR de la banda 1 para valores de  $\mu = 1/3$  y 1 frente al número de iteraciones.

La señal de sondeo de doble banda estaba compuesta por dos señales OFDM de acuerdo con el estándar 5G-NR, cada una con un ancho de banda de 10 MHz y una separación entre subportadoras de 30 kHz, presentando una PAPR en torno a 11 dB. Antes de combinar las señales de banda base, generadas con una tasa de muestreo  $f_s = 92.16$  MHz, se aplicó un desplazamiento de frecuencia  $-f_s/4$  a la señal de la banda 1 y un desplazamiento  $+f_s/4$  a la señal de la banda 2. Estos valores se escogieron para optimizar el rango de frecuencias observables con el equipamiento disponible.

En el analizador, se adquirieron muestras de la envolvente compleja de la señal a la salida del amplificador bajo prueba considerando las portadoras de cada una de las bandas y una tasa de muestreo  $f_s/2$ , lo que supone un factor de sobremuestreo de 3 para cada banda. El ruido de medida se redujo promediando 10 realizaciones de la señal adquirida y posteriormente se sincronizaron las señales de entrada y salida. Cabe mencionar que, en la arquitectura de aprendizaje directo empleada para predistorsión digital, a las señales predistorsionadas se les aplicó una técnica de *clipping* que impedía que su PAPR superara los 12 dB como protección.

### B. Comportamiento del modelo

Se planteó una configuración del modelo 2D-GMP [8], con 534 regresores en cada banda, lo que supone un número demasiado elevado y da lugar a problemas numéricos, de manera que, sin la aplicación de la regularización de Ridge, el funcionamiento de la arquitectura DLA tiende a la diver-

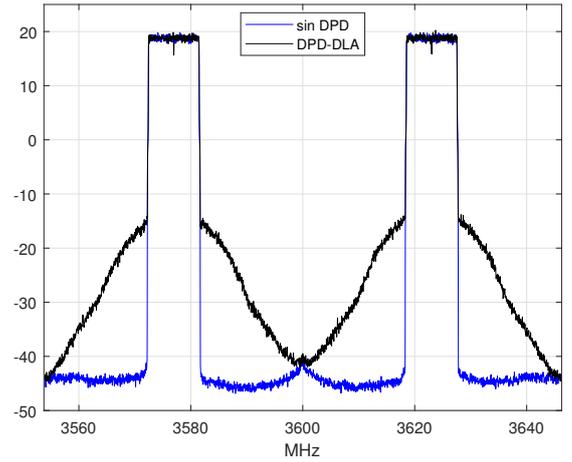


Fig. 3: Densidad espectral de potencia de la señal de salida del amplificador sin emplear DPD frente al uso de DPD.

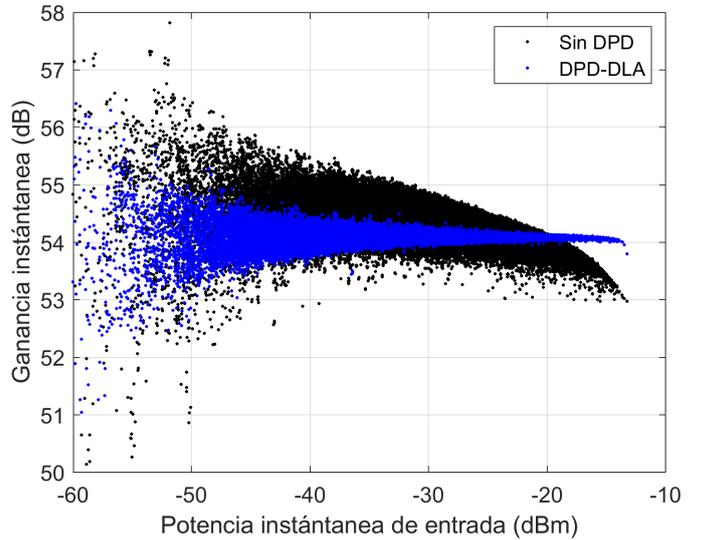


Fig. 4: AM-AM sin emplear DPD comparado con AM-AM al emplear DPD-DLA.

gencia. A este mismo modelo se le aplicó la regularización de Ridge, teniendo como resultado las prestaciones que se detallan a continuación.

