

UNIVERSIDAD DE LA CORUÑA
Departamento de Electrónica y Sistemas



TESIS DOCTORAL

**ALGORITMOS ADAPTATIVOS
CIEGOS
PARA IGUALACIÓN Y SEPARACIÓN
DE SEÑALES**

Autor: CARLOS J. ESCUDERO CASCÓN

Director: LUIS CASTEDO RIBAS

Marzo, 1998

Dr. Luis Castedo Ribas
Profesor Titular de Universidad
Departamento de Electrónica y Sistemas
Universidad de La Coruña.

CERTIFICA:

Que la memoria titulada “*Algoritmos Adaptativos Ciegos para Igualación y Separación de Señales*”, ha sido realizada por D. Carlos J. Escudero Cascón bajo mi dirección en el Departamento de Electrónica y Sistemas de la Universidad de La Coruña y concluye la Tesis que presenta para optar al grado de Doctor.

La Coruña, 8 de Enero de 1998

Fdo: Dr. Luis Castedo Ribas
Director de la Tesis Doctoral

Tesis Doctoral: ALGORITMOS ADAPTATIVOS CIEGOS PARA
IGUALACIÓN Y SEPARACIÓN DE SEÑALES

Autor: D. Carlos J. Escudero Cascón

Director: D. Luis Castedo Ribas

Fecha: 2 de Marzo de 1998

Tribunal

Presidente:

Vocal 1:

Vocal 2:

Vocal 3:

Secretario:

Dedicado a mi familia.

Agradecimientos

Quiero agradecer a mi compañero y amigo el Profesor Luis Castedo el haber sido mi director de Tesis. Gracias a su ayuda he podido realizar este trabajo. Además, no quiero olvidarme de la Profesora Urbashi Mitra que aceptándome en el Departamento de Electrical Engineering de la Universidad de Ohio State contribuyó a acelerar el término de este trabajo.

Quiero recordar también en esta sección de agradecimientos a mis compañeros, y amigos, del grupo y del departamento al que pertenezco, sin olvidar a los de otros departamentos con los que tanto he investigado en nuestro despacho compartido de la cafetería.

No me olvido de mis amigos con los que he pasado, y me quedan por pasar, muchas “curas de estrés” de fin de semana. Entre ellos quiero destacar a Li-Chung que contribuyó a hacer de mi estancia en Estados Unidos una experiencia inolvidable.

Finalmente, quiero mostrar un especial agradecimiento a mi familia a quien dedico esta Tesis. Mis padres, hermano y esposa que siempre han sabido aguantar mis momentos de mal humor.

Resumen

Esta tesis aborda dos problemas fundamentales en la recepción de señales moduladas digitalmente: la igualación de canales y la separación de señales. Se proponen nuevas soluciones basadas en el conocido criterio de optimización de módulo constante (CM, Constant Modulus).

El objetivo en igualación de canales es eliminar la interferencia intersimbólica (ISI, Intersymbol Interference) causada por los canales dispersivos. En numerosas situaciones prácticas, esto se lleva a cabo mediante un filtro transversal cuya salida es una combinación lineal de versiones retardadas de la señal recibida. Debido a que las características del canal no son conocidas *a priori*, los coeficientes del filtro han de ser ajustados adaptativamente mediante un algoritmo recursivo. Uno de los algoritmos más utilizados es el CMA (Constant Modulus Algorithm) debido a su sencillez y a no requerir la transmisión de secuencias de entrenamiento. No obstante, estudios recientes han mostrado que este algoritmo puede converger hacia puntos no deseados cuando el igualador tiene una longitud finita y el canal se modela como un sistema autorregresivo (AR, Auto-Regressive).

Este problema es abordado en la presente tesis presentándose una modificación a la función de coste minimizada por el CMA que no posee mínimos locales para canales AR. Se muestra que en los mínimos locales existe una correlación entre la salida y la entrada desplazada, la cual no existe en los mínimos deseados. Esto motiva la introducción de un nuevo criterio en el que conjuntamente se minimiza la función de coste CM y la correlación entre la salida y versiones retardadas de la entrada. Aunque un análisis general del nuevo criterio es extremadamente complejo, se presenta un análisis de los puntos estacionarios de la función de coste resultante para un igualador de dos coeficientes y un canal AR de primer orden, demostrándose que en este caso no existen mínimos no deseados. Asimismo, también se presenta un análisis en presencia de ruido en el que se demuestra que el error cuadrático medio (MSE, Mean Square Error) alcanzado en los mínimos no deseados coincide con el mínimo posible (MMSE, Minimum Mean Square Error).

El otro problema que se aborda es la separación de señales mezcladas linealmente. Este problema surge con frecuencia en los receptores de sistemas de comunicaciones multiusuario en los cuales existe un *front-end* de comunicaciones (array de antenas, banco de filtros adaptados, ...) que suministran una mezcla lineal de las señales que se desean demodular. La separación se lleva a cabo con un sistema lineal de múltiples salidas. De nuevo nos encontramos con que la mezcla es desconocida y que los parámetros del sistema de separación deben ajustarse de forma adaptativa. La utilización del CMA en esta aplicación resulta enormemente atractiva porque es capaz de extraer en cada salida una única señal del entorno. Sin embargo, este criterio no puede aplicarse por sí solo ya que corre el peligro de extraer la misma señal en distintas salidas.

Para evitar este problema de captura, se presenta un nuevo criterio que considera la adaptación conjunta del criterio CM y la correlación cruzada entre salidas distintas. Estas correlaciones serán distintas de cero cuando se produzca el problema de captura y su minimización penaliza la convergencia hacia los mínimos donde esto ocurre. Un análisis general del nuevo criterio propuesto resulta extremadamente complejo. No obstante, se presenta un análisis particular para dos usuarios en el que se demuestra la eliminación de los mínimos no deseados. Asimismo, también se presenta un análisis en presencia de ruido en el que se demuestra que el MSE alcanzado en los mínimos es igual al mínimo MSE alcanzable (MMSE).

Abstract

This thesis deals with two fundamental problems in the reception of digital signals: channel equalization and signal separation. New solutions are proposed using the well-know Constant Modulus (CM) optimization criterion.

The objective of channel equalization is to remove the Intersymbol Interference (ISI) introduced by dispersive channels. In many practical situations, the equalizer is a transversal filter whose output is a linear combination of delayed versions of the received signal. Since the channel features are unknown in advanced, the filter coefficients must be adjusted in an adaptive way by means of a recursive algorithm. The CMA (Constant Modulus Algorithm) is one of the most widely used because it is simple, robust and does not require training sequences. However, recent stability analysis have demonstrated that this algorithm may converge to undesired points when considering finite length equalizers and the channel is modeled as an auto-regressive (AR) system.

This thesis proposes a solution to this problem by introducing a modification of the CMA cost function that does not present local minima. It is demonstrated that in the local minima there is a correlation between the output and the delayed input. This correlation vanishes at the desired ones. Though it is very difficult to obtain a general analysis of the proposed cost function, it is presented an analysis of the stationary points for a 2-tap equalizer and a first order AR channel that shows the non existence of undesired local minima. Moreover, it is also performed an analysis in presence of noise demonstrating that the Mean Square Error (MSE) at the minima of the new cost function is the minimum achievable (MMSE, Minimum Mean Square Error).

Another problem considered in this thesis is the separation of linear mixtures of signals. This problem frequently arises in multiuser communications receivers where there exist a front-end (antenna array, match bank filters, ...) that provides a linear mixture of the desired signals. The separation is made by means of a multiple outputs linear system. Again, the mixture is unknown and the parameters of the separation system must be adjusted in an adaptive way. The CMA is very attractive because it always extracts a single signal from the environment. However, this criterion cannot be used by itself because it may extract the same signal at different outputs.

In order to avoid this capture problem, a new criterion is proposed that considers the joint minimization of the CM criterion and the cross-correlation between different outputs. These correlations are non-zero when the capture problem appears and their minimization penalizes convergence to these kind of solutions. It is very difficult to perform a general analysis of the new criterion. However, an analysis is presented for two users that shows how the non-desired minima disappear. Moreover, it is presented an analysis in presence of noise where it is shown as the MSE reached at the minima of the new cost function corresponds to the MMSE value.

Índice General

1	Introducción	1
1.1	Motivación	1
1.2	Técnicas de Diversidad	3
1.2.1	Diversidad Temporal	3
1.2.2	Diversidad Frecuencial	4
1.2.3	Diversidad Espacial	4
1.2.4	Diversidad por Espectro Ensanchado	5
1.2.5	Otras Formas de Diversidad	6
1.3	Modelo de Señal para Canales no Dispersivos.	7
1.3.1	Diversidad Temporal y Frecuencial	9
1.3.2	Diversidad Espacial	11
1.3.3	Diversidad por Código	12
1.3.4	Discusión	13
1.4	Modelo de Señal para Canales Dispersivos	14
1.4.1	Línea de Retardos	15
1.4.2	Array de Antenas	16
1.4.3	Banco de Filtros Adaptados	17
1.4.4	Discusión	19
1.5	Arquitecturas Receptoras	19
1.5.1	Receptores MISO para Sistemas con Diversidad Temporal	20
1.5.2	Receptores MISO para Sistemas con Diversidad Espacial o por Código	21
1.5.3	Receptores MISO para Sistemas con Doble Diversidad	22

1.5.4	Receptores MISO para Sistemas con Triple Diversidad	25
1.5.5	Receptores MIMO	25
1.6	Planteamiento y Objetivos de la Tesis	25
2	Igualación de Canales AR	29
2.1	Introducción	29
2.2	Modelo de Señal	33
2.3	Igualadores Zero-Forcing	35
2.4	Igualadores MMSE	36
2.5	Igualadores Ciegos	37
2.5.1	Mínimos Locales Dependientes de la Longitud	42
2.6	Modificación de los Igualadores CMA	42
2.6.1	Algoritmo Adaptativo	44
2.6.2	Generalización a Canales AR de Mayor Orden	45
2.7	Análisis de los Puntos Estacionarios	47
2.7.1	Análisis sin Ruido	48
2.7.2	Análisis con Ruido	54
2.7.3	Valor del MSE en los Mínimos con Ruido	57
2.8	Simulaciones	59
2.9	Conclusiones	64
3	Separación Ciega de Señales	67
3.1	Introducción	67
3.2	Modelo de Señal	69
3.3	Receptor de Decorrelación	70
3.4	Receptor MMSE	71
3.5	Receptores Ciegos	72
3.5.1	Receptores LCMV	72
3.5.2	Receptores Basados en Técnicas de Separación Ciega de Fuentes	73
3.5.3	Receptores CMA	74
3.6	Modificación Receptores CMA	75

3.6.1	Algoritmo Adaptativo	76
3.7	Análisis de los Puntos Estacionarios	77
3.7.1	Análisis sin Ruido	77
3.7.2	Análisis con Ruido	82
3.7.3	Valor del MSE en los Mínimos con Ruido	85
3.8	Simulaciones	87
3.9	Conclusiones	95
4	Conclusiones y Líneas Futuras de Trabajo	97
4.1	Conclusiones	97
4.1.1	Igualación de Canales	98
4.1.2	Separación de Señales	99
4.2	Líneas Futuras de Trabajo	100
A	Diferenciación con Respecto a Vectores y Matrices	103
A.1	Diferenciación con Respecto a un Escalar	103
A.2	Diferenciación con Respecto a un Vector	104
A.2.1	Ejemplos	104
A.3	Diferenciación con respecto a una matriz	105
A.3.1	Ejemplos	105
B	Puntos Estacionarios de un Igualador Infinito CMA	107
C	Desarrollos Correspondientes al Capítulo 2	111
C.1	Desarrollo de la Ecuación (2.47)	111
C.2	Desarrollo de la ecuación (2.77)	113
C.3	Desarrollo del sistema de ecuaciones (2.81)	114
D	Desarrollos Correspondientes al Capítulo 3	117
D.1	Naturaleza de los Puntos Estacionarios de la Función de coste	117
D.1.1	Puntos Estacionarios	119
D.2	Cálculo del MSE	125

D.2.1	Desarrollo de la Ecuación (3.53)	125
D.2.2	Cálculo del Gradiente de (3.53)	126
D.2.3	Desarrollo de la Ecuación (3.55)	127
D.2.4	Cálculo del Gradiente de (3.55)	129
D.2.5	Obtención de la Expresión (3.57)	130
E	Notación y Abreviaturas Empleadas	133
E.1	Notación	133
E.2	Abreviaturas	134
	Bibliografía	135

Índice de Figuras

1.1	Diversidad temporal.	3
1.2	Diversidad frecuencial.	4
1.3	Diversidad espacial. Array de antenas.	5
1.4	Diversidad por código.	6
1.5	Diagrama de bloques un sistema de comunicaciones monousuario con diversidad y canales no dispersivos.	8
1.6	Diagrama de bloques un sistema de comunicaciones multiusuario con diversidad.	10
1.7	Diversidad temporal mediante una línea de retardos.	15
1.8	Arquitectura de receptor MISO para sistemas con diversidad temporal.	21
1.9	Arquitectura de receptor MISO para sistemas con diversidad espacial o por código.	22
1.10	Arquitectura de receptores MISO para sistemas con doble diversidad: código o espacial y temporal.	23
1.11	Arquitectura de receptores MISO para sistemas con doble diversidad: espacial y por código.	24
1.12	Arquitectura de receptores MISO para sistemas con triple diversidad.	26
1.13	Arquitectura de receptores MIMO.	27
2.1	Diagrama de bloques de un receptor monousuario con igualador lineal.	30
2.2	Diagrama de bloques de un receptor monousuario con igualador con realimentación de decisiones (DFE)	31
2.3	Respuesta al impulso de un canal de microondas.	32
2.4	Diagrama de bloques de un sistema de igualación de canal.	33
2.5	Representación de la función (2.69).	52

2.6	Representación de las funciones de coste correspondientes al nuevo algoritmo y al CMA para un igualador de dos coeficientes y un canal AR(1) con $a = 0.5$	53
2.7	Representación del contorno de la nueva función de coste y de la de CMA para un igualador de dos coeficientes y un canal AR(1) con $a = 0.5$ y ruido aditivo con una $SNR = 10$ dB.	58
2.8	Curvas de comparación del MSE respecto a la SNR en recepción del nuevo algoritmo, MMSE y de la aproximación obtenida en la sección 2.7.3.	60
2.9	Curvas de comparación de las probabilidades de error respecto a la SNR en recepción del nuevo algoritmo y del MMSE	61
2.10	Evolución temporal de la ISI cuando se considera un igualador de dos coeficientes y un canal AR(1) con función de transferencia $H(z) = \frac{1}{1-0.5z^{-1}}$. ($\mu = 0.002$).	62
2.11	Evolución temporal de la ISI cuando se considera un igualador de cinco coeficientes y un canal AR(1) con función de transferencia $H(z) = \frac{1}{1-0.5z^{-1}}$. ($\mu = 0.002$).	62
2.12	Evolución temporal de la ISI cuando se considera un igualador de tres coeficientes y un canal AR(2) con función de transferencia $H(z) = \frac{1}{(1-az^{-1})(1-bz^{-1})}$. ($\mu = 0.004$).	63
2.13	Evolución temporal de la ISI cuando se considera un igualador de dos coeficientes, un canal AR(1) con función de transferencia $H(z) = \frac{1}{1-0.5z^{-1}}$ y ruido aditivo con $SNR = 10$ dB. ($\mu = 0.002$).	64
3.1	Evolución temporal de la MAI para cada una de las salidas del sistema con una antena en array de 10 sensores. La primera curva corresponde con la extracción de un usuario de potencia 10 dB superior a la del usuario de la curva inferior. ($\alpha = 1$ y $\mu = 2 \times 10^{-4}$).	89
3.2	Evolución temporal de la SINR para cada una de las salidas del sistema con una antena en array de 10 sensores. SNR primer usuario 6 dB, SNR segundo usuario 3 dB. ($\alpha = 1$ y $\mu = 5 \times 10^{-4}$).	90
3.3	Evolución temporal de la MAI para cada uno de los usuarios de un sistema CDMA. La curva superior corresponde con la MAI de un usuario con una potencia de recepción 5 dB superior a la del usuario de la curva inferior. ($\alpha = 1$ y $\mu = 5 \times 10^{-5}$).	91

3.4	Error cuadrático medio (MSE) frente a la SNR del nuevo criterio CM y del MMSE obtenido en la recepción de uno de los dos usuarios que transmiten con una modulación DS/BPSK y un código <i>kasami</i> de 15 chips de longitud.	92
3.5	Probabilidad de error frente a la SNR alcanzada por el nuevo criterio CM y el MMSE en la recepción de uno de los dos usuarios que transmiten con una modulación DS/BPSK y códigos <i>kasami</i> de 15 chips de longitud.	93
3.6	Evolución temporal de la SINR para el nuevo criterio y el MMSE para uno de los dos usuarios que transmiten con una modulación DS/BPSK y códigos <i>kasami</i> de 15 chips de longitud.	94

Capítulo 1

Introducción

1.1 Motivación

El comportamiento de todos los sistemas de comunicaciones está limitado por la distorsión que los canales introducen sobre la señal transmitida. Esta situación se ve agravada en la realidad por el hecho de que las características de los canales son desconocidas y cambian aleatoriamente con el tiempo debido a las constantes variaciones de las propiedades físicas de los medios de transmisión.