En primer lugar, el error cuadrático medio normalizado (NMSE) frente al número de iteraciones DLA se muestra en la Fig. 2(a). Se observa un decrecimiento a medida que avanzan las iteraciones de aprendizaje, siendo más rápido a mayor valor del paso de aprendizaje  $\mu$ . Por otro lado, en la Fig. 2(b) se muestra el comportamiento de la relación de potencia de canal adyacente (ACPR), que experimenta una evolución muy similar a la del NMSE. Esto quiere decir que se ha conseguido mitigar el recrecimiento espectral, tal y como se muestra en Fig. 3.

Por otro lado cabe destacar que el hecho de que la regularización de Ridge establezca el comportamiento del bucle de control en la arquitectura DLA permite aumentar  $\mu$  a su valor máximo. Esto implica que es posible alcanzar el valor óptimo de los coeficientes  $w_i$  en un menor número de iteraciones, tal y como se observa en la Fig. 2.

TABLA I: Indicadores de Linealización para los Casos sin DPD y con DPD-DLA.

Caso	NMSE <sub>1</sub>	NMSE <sub>2</sub>	ACPR <sub>1,-1</sub>	ACPR <sub>1,+1</sub>	ACPR <sub>2,-1</sub>	ACPR <sub>2,+1</sub>	EVM <sub>1</sub>	EVM <sub>2</sub>	P <sub>out</sub>	PAE
	(dB)		(dBc)				(%)		(dBm)	(%)
Sin DPD	-30.5	-30.0	-39.4	-39.2	-38.7	-38.7	2.7	2.8	30.1	17.9
DPD-DLA	-49.7	-49.8	-62.8	-63.3	-62.8	-62.1	1.2	1.2	30.1	18.0

La linealización también se puede observar en la Fig. 4, la cual relaciona potencia instantánea de entrada con la ganancia instantánea, y donde el caso óptimo se encuentra al conseguir que la nube de puntos sea horizontal, lo cual se logra con este método.

Por último, el rendimiento de la arquitectura empleada se muestra en la Tabla I. Es importante destacar que la eficiencia de potencia añadida (PAE) no se ha visto alterada. Esto quiere decir que el empleo de estas técnicas no empeoran la eficiencia del amplificador. También resalta la reducción de la magnitud del vector de error (EVM), que implica una mejora en la recepción de la señal, siendo sinónimo de una menor dispersión en la constelación de la señal recibida.

## V. CONCLUSIONES

La aplicación de la arquitectura DLA junto con la regularización de Ridge ha resultado en una mejora en prestaciones de linealización, ya que la señal a la salida del amplificador cumple con el criterio de NMSE y las máscaras espectrales impuestas por 5G-NR sin degradar su eficiencia. Además, la combinación del 2D-GMP con la regularización de Ridge ha conseguido regresionar un modelo que a priori era inestable por la gran cantidad de regresores que contiene.

Por otro lado, en favor de la técnica de regularización de Ridge, cabe mencionar que la única modificación en el modelo ha sido la inserción de la constante en la diagonal de la matriz de correlación. Su implementación es muy sencilla y garantiza estabilidad en situaciones en las que el número de regresores es demasiado elevado.