Una de las principales fuentes de distorsión es el ruido que se superpone a las señales recibidas. Este ruido aditivo se modela habitualmente como un proceso aleatorio gaussiano ya que consiste en la suma de pequeñas interferencias, ruidos térmicos y otro tipo de contribuciones tanto naturales como generadas por el hombre. También es práctica común considerar que el ruido está incorrelado en el tiempo denominándose, entonces, ruido blanco. Los canales que introducen ruido blanco gaussiano se conocen en la bibliografía con el nombre de AWGN (Additive White Gaussian Noise).

La presencia de ruido es particularmente dañina en las comunicaciones por radio afectadas de propagación multitrayecto. Cuando la señal transmitida llega al receptor a través de diferentes caminos, su suma es en ocasiones constructiva, reforzándose la señal recibida, y en otras destructiva, produciéndose una atenuación de la señal recibida. Este fenómeno se conoce con el nombre de desvanecimiento (*fading*) y, cuando ocurre, el ruido puede llegar a enmascarar la señal transmitida.

Si las diferencias entre los retardos de propagación de los multitrayectos son despreciables frente a la duración de un símbolo, el desvanecimiento sólo causa una modificación de la amplitud pero no distorsiona su forma. Es lo que se llama desvanecimiento plano (*flat fading*). Por el contrario, si las diferencias entre los retardos son comparables a la duración del símbolo, el canal modifica la forma de los pulsos correspondientes a los símbolos transmitidos diciéndose que el desvanecimiento es selectivo

en frecuencia. La dispersión introducida por el canal hace que cuando se transmita una secuencia de pulsos concatenados, éstos se solapen entre sí apareciendo el fenómeno denominado interferencia intersimbólica (ISI, Intersymbol Interference). Además de los canales con propagación multitrayecto mencionados, existen otros muchos canales de comunicaciones que también presentan efectos dispersivos. Un ejemplo es el canal telefónico donde las limitaciones prácticas de los codificadores y las líneas de transmisión introducen una distorsión lineal sobre las señales transmitidas.

Los desvanecimientos e ISI descritos anteriormente son las principales causas de degradación en sistemas de comunicaciones monousuario en los cuales el canal sólo transporta la señal de un único usuario. En la práctica, sin embargo, es cada vez más frecuente que un canal sea compartido por varios usuarios en cuyo caso surge una nueva forma de degradación que se denomina interferencia de acceso múltiple (MAI, Multiple Access Interference). Esta interferencia se produce debido a imperfecciones en el proceso de multiplexación y es especialmente grave cuando la amplitud de la señal de interés es menor que la de los usuarios interferentes. Esto ocurre habitualmente en comunicaciones móviles donde los usuarios interferentes pueden estar mucho más cerca del receptor que el usuario deseado provocando el efecto denominado de *near-far* [39].

Para que los receptores puedan combatir los efectos nocivos de los canales hay que dotarlos de dispositivos y técnicas adecuadas. Así, la ISI se combate con filtros que compensen (igualen) la respuesta en frecuencia del canal y que eliminen la distorsión lineal. El desvanecimiento plano, por su parte, se reduce empleando técnicas de diversidad en las que se persigue obtener en el receptor diferentes réplicas de la misma señal transmitida. Cuando las réplicas son independientes entre sí, la probabilidad de tener fuertes desvanecimientos en todas ellas se reduce considerablemente.

En el caso de comunicaciones multiusuario, la utilización de técnicas de diversidad es clave para poder separar los usuarios en recepción. En muchas situaciones prácticas el proceso de multiplexación no es perfecto y las réplicas obtenidas resultan ser una mezcla de las señales transmitidas. Estas mezclas deben procesarse mediante filtros que permitan recuperar las señales originales.

Todos los mecanismos degradadores que acabamos de mencionar tienen una naturaleza aleatoria de forma que, aunque podemos encontrar un modelo bastante preciso para describirlos, sus parámetros son desconocidos y cambian con el tiempo. Esto obliga a utilizar en los receptores técnicas adaptativas que sean capaces de aprender las características del canal y seguir sus variaciones temporales.

En la presente tesis vamos a abordar dos problemas fundamentales en comunicaciones digitales. Uno es la eliminación (o reducción) de la ISI mediante el uso de igualadores y otro es la cancelación de la MAI en sistemas de comunicación multiusuario. Dado que estos últimos hacen uso de técnicas de diversidad, hemos estimado conveniente dedicar la siguiente sección a hacer una breve revisión de ellas.

1.2 Técnicas de Diversidad

El objetivo fundamental de las técnicas de diversidad es la obtención de diferentes réplicas, o versiones, correspondientes a una misma señal transmitida a través de canales independientes. Como ya se mencionó anteriormente, las técnicas de diversidad pueden ser utilizadas para combatir el desvanecimiento o bien para transmitir diferentes usuarios por un mismo canal.

Muchas son las técnicas de diversidad que han sido propuestas de entre las cuales las más importantes son las siguientes:

1.2.1 Diversidad Temporal

Las diferentes versiones de la señal transmitida se obtienen observando la recibida en distintos instantes de tiempo. Cuando se desea combatir el desvanecimiento se debe transmitir la misma señal por diferentes ranuras temporales de forma que todas ellas alcancen el receptor con características de perturbación distintas.

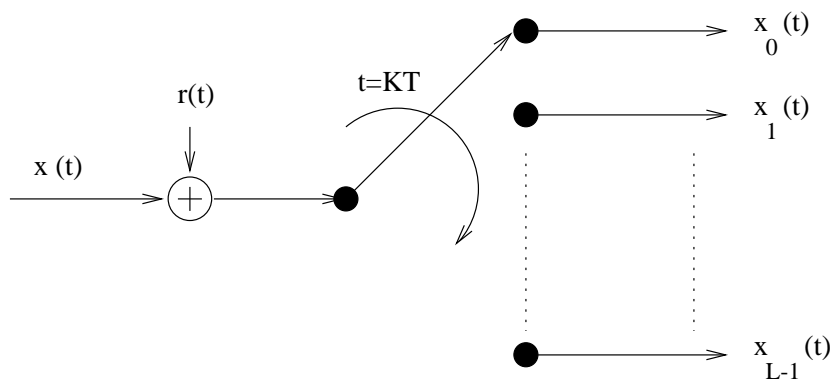


Figura 1.1: Diversidad temporal.

Por otra parte, la técnica de acceso multiusuario basada en este tipo de diversidad se denomina TDMA (Time Division Multiple Access). En esta técnica las señales transmitidas en las diferentes ranuras temporales se corresponden con las de cada uno de los usuarios del sistema. La figura 1.1 representa la forma de separar los usuarios a partir de la señal recibida. Cada una de las versiones es obtenida cada KT unidades de tiempo donde K es el número de símbolos que contiene una ranura temporal, T es el período de símbolo y L es el número de usuarios. El grado de diversidad del sistema determina el número máximo de ranuras asignadas a diferentes usuarios.

1.2.2 Diversidad Frecuencial

Las versiones de señal son obtenidas observando la señal recibida en bandas frecuenciales diferentes. Al igual que en las técnicas de diversidad temporal, cuando se desea combatir el desvanecimiento una misma señal ha de ser transmitida simultáneamente en diferentes bandas de frecuencia.

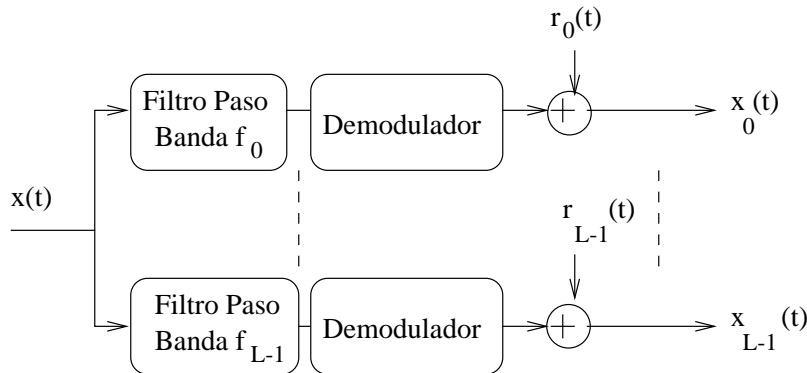


Figura 1.2: Diversidad frecuencial.

La técnica de acceso multiusuario basada en este tipo de diversidad se denomina FDMA (Frequency Division Multiple Access). En ella, cada uno de los usuarios transmite a frecuencias diferentes y el receptor sólo ha de filtrar una para obtener la señal del usuario deseado. La figura 1.2 representa la forma de separar las señales en este tipo de diversidad.

1.2.3 Diversidad Espacial

Las versiones de la señal se obtienen por medio de su recepción en diversos puntos del espacio. La diversidad espacial se consigue desplegando un array de antenas que se compone de múltiples sensores separados físicamente entre sí y dispuestos en un orden predeterminado para garantizar ciertas propiedades de las señales recibidas. En la figura 1.3 se representa una antena en array con K sensores.

Un sistema monousuario aprovecha la separación espacial de los sensores para obtener versiones de señal independientes y reducir así la probabilidad de alcanzar valores de desvanecimientos críticos. La técnica de acceso multiusuario basada en este tipo de diversidad se denomina SDMA (Spatial Division Multiple Access). Esta técnica permite la compartición de un medio a usuarios cuyas señales alcanzan el receptor por ángulos diferentes.

El número de sensores influye directamente en el grado de diversidad (número de versiones) que proporciona la antena. Este grado de diversidad se traduce en un

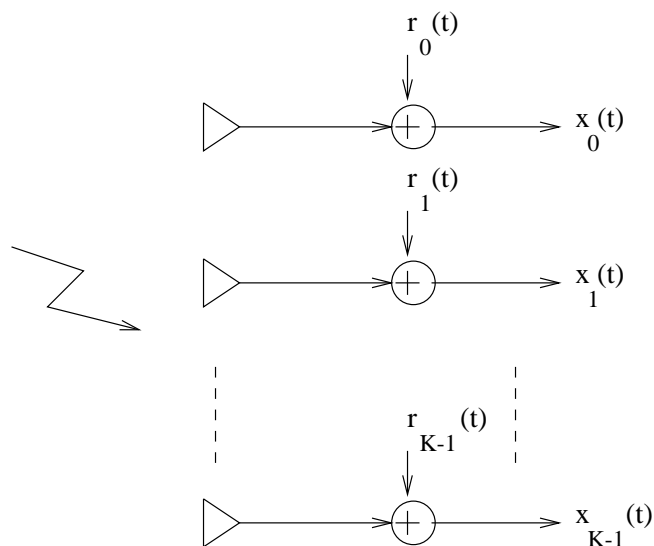


Figura 1.3: Diversidad espacial. Array de antenas.

aumento de la SNR en recepción y, en el caso de comunicaciones multiusuario, en un incremento del número de usuarios que un receptor puede detectar [16].

1.2.4 Diversidad por Espectro Ensanchado

En esta técnica de diversidad el espectro de frecuencias utilizado para transmitir una señal es mucho mayor que el de la señal moduladora. La operación de ensanchado de espectro se realiza por medio de códigos¹ independientes de la señal a transmitir, debiéndose deshacer esta operación para recuperar el espectro original. La modulación de espectro ensanchado permite combatir las interferencias en el canal, especialmente si son de banda estrecha, ya que el receptor sólo comprime el espectro de la señal deseada y expande el de las interferencias.

Existen varias técnicas de diversidad por espectro ensanchado siendo dos las más utilizadas: por secuencia directa (DS, Direct Sequence) y por salto de frecuencia (FH, Frequency Hopping). Ambas hacen uso de un código pseudo-aleatorio compuesto por una secuencia binaria con propiedades similares a las de un ruido. Los componentes de esta secuencia binaria se denominan *chips* y poseen un período mucho menor que el de un símbolo. Las técnicas DS logran el incremento del ancho de banda multiplicando los símbolos a transmitir por el código pseudo-aleatorio de forma que la señal obtenida tiene un mayor número de oscilaciones por unidad de tiempo. Por

¹La diversidad por espectro ensanchado también se conoce con el nombre de diversidad por código.

el contrario, las técnicas FH utilizan el código para producir saltos en la frecuencia de la señal logrando así el incremento del ancho de banda.

En un entorno multiusuario, este tipo de diversidad se denomina CDMA (Code Division Multiple Access) y considera que cada usuario tiene asociado un código diferente cuya correlación cruzada con los utilizados por otros usuarios es muy reducida. La figura 1.4 representa cómo se deshace el proceso de ensanchado de espectro de la señal recibida por medio de filtros adaptados a los códigos en un sistema DS/CDMA. Las versiones de señal utilizadas en este caso son las correspondientes a las salidas de los filtros adaptados.

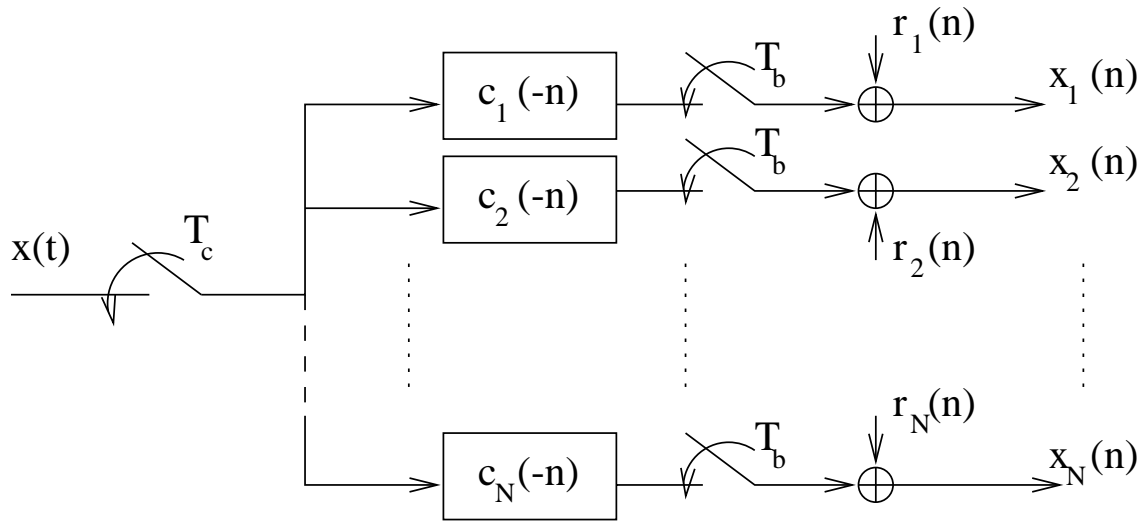


Figura 1.4: Diversidad por código.

1.2.5 Otras Formas de Diversidad

En este grupo se consideran las técnicas de diversidad por polarización y por componentes de campo. Estas técnicas proporcionan un bajo grado de diversidad (únicamente dos versiones de señal) y su uso no está tan extendido como las anteriores.

- Diversidad por polarización. Una antena puede considerar dos tipos de polarización: horizontal y vertical. De cada una de estas polarizaciones se puede extraer una señal que esta incorrelada con la otra [38].
- Diversidad por componentes de campo. Este tipo de diversidad aprovecha la teoría de campos electromagnéticos. Aunque los campos eléctrico, E , y magnético, H , de una señal electromagnética están mutuamente relacionados, las señales

recibidas por cada uno de éstos sufrirán distorsiones diferentes ya que fenómenos como la reflexión actúan de distinta forma según el campo considerado [38].

Como puede verse, las técnicas de diversidad existentes son muchas y muy diferentes. No obstante, es curioso observar que el modelo de señal correspondiente a todas ellas tiene enormes similitudes. En las subsecciones siguientes se plantearán los distintos modelos de señal para sistemas de comunicación monousuario y multiusuario. Primero se abordará el caso más sencillo de canales no dispersivos y después el de canales dispersivos.

1.3 Modelo de Señal para Canales no Dispersivos.

Un canal se dice que es no dispersivo si su respuesta en frecuencia es de la forma

$$H(f) = H e^{j2\pi f n_0} \quad (1.1)$$

donde H representa un factor de atenuación y n_0 es el retardo de propagación. Ambos parámetros son considerados constantes durante el tiempo que se tarda en transmitir un símbolo, aunque su valor sea desconocido y pueda fluctuar de manera aleatoria de un símbolo a otro. Esto es lo que se conoce con el nombre de desvanecimiento plano (*flat fading*) [56].

En el dominio del tiempo, la señal recibida se puede expresar en discreto de la siguiente manera

$$x(n) = H s(n - n_0) + r(n) \quad (1.2)$$

donde $s(n)$ representa la información transmitida y $r(n)$ contiene la suma de todas las señales no deseadas (ruido e interferencias) que existen en el canal.