## REFERENCIAS

- [1] A. Katz, J. Wood, and D. Chokola, "The evolution of PA linearization: From classic feedforward and feedback through analog and digital predistortion," *IEEE Microw. Mag.*, vol. 17, no. 2, pp. 32–40, Feb. 2016.
- [2] P. Roblin, S. K. Myoung, D. Chaillot, Y. G. Kim, A. Fathimulla, J. Strahler, and S. Bibyk, "Frequency-selective predistortion linearization of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 1, pp. 65–76, Jan. 2008.
- [3] S. A. Bassam, W. Chen, M. Helaoui, F. M. Ghannouchi, and Z. Feng, "Linearization of concurrent dual-band power amplifier based on 2D-DPD technique," *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 12, pp. 685–687, Dec. 2011.
- [4] P. L. Gilabert, R. N. Braithwaite, and G. Montoro, "Beyond the Moore-Penrose inverse: Strategies for the estimation of digital predistortion linearization parameters," *IEEE Microw. Mag.*, vol. 21, no. 12, pp. 34–46, Dec. 2020.
- [5] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.
- [6] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning – with Applications in R*, ser. Springer Texts in Statistics. New York: Springer, 2013, vol. 103.
- [7] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.
- [8] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, "2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 10, pp. 2547–2553, Oct. 2011.

# Índice de Figuras

---

2.1	Bloques principales del transmisor OFDM	4
2.2	Compresión de ganancia de un amplificador	5
2.3	Comparativa de los efectos de la saturación en tiempo y frecuencia y descomposición de la señal saturada en frecuencias puras por serie de Fourier	6
2.4	Característica AM-AM	7
2.5	Notación de los canales en doble banda concurrente	8
2.6	Modelado en el espacio de señal	10
3.1	Esquema aprendizaje indirecto	14
3.2	Esquema aprendizaje directo	15
3.3	Efectos de valores de $\mu$ extremos	18
3.4	Uso de <i>clipping</i> sobre señal de alta PAPR	20
3.5	Divergencia de los parámetros principales NMSE, EVM y ACPR	20
4.1	Esquema de la inicialización de los elementos del DPD en doble banda	22
4.2	Esquema de los elementos del bucle del DPD en doble banda	23
4.3	Matrices de correlación de los regresores de las bandas derecha e izquierda. Un valor más alto indica mayor correlación, alcanzando el máximo en 1	23
4.4	Tipos de regresores del modelo 2D-GMP de Erzinger, aplicados a la banda izquierda	25
4.5	Esquema final usado, modelo 2D-GMP	27
5.1	Diferentes tipos de funciones para el enfriamiento	30
5.2	Diagrama de flujo del algoritmo <i>orthogonal matching pursuit</i>	31
5.3	Diagrama de flujo del algoritmo <i>doubly orthogonal matching pursuit</i>	32
5.4	NMSE según las iteraciones para el OMP y DOMP, junto con el número óptimo de coeficientes	33
5.5	Gráfica de los valores de NMSE en validación e identificación para ciertos valores de $\lambda$	34
6.1	Montaje del laboratorio de la Universidad de Chalmers	36
6.2	Espectro de la señal dual con las iteraciones	37
6.3	Constelación de las portadoras de la señal OFDM según pasan las iteraciones	37
6.4	Montaje del laboratorio de Sistemas de Radiocomunicación de la Universidad de Sevilla	39
6.5	Muestra de la señal en el analizador de espectro	40
6.6	Efecto de la divergencia en el espectro	40
6.7	Comparación de la evolución de NMSE, ACPR y EVM con $\mu$ igual a 1/3 y 1/2	41



# Índice de Tablas

---

6.1	Indicadores de linealización para diferentes casos de DPD-DLA en Chalmers	38
6.2	Indicadores de linealización para diferentes casos de DPD-DLA en Sevilla	39