Cuando el sistema de comunicaciones no se ve afectado por interferencias externas, $r(n)$ sólo contiene al ruido térmico presente en los dispositivos electrónicos del receptor. En este caso, es práctica común modelarlo como un proceso estocástico gaussiano, estadísticamente independiente de $s(n)$, con media cero y función de autocorrelación

$$E[r^*(n)r(n+m)] = \frac{N_0}{2} \delta(m) \quad (1.3)$$

donde $\frac{N_0}{2}$ es la densidad espectral de ruido y $\delta(m)$ es el impulso unitario. Cuando existen interferencias en el sistema $r(n)$ también se modela como un proceso estocástico estadísticamente independiente de $s(n)$ aunque su distribución y función de autocorrelación pueden adoptar formas muy diversas.

El modelo de señal que se ha descrito corresponde al de un receptor extremadamente simple que carece de diversidad y sólo proporciona una versión de la señal transmitida. Esto limita considerablemente las técnicas que pueden utilizarse para

mejorar la recepción ya que, si la señal deseada sufre un fuerte desvanecimiento, el ruido y/o las interferencias pueden enmascararla y el sistema no podrá recuperar la señal. Si se desea dotar al receptor de algún mecanismo para combatir los desvanecimientos, hay que recurrir a técnicas de diversidad que proporcionen varias versiones de la misma señal transmitida. La figura 1.5 muestra el diagrama de bloques de un sistema de comunicaciones monousuario en el que sólo se transmite $s(n)$ a través de un canal y que utiliza alguna de las técnicas de diversidad descritas anteriormente. Cada una de las L señales proporcionadas por la técnica de diversidad se puede escribir de la siguiente manera

$$x_i(n) = H_i s(n) + r_i(n) \quad (1.4)$$

donde H_i y $r_i(n)$ representan la atenuación y el ruido correspondientes a la versión i -ésima recibida. Por razones de sencillez, se ha considerado que la señal alcanza el receptor con un retardo de propagación nulo, $n_0 = 0$. Cuando las técnicas de diversidad se diseñan con el objetivo de combatir el desvanecimiento, se persigue que los valores de H_i sean estadísticamente independientes. De esta forma, si la probabilidad de que H_i atenúe la señal por debajo de un determinado valor crítico, haciendo que el ruido enmascare la señal recibida, es P , la probabilidad de que esto ocurra en todas las versiones recibidas es P^L .

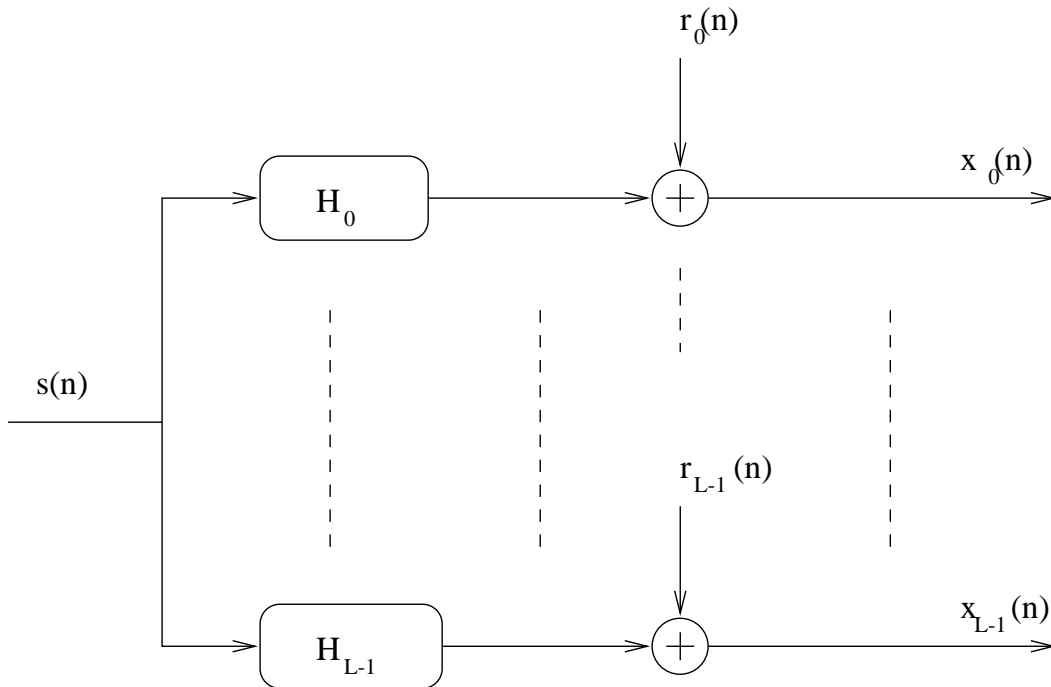


Figura 1.5: Diagrama de bloques un sistema de comunicaciones monousuario con diversidad y canales no dispersivos.

Además de combatir el desvanecimiento, los receptores con diversidad también

pueden utilizarse para separar las señales correspondientes a usuarios que comparten el mismo canal. La figura 1.6 representa el diagrama de bloques correspondiente a un sistema de comunicaciones con N usuarios, donde las señales transmitidas se superponen en recepción después de haber sufrido atenuaciones diferentes. En este caso, cada una de las señales recibidas se puede escribir de la siguiente manera

$$x_i(n) = \sum_{j=1}^N H_{ij} s_j(n) + r_i(n) \quad (1.5)$$

donde H_{ij} es la atenuación que sufre el usuario j en la versión i de las señales recibidas y $s_j(n)$ es la señal transmitida correspondiente al usuario j . La expresión (1.5) se puede reescribir de la siguiente manera

$$x_i(n) = H_{il} s_l(n) + r_i(n) + \sum_{\substack{j=1 \\ j \neq l}}^N H_{ij} s_j(n) \quad (1.6)$$

donde el último término representa las interferencias que sobre el usuario l causan los otros usuarios en el entorno. Esto es lo que se conoce con el nombre de interferencia de acceso múltiple (MAI).

Para describir el modelo de señal de una forma más compacta, las señales $x_i(n)$ se pueden agrupar en un vector de la forma

$$\mathbf{x}(n) = [x_0(n), x_1(n), \dots, x_{L-1}(n)]^T \quad (1.7)$$

donde el superíndice T representa transposición.

La forma del vector de versiones (1.7) depende de las señales del entorno y de la técnica de diversidad utilizada, pero en todas ellas se puede escribir de la siguiente forma

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{r}(n) \quad (1.8)$$

donde \mathbf{H} es una matriz de dimensiones $L \times N$ que contiene todas las atenuaciones H_{ij} , $\mathbf{s}(n) = [s_1(n), \dots, s_N(n)]^T$ y $\mathbf{r}(n) = [r_0(n), \dots, r_{L-1}(n)]^T$.

A continuación se describe la forma concreta que adopta este modelo según el tipo de diversidad.

1.3.1 Diversidad Temporal y Frecuencial

En estas dos técnicas los N usuarios transmiten en N ranuras temporales (diversidad temporal) o N bandas frecuenciales (diversidad frecuencial) distintas. Es habitual intercalar tiempos y bandas de frecuencia de guarda entre las señales para que al multiplexar los usuarios apenas se interfieran entre sí. Como consecuencia, cada una

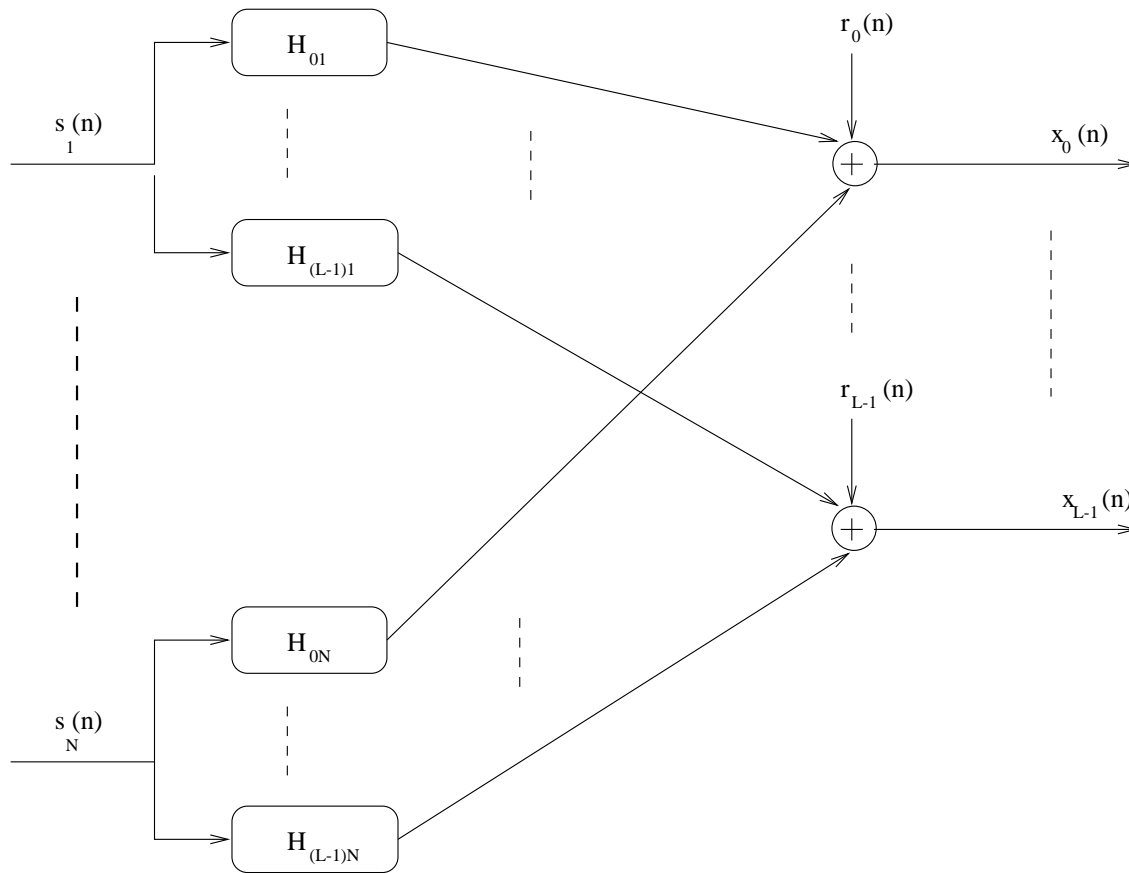


Figura 1.6: Diagrama de bloques un sistema de comunicaciones multiusuario con diversidad.

de las versiones de las señales observadas sólo tiene contribuciones de un sólo usuario y se pueden expresar de la forma

$$x_i(n) = H_{il}s_l(n) + r_i(n) \quad (1.9)$$

Si el modelo de señal se escribe en forma vectorial como la ecuación (1.8), la matriz \mathbf{H} resulta ser una matriz de permutación con un único elemento distinto de cero en cada columna. También es habitual encontrarse con que la matriz de covarianza del ruido resulta ser una matriz diagonal.

1.3.2 Diversidad Espacial

Este tipo de diversidad trata de aprovechar el hecho de que en muchos sistemas de comunicación por radio las señales en el entorno suelen llegar al receptor por ángulos diferentes. Puede demostrarse [16] que cuando una señal radioeléctrica de banda estrecha² se propaga como una onda plana, las señales suministradas por sensores dispuestos en distintos puntos del espacio se diferencian en un desfase que depende del ángulo de incidencia. La práctica habitual es situar los sensores separados entre sí una distancia igual a media longitud de onda lo que permite que haya una correspondencia biunívoca entre los ángulos de llegada y los desfases. Cuando hay desvanecimiento plano, sin embargo, la separación entre sensores suele ser muy superior a la longitud de onda para garantizar la independencia estadística de las atenuaciones en éstos. Esta última situación no está en los objetivos de esta tesis.

Considerando un entorno en el que N señales inciden sobre un array de K sensores, el vector de tamaño $K \times 1$ que contiene las señales de entrada se puede expresar de la siguiente forma

$$\mathbf{x}(n) = \sum_{i=1}^N A_i \mathbf{a}(\theta_i) s_i(n) + \mathbf{r}(n) \quad (1.10)$$

donde A_i es la amplitud con que es recibida la señal $s_i(n)$ y $\mathbf{a}(\theta_i)$ es el vector de *steering*³ correspondiente al ángulo θ_i por el que incide $s_i(n)$.

Cuando se emplea notación vectorial, la matriz \mathbf{H} en (1.8) es de la forma $\mathbf{H} = [A_1 \mathbf{a}(\theta_1), \dots, A_N \mathbf{a}(\theta_N)]$, es decir, los vectores de *steering* multiplicados por la amplitud con que los usuarios son recibidos. Siempre supondremos que las columnas de \mathbf{H} son linealmente independientes. Al menos ésto siempre puede garantizarse que es cierto cuando el array es lineal con sensores equiespaciados media longitud de onda y todas las señales inciden por ángulos distintos [55].

²Se habla de un modelo de propagación de banda estrecha cuando el retardo entre la llegada de una señal al primer sensor de la antena con respecto al último es despreciable.

³El vector de *steering* de una señal de banda estrecha que incide con un ángulo θ sobre un array de antenas con sus sensores linealmente equiespaciados una distancia de media longitud de onda, es $\mathbf{a}(\theta) = [1, e^{-j\phi}, e^{-j2\phi}, \dots, e^{-j(K-1)\phi}]^T$ donde $\phi = \pi \text{sen}(\theta)$.

En los sistemas con diversidad espacial, el ruido térmico es introducido por el receptor presente en cada uno de los sensores. Como estos receptores son dispositivos físicos diferentes sus ruidos son estadísticamente independientes entre sí y la matriz de correlación del ruido también resulta ser diagonal.

1.3.3 Diversidad por Código

En este apartado se va a describir el modelo de señal correspondiente a un sistema con diversidad de código que utiliza la modulación Direct Sequence (DS) expuesta anteriormente. Denominando $v(n)$ a la señal recibida cuando se muestrea a período de chip, T_c , se obtiene

$$v(n) = \sum_{i=1}^N v_i(n) + r(n) \quad (1.11)$$

donde $r(n)$ es el ruido gaussiano en el canal y $v_i(n)$ es la señal recibida correspondiente al usuario i -ésimo que tiene la forma

$$v_i(n) = A_i s_i(\lfloor \frac{n}{M} \rfloor) c_i(n - \lfloor \frac{n}{M} \rfloor M) \quad (1.12)$$

donde $\lfloor \cdot \rfloor$ representa la operación de redondeo al entero más bajo, A_i es la amplitud con que se recibe al usuario i -ésimo y $c_i(n)$ es su código asociado que posee una longitud de M chips. Obsérvese que los símbolos transmitidos se representan por la señal $s_i(\lfloor \frac{n}{M} \rfloor)$ que tiene el mismo valor durante los M chips correspondientes a un período de símbolo.

Las versiones de señal en esta técnica de diversidad son proporcionadas por las salidas de los filtros adaptados a los distintos códigos de usuario tal y como se muestra en la figura (1.4). El filtro l -ésimo está adaptado al código del usuario l , proporcionando a su salida la correlación entre la señal (1.11) y el código l en un período de símbolo, T_b . Cada M chips recibidos se obtiene una salida, a velocidad de símbolo, de la forma

$$x_l(n) = v(n) * c_l(-n) = \sum_{m=0}^{M-1} v(m + nM) c_l(m) \quad l = 1, \dots, N \quad (1.13)$$

Por otro lado, se puede comprobar que la señal $v(m + nM)$ en un período de símbolo es

$$v(m + nM) = \sum_{i=1}^N A_i s_i(n) c_i(m) + r(m + nM) \quad m \in [0, M - 1] \quad (1.14)$$

Sustituyendo (1.14) en (1.13) se obtiene

$$x_l(n) = \sum_{i=1}^N A_i s_i(n) \rho_{il} + r_l(n) \quad l = 1, \dots, N \quad (1.15)$$

donde ρ_{il} es la correlación entre los códigos $c_i(n)$ y $c_l(n)$

$$\rho_{il} = \sum_{m=0}^{M-1} c_i(m)c_l(m) \quad (1.16)$$

y

$$r_l(n) = \sum_{m=0}^{M-1} r(m + nM)c_l(m) \quad l = 1, \dots, N \quad (1.17)$$

es la componente de ruido a la salida del filtro l . Así pues, las señales obtenidas a la salida de un banco de filtros adaptados a los N códigos de usuario se pueden expresar matricialmente de la siguiente forma

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{r}(n) \quad (1.18)$$

donde \mathbf{H} es una matriz de correlación de tamaño $N \times N$ cuyo elemento (i, l) es $A_l\rho_{il}$. El vector $\mathbf{s}(n) = [s_1(n), \dots, s_N(n)]^T$ contiene los símbolos transmitidos por los N usuarios en el período n y $\mathbf{r}(n) = [r_1(n), r_2(n), \dots, r_N(n)]^T$ contiene las componentes de ruido a la salida de cada filtro. Obsérvese que, a diferencia de las técnicas de diversidad anteriores, las componentes del vector de ruido no son estadísticamente independientes entre sí. Su matriz de autocorrelación es $\mathbf{R}_{ruido} = \frac{N_0}{2}\mathbf{H}'$ donde \mathbf{H}' sólo contiene las componentes ρ_{il} .