# Bibliografía

---

- [1] C. Reiners and H. Rohling, "Multicarrier transmission technique in cellular mobile communications systems," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*. IEEE, 1994.
- [2] A. C.A. and P. Prakasam, "Design of high speed FFT algorithm for OFDM technique," in *2016 Conference on Emerging Devices and Smart Systems (ICEDSS)*. IEEE, mar 2016.
- [3] D. V. Bhimsing and A. Bhagali, "Performance of channel estimation and equalization in OFDM system," in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*. IEEE, sep 2017.
- [4] T. Inoue, "5G NR Release 16 and millimeter wave integrated access and backhaul," in *2020 IEEE Radio and Wireless Symposium (RWS)*. IEEE, jan 2020.
- [5] S. Henry, A. Alsohaily, and E. S. Sousa, "5G is real: Evaluating the compliance of the 3GPP 5G New Radio system with the ITU IMT-2020 requirements," *IEEE Access*, vol. 8, pp. 42 828–42 840, 2020.
- [6] S. Manzoor, A. S. Bamuhaisoon, and A. N. Alifa, "Channel estimation for MIMO-OFDM systems," in *2015 5th National Symposium on Information Technology: Towards New Smart World (NSITNSW)*. IEEE, feb 2015.
- [7] P. Yuan, D. Xiao, J. Han, and X. Jing, "Neighbor carrier signal strength estimation for carrier aggregation in LTE-A," in *2010 WASE International Conference on Information Engineering*. IEEE, aug 2010.
- [8] L. Larson, D. Kimball, P. Asbeck, P. Draxler, J. Deng, and M. Li, "Digital predistortion techniques for linearized power amplifiers," in *2006 Asia-Pacific Microwave Conference*. IEEE, dec 2006.
- [9] R. H. Flake, "Volterra series representation of nonlinear systems," *Transactions of the American Institute of Electrical Engineers, Part II: Applications and Industry*, vol. 81, no. 6, pp. 330–335, 1963.
- [10] C. Crespo-Cadenas, M. J. Madero-Ayora, J. Reina-Tosina, and J. A. Becerra, "Formal deduction of a Volterra series model for complex-valued systems," *Signal Processing*, vol. 131, pp. 245–248, feb 2017.
- [11] C. Crespo-Cadenas, M. J. Madero-Ayora, and J. A. Becerra, "A bivariate Volterra series approach to modeling and linearization of power amplifiers," in *2021 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*. IEEE, jan 2021.
- [12] A. Katz, J. Wood, and D. Chokola, "The evolution of PA linearization: From classic feedforward and feedback through analog and digital predistortion," *IEEE Microwave Magazine*, vol. 17, no. 2, pp. 32–40, feb 2016.
- [13] A. Katz and D. Chokola, "The evolution of linearizers for high power amplifiers," in *2015 IEEE MTT-S International Microwave Symposium*. IEEE, may 2015.
- [14] C. Eun and E. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 223–227, 1997.