Cuando los códigos son perfectamente ortogonales, $\rho_{il} = 0 \quad \forall i \neq l$, la matriz de correlación \mathbf{H} es diagonal y la salida de cada filtro adaptado contiene componentes de un único usuario. Sin embargo, esta hipótesis no se cumple en la práctica debido a asincronías en la transmisión o distorsiones del canal. Como consecuencia, la matriz \mathbf{H} no es diagonal y la salida de cada filtro adaptado suministra una mezcla de los símbolos transmitidos por diferentes usuarios. No obstante, en la mayor parte de las situaciones prácticas se cumple que las columnas de la matriz \mathbf{H} son linealmente independientes [57].

1.3.4 Discusión

Examinando los distintos modelos de señal que acabamos de presentar puede observarse que, con independencia del tipo de diversidad considerado, todo receptor con diversidad dispone de un *front-end* que suministra un vector de observaciones que es una mezcla lineal e instantánea de las señales que se transmiten. La mezcla está representada por \mathbf{H} , que de aquí en adelante llamaremos matriz de mezcla. Cuando la matriz de mezcla es diagonal (o una matriz resultante de permutar las filas en una matriz diagonal), cada componente del vector de observaciones se corresponde con la señal de un único usuario superpuesto con una componente de ruido. Sin embargo, en muchas situaciones prácticas la matriz de mezcla no será diagonal y en cada

componente de $\mathbf{x}(n)$ aparecen las contribuciones de diferentes usuarios denominada interferencia de acceso múltiple (MAI). Estas interferencias complican especialmente la recepción de una señal cuando existen grandes diferencias en las potencias recibidas como ocurre en comunicaciones móviles cuando se reciben señales de potencias muy diferentes (problema del *near-far*).

1.4 Modelo de Señal para Canales Dispersivos

En el apartado anterior hemos supuesto que, exceptuando al ruido y a las interferencias, la única distorsión que los canales introducen sobre las señales transmitidas es una atenuación que sólo modifica la potencia con que son recibidas. Esto simplifica notablemente el modelo de señales de entrada que, en el caso multiusuario, resulta ser una mezcla instantánea de los símbolos transmitidos. En la práctica los canales reales reflejan, refractan y dispersan las señales que pasan a través de ellos. El resultado es que la salida correspondiente a un único símbolo transmitido será una combinación de versiones retardadas, desfasadas y atenuadas de la entrada lo que tiene como consecuencia que una señal que a la entrada tenga una determinada duración aparezca a la salida con una duración mayor. Esta es la razón de que a estos canales se les llame dispersivos.

Los canales con multitrayecto son un claro ejemplo de canales dispersivos. En éstos una misma señal se propaga rebotando en diferentes obstáculos encontrados en el camino entre el transmisor y el receptor. Las diferentes reflexiones y refracciones que sufre la señal dan lugar a que se reciban a través de caminos diferentes, y con retardos de propagación distintos, múltiples señales provenientes de un mismo origen, que se superponen a la entrada del receptor. Este tipo de canales constituyen la principal limitación de la velocidad de transmisión en los sistemas de comunicaciones móviles [56]. Otros ejemplos de canales dispersivos pueden encontrarse en el contexto de medios guiados (canales telefónicos, par trenzado y cable coaxial). Prácticamente podemos decir que todo canal que tenga severamente limitado su ancho de banda presenta un comportamiento dispersivo.

Un canal dispersivo se puede representar como un filtro lineal [6]. En concreto, un canal causal con respuesta de duración limitada puede ser representado por un filtro FIR (Finite Impulse Response) con una respuesta al impulso unitario de la forma

$$h(n) = \sum_{k=0}^{P-1} h(k)\delta(n-k) \quad (1.19)$$

donde $\delta(n)$ es el impulso unitario y P es la longitud del canal. La señal obtenida en recepción cuando se transmite una secuencia $s(n)$ a través de este tipo de canales es

de la siguiente forma

$$x(n) = h(0)s(n) + \underbrace{\sum_{i=1}^{P-1} h(i)s(n-i)}_{ISI} + r(n) = \mathbf{h}^T \mathbf{s}(n) + \mathbf{r}(n) \quad (1.20)$$

donde $\mathbf{h} = [h(0), h(1), \dots, h(P-1)]^T$ es un vector con la respuesta al impulso y $\mathbf{s}(n) = [s(n), s(n-1), \dots, s(n-P+1)]^T$ es un vector que contiene parte de la secuencia transmitida. Por simplicidad, se ha supuesto que el símbolo deseado alcanza el receptor con un retardo nulo. Como puede observarse, la ISI está representada por el término del sumatorio que incluye el efecto de los $P-1$ símbolos posteriormente transmitidos.

El modelo de la ecuación (1.20) describe la señal recibida en un sistema sin diversidad debida a un único usuario. Cuando existen varios usuarios en el entorno y se emplean receptores con diversidad, el modelo de señal se complica notablemente tal y como se describe a continuación.

1.4.1 Línea de Retardos

La figura 1.7 presenta el diagrama de bloques correspondiente a una línea de retardos que proporciona versiones retardadas a valor T segundos de la señal recibida. En cierto sentido, este sistema puede considerarse como un *front-end* de comunicaciones dotado de diversidad temporal que puede ser aprovechado para combatir la ISI introducida por el canal dispersivo. Esto es precisamente lo que hace un igualador, tal y como se explica en el siguiente capítulo. Utilizando notación vectorial, las L señales presentes en la línea de retardos se pueden agrupar en un vector $\mathbf{x}(n)$ que se representa de la siguiente forma

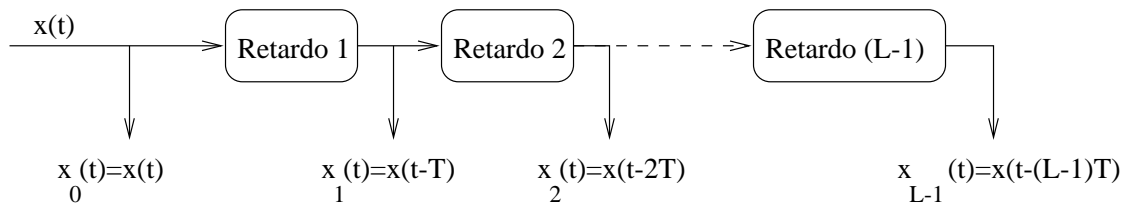


Figura 1.7: Diversidad temporal mediante una línea de retardos.

$$\mathbf{x}(n) = [x(n), \dots, x(n-L+1)]^T = \mathbf{H}\mathbf{s}(n) + \mathbf{r}(n) \quad (1.21)$$

donde \mathbf{H} es una matriz convolucional de tamaño $L \times (P + L - 1)$ de la forma

$$\mathbf{H} = \begin{pmatrix} h(0) & \cdots & h(P-1) & 0 & \cdots & \cdots & 0 \\ 0 & h(0) & \cdots & h(P-1) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & h(0) & \cdots & h(P-1) & \cdots & 0 \end{pmatrix} \quad (1.22)$$

$\mathbf{s}(n) = [s(n), s(n-1), \dots, s(n-L-P+2)]^T$ es un vector que contiene parte de la secuencia transmitida por el usuario y $\mathbf{r}(n) = [r(n), r(n-1), \dots, r(n-L+1)]^T$ es un vector de tamaño $L \times 1$ conteniendo al ruido aditivo.