- [15] L. Ding, F. Mujica, and Z. Yang, "Digital predistortion using direct learning with reduced bandwidth feedback," in *2013 IEEE MTT-S International Microwave Symposium Digest (MTT)*. IEEE, jun 2013.
- [16] E. Abd-Elrady and B. Mulgrew, "Direct learning architectures for digital predistortion of nonlinear Volterra systems," in *Sensor Signal Processing for Defence (SSPD 2010)*. IET, 2010.
- [17] L. Schuartz, A. T. Hara, A. A. Mariano, B. Leite, and E. G. Lima, "Comparison between direct and indirect learnings for the digital pre-distortion of concurrent dual-band power amplifiers," in *Proceedings of the 32nd Symposium on Integrated Circuits and Systems Design - SBCCI '19*, A. for Computing Machinery, Ed. New York, New York: Association for Computing Machinery, 2019.
- [18] S. Wang, M. A. Hussein, O. Venard, and G. Baudoin, "Impact of the normalization gain of digital predistortion on linearization performance and power added efficiency of the linearized power amplifier," in *2017 47th European Microwave Conference (EuMC)*. IEEE, oct 2017.
- [19] L. Aladrén, P. García-Dúcar, P. L. Carro, and J. de Mingo, "Digital predistortion optimization using normalization gain adjustment in wideband systems," *European Microwave Week 2013 : Conference proceedings : 6-11 October 2013, Nuremberg, Germany*, 2013.
- [20] P. L. Gilabert, R. N. Braithwaite, and G. Montoro, "Beyond the Moore-Penrose inverse: Strategies for the estimation of digital predistortion linearization parameters," *IEEE Microwave Magazine*, vol. 21, no. 12, pp. 34–46, dec 2020.
- [21] W. Polzleitner and H. Wechsler, "Simulating preattentive and attentive vision with moore-penrose associative memories," in *Conference Proceedings 1991 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 1991.
- [22] P. Horata, S. Chiewchanwattana, and K. Sunat, "A comparative study of pseudo-inverse computing for the extreme learning machine classifier," in *The 3rd International Conference on Data Mining and Intelligent Information Technology Applications*, 2011, pp. 40–45.
- [23] H. Enzinger, K. Freiberger, and C. Vogel, "Competitive linearity for envelope tracking: Dual-band crest factor reduction and 2D-vector-switched digital predistortion," *IEEE Microwave Magazine*, vol. 19, no. 1, pp. 69–77, jan 2018.
- [24] T. Zhang, C. L. P. Chen, C.-H. Wang, and S. C. Tam, "A new dynamic optimal learning rate for a two-layer neural network," in *2012 International Conference on System Science and Engineering (ICSSE)*. IEEE, jun 2012.
- [25] W. Qi, J. Huang, and X. Wu, "A 2D rapidly-convergent iterative digital predistortion method using direct learning structure," in *2020 IEEE MTT-S International Wireless Symposium (IWS)*. IEEE, sep 2020.
- [26] S. A. Sandeep Kumar Vangala, "Analysis and comparison of clipping techniques for OFDM/OQAM peak-to-average power ratio reduction," in *2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*. IEEE, oct 2015.
- [27] R. N. Braithwaite, "Peak limited digital predistortion of a RF power amplifier using a closed loop estimator," in *2021 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*. IEEE, jan 2021.
- [28] ———, "A combined approach to digital predistortion and crest factor reduction for the linearization of an RF power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 1, pp. 291–302, jan 2013.
- [29] A. Islam, P. J. Xia, H. Huang, and S. Boumaiza, "Constrained identification of rational functions for robust digital predistortion," in *2017 IEEE Radio and Wireless Symposium (RWS)*. IEEE, jan 2017.
- [30] A. Khlass, S. E. Elayoubi, and T. Bonald, "Multi-flow transmission and carrier aggregation inter-operation in HSPA+ advanced," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*. IEEE, sep 2014.

- [31] C. Quindroit, N. Narahariseti, P. Roblin, S. Gheitanchi, V. Mauer, and M. Fitton, "FPGA implementation of orthogonal 2D digital predistortion system for concurrent dual-band power amplifiers based on time-division multiplexing," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 12, pp. 4591–4599, dec 2013.
- [32] E. M. Mirkes, J. Allohibi, and A. N. Gorban, "Do fractional norms and quasinorms help to overcome the curse of dimensionality?" in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, jul 2019.
- [33] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 159–165, jan 2004.
- [34] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, oct 2006.
- [35] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, "2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 10, pp. 2547–2553, oct 2011.
- [36] C. Tarver, M. Abdelaziz, L. Anttila, M. Valkama, and J. R. Cavallaro, "Low-complexity, sub-band DPD with sequential learning: Novel algorithms and WARPLab implementation," in *2016 IEEE International Workshop on Signal Processing Systems (SiPS)*. IEEE, oct 2016.
- [37] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, dec 2007.
- [38] J. A. Becerra, M. J. Madero-Ayora, J. Reina-Tosina, C. Crespo-Cadenas, J. Garcia-Frias, and G. Arce, "A doubly orthogonal matching pursuit algorithm for sparse predistortion of power amplifiers," *IEEE Microwave and Wireless Components Letters*, vol. 28, no. 8, pp. 726–728, aug 2018.
- [39] ———, "A reduced-complexity doubly orthogonal matching pursuit algorithm for power amplifier sparse behavioral modeling," in *2019 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*. IEEE, jan 2019.
- [40] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, mar 1978.
- [41] J. A. Becerra, M. J. Madero-Ayora, R. G. Noguera, and C. Crespo-Cadenas, "On the optimum number of coefficients of sparse digital predistorters: A bayesian approach," *IEEE Microwave and Wireless Components Letters*, vol. 30, no. 12, pp. 1117–1120, dec 2020.
- [42] M. Nogales, E. Marqués-Valderrama, J. A. Becerra, M. J. Madero-Ayora, and C. Crespo-cadenas, "Arquitectura de aprendizaje directo en DPD de doble banda con regularización de Ridge," in *Aceptado en XXXVI Simposium Nacional de la Unión Científica Internacional de Radio (URSI 2021), 2-4 Septiembre 2021.*, 2021.
- [43] Y. Liu, X. Xia, Q. Xu, W. Pan, W. Ma, S. Shao, and Y. Tang, "Relaxing requirements on training samples in digital predistortion by using Ridge regression," *IEEE Microwave and Wireless Components Letters*, vol. 31, no. 6, pp. 616–619, jun 2021.



# Glosario

---

- 2D-GMP** *Two Dimensional Generalized Memory Polynomial*. 24
- 2D-MP** *Two Dimensional Memory Polynomial*. 24
- 2D-DPD** *Two Dimensional Digital Predistortion*, Predistorsión Digital Bidimensional. 21
- 3GPP** *3rd Generation Partnership Project*. 4
- 5G-NR** *5G New Radio*. 4
- ACPR** *Adjacent Channel Power Ratio*, Ratio de Potencia del Canal Adyacente. 1
- BIC** *Bayesian Information Criterion*, Criterio de Información Bayesiano. 30
- CA** *Carrier Aggregation*, Agregación de Portadoras. 2
- CFR** *Crest Factor Reduction*, Reducción del Factor de Cresta. 1
- CP** *Cyclic Prefix*, Prefijo Cíclico. 3
- DOMP** *Doubly Orthogonal Matching Pursuit*. 30
- DPD** *Digital Predistortion*, Predistorsión Digital. 1
- DUT** *Devide Under Test*, Dispositivo Bajo Prueba. 38
- EVM** *Error Vector Magnitude*, Magnitud del Vector de Error. 1
- FFT** *Fast Fourier Transform*, Transformada Rápida de Fourier. 3
- FPGA** *Field-Programmable Gate Array*, Matriz de Puertas Lógicas Programable en Campo. 22
- GMP** *Generalized Memory Polynomial*. 18
- HSPA+** *High Speed Packet Access +*, Acceso de Paquetes de Alta Velocidad +. 21
- IFFT** *Inverse Fast Fourier Transform*, Transformada Rápida de Fourier Inversa. 3
- IP** *Internet Protocol*, Protocolo de Internet. 3
- MIMO** *Multiple Inputs Multiple Outputs*, Múltiples Entradas Múltiples Salidas. 4
- MP** *Memory Polynomial*, Memory Polynomial. 24
- NMSE** *Normalized Mean Squared Error*, Error Cuadrático Medio Normalizado. 1
- OMP** *Orthogonal Matching Pursuit*. 30

**PA** *Power Amplifier*, Amplificador de Potencia. 1

**PAPR** *Peak to Average Power Ratio*, Ratio de Potencia Media a Pico. 1

**RC-DOMP** *Reduced Complexity Doubly Orthogonal Matching Pursuit*. 31

**SSPI-DOMP** *Simplified Sparse Parameters Identification Doubly Orthogonal Matching Pursuit*. 31

**URSI** Simposium Nacional de la Unión Científica Internacional de Radio. 34