1.4.2 Array de Antenas

Los sensores dotados de diversidad espacial utilizan un array de antenas. Si suponemos que una señal transmitida $s_i(n)$ alcanza al array a través de múltiples reflexiones, cada una de ellas con un retardo y ángulo de llegada diferente, el vector $\mathbf{x}(n)$ de señales recibidas se puede escribir de la forma

$$\mathbf{x}(n) = \sum_{l=0}^{P-1} h_i(l) \mathbf{a}(\theta_{il}) s_i(n-l) + \mathbf{r}(n) \quad (1.23)$$

donde h_l es la amplitud del multitrayecto l -ésimo y $\mathbf{a}(\theta_{il})$ su vector de steering. La ecuación (1.23) se puede reescribir de forma más compacta como

$$\mathbf{x}(n) = \mathbf{H}_i \mathbf{s}_i(n) + \mathbf{r}(n) \quad (1.24)$$

donde $\mathbf{s}_i(n) = [s_i(n), s_i(n-1), \dots, s_i(n-P+1)]^T$ y \mathbf{H}_i es una matriz de tamaño $L \times P$ que contiene los efectos producidos por el canal dispersivo y los desfases provocados por el array de antenas. En el caso de un array lineal con sensores equiespaciados media longitud de onda, esta matriz adopta la forma

$$\mathbf{H}_i = \begin{pmatrix} h_i(0) & h_i(1) & \cdots & h_i(P-1) \\ h_i(0)e^{j\phi_{i0}} & h_i(1)e^{j\phi_{i1}} & \cdots & h_i(P-1)e^{j\phi_{i(P-1)}} \\ \vdots & \vdots & \ddots & \vdots \\ h_i(0)e^{j(K-1)\phi_{i0}} & h_i(1)e^{j(K-1)\phi_{i1}} & \cdots & h_i(P-1)e^{j(K-1)\phi_{i(P-1)}} \end{pmatrix} \quad (1.25)$$

donde los desfases ϕ_{ij} dependen de los ángulos de incidencia de los multitrayectos.

Cuando existen N usuarios en el entorno, la señal recibida es

$$\mathbf{x}(n) = \sum_{i=1}^N \mathbf{H}_i \mathbf{s}_i(n) + \mathbf{r}(n) = \mathbf{H} \mathbf{s}(n) + \mathbf{r}(n) \quad (1.26)$$

donde $\mathbf{s}(n) = [s_1^T(n), \dots, s_N^T(n)]^T$ es un supervector de tamaño $NP \times 1$ conteniendo las secuencias transmitidas y $\mathbf{H} = [\mathbf{H}_1, \dots, \mathbf{H}_N]$ es una supermatriz de dimensión $K \times NP$ que contiene los efectos del canal.

1.4.3 Banco de Filtros Adaptados

El modelo de las señales proporcionadas por un banco de filtros adaptados en un sistema DS/CDMA es muy similar al del apartado anterior como se verá a continuación. Consideremos, inicialmente, la señal transmitida por el usuario i -ésimo que resulta de multiplicar los símbolos $s_i(n)$ por el código $c_i(n)$, es decir,

$$s_i^c(n) = s_i(\lfloor \frac{n}{M} \rfloor) c_i(n - \lfloor \frac{n}{M} \rfloor M) \quad (1.27)$$

Este producto puede ser expresado por una suma de convolución de la forma

$$s_i^c(n) = (s_i(\lfloor \frac{n}{M} \rfloor) \delta(n - \lfloor \frac{n}{M} \rfloor M)) * c_i(n) \quad (1.28)$$

donde $*$ denota convolución y $\delta(n - \lfloor \frac{n}{M} \rfloor M)$ es un tren de impulsos unitarios situados en los instantes kM . La señal recibida es el resultado de transmitir $s_i^c(n)$ por el canal dispersivo de forma que la señal obtenida se puede representar como

$$v_i(n) = s_i^c(n) * h_i(n) = \sum_{k=0}^{P-1} s_i^c(n-k) h_i(k) + r(n) \quad (1.29)$$

donde $h_i(n)$ es la respuesta al impulso del canal correspondiente al usuario i -ésimo. Consideraremos que la longitud del canal es inferior a la de un símbolo⁴, es decir, $P < M$. Quiere esto decir que un símbolo sólo se ve afectado por el transmitido inmediatamente después. A continuación se representa el efecto del código y del canal en conjunto por

$$g_i(n) = h_i(n) * c_i(n) = \sum_{k=0}^{M-1} h_i(n-k) c_i(k) \quad n = 0, \dots, M+P-2 \quad (1.30)$$

lo que nos permite expresar la ecuación (1.29) de la siguiente forma

$$\begin{aligned} v_i(n) &= (s_i(\lfloor \frac{n}{M} \rfloor) \delta(n - \lfloor \frac{n}{M} \rfloor M)) * g_i(n) + r(n) = \\ &= \sum_{k=0}^{M+P-2} (s_i(\lfloor \frac{n-k}{M} \rfloor) \delta(n-k - \lfloor \frac{n-k}{M} \rfloor M)) g_i(k) + r(n) \end{aligned} \quad (1.31)$$

⁴Esta hipótesis no es muy restrictiva en la práctica.

Esta ecuación representa a la señal recibida $s_i(n)$ correspondiente al usuario i -ésimo y constituye la entrada al banco de filtros adaptados a los códigos $c_l(n)$, $l = 1, \dots, N$. Suponiendo un sistema de transmisión síncrono, las salidas de estos filtros se pueden representar como

$$x_l(n) = \sum_{k=0}^{M-1} v_i(n+k)c_l(k) = \mathbf{c}_l^T \mathbf{v}_i(n) \quad l = 1, \dots, N \quad (1.32)$$

donde $\mathbf{c}_l = [c_l(0), c_l(1), \dots, c_l(M-1)]^T$ es el vector del código l -ésimo y $\mathbf{v}_i(n) = [v_i(n), \dots, v_i(n+M-1)]^T$ contiene los M chips que entran en el filtro adaptado cada período de símbolo. Teniendo en cuenta (1.31), este vector se puede expresar como

$$\mathbf{v}_i(n) = \mathbf{G}_i \mathbf{s}_i(n) + \mathbf{r}(n) \quad (1.33)$$

donde $\mathbf{r}(n) = [r(n), \dots, r(n+M-1)]^T$. Como se ha supuesto que la duración del canal es inferior a la del período de símbolo (hipótesis $P < M$), el vector $\mathbf{s}_i(n) = [s_i(\lfloor \frac{n}{M} \rfloor - 1), s_i(\lfloor \frac{n}{M} \rfloor)]^T$ sólo tiene dos componentes y la matriz \mathbf{G}_i tiene dimensiones $M \times 2$ siendo

$$\mathbf{G}_i = \begin{pmatrix} g_i(M) & g_i(0) \\ \vdots & \vdots \\ g_i(M+P-2) & g_i(P-2) \\ 0 & g_i(P-1) \\ \vdots & \vdots \\ 0 & g_i(M-1) \end{pmatrix} \quad (1.34)$$

Reescribiendo (1.32) en forma matricial

$$\mathbf{x}(n) = [x_1(n), \dots, x_N(n)]^T = \mathbf{C} \mathbf{v}_i(n) \quad (1.35)$$

donde

$$\mathbf{C} = \begin{pmatrix} \mathbf{c}_1^T(n) \\ \vdots \\ \mathbf{c}_N^T(n) \end{pmatrix} \quad (1.36)$$

Sustituyendo (1.33) en (1.35) se obtiene

$$\mathbf{x}(n) = \mathbf{C} \mathbf{G}_i \mathbf{s}_i(n) + \mathbf{C} \mathbf{r}(n) \quad (1.37)$$

Finalmente, la consideración de los N usuarios del sistema es inmediata ya que el vector de entrada al banco de filtros adaptados consiste en la superposición de las señales, $v_i(n)$, correspondientes a los diferentes usuarios de la siguiente forma

$$\begin{aligned} \mathbf{v}(n) &= \sum_{i=1}^N \mathbf{G}_i \mathbf{s}_i(n) + \mathbf{r}(n) = \\ &= [\mathbf{G}_1, \dots, \mathbf{G}_N] \mathbf{s}(n) + \mathbf{r}(n) \end{aligned} \quad (1.38)$$

donde $\mathbf{s}(n) = [\mathbf{s}_1^T(n), \dots, \mathbf{s}_N^T(n)]^T$ es un supervector que contiene símbolos correspondientes a varios períodos de los diferentes usuarios del sistema.

Sustituyendo (1.38) en (1.35) se obtiene la expresión general del vector de entrada

$$\mathbf{x}(n) = (\mathbf{C}[\mathbf{G}_1, \dots, \mathbf{G}_N])\mathbf{s}(n) + \mathbf{C}\mathbf{r}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{C}\mathbf{r}(n) \quad (1.39)$$

donde $\mathbf{H} = \mathbf{C}[\mathbf{G}_1, \dots, \mathbf{G}_N]$.

1.4.4 Discusión

Examinando los resultados de las subsecciones anteriores se puede concluir que, al igual que ocurría en el apartado anterior para canales no dispersivos, el modelo de señal resultante es independiente de la técnica de diversidad utilizada. En todos los casos el vector de observaciones consiste en una mezcla lineal de las señales transmitidas por los N usuarios en diferentes instantes de tiempo que se puede representar de la forma

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{r}(n) \quad (1.40)$$

donde \mathbf{H} depende de las características del canal y del tipo de diversidad, $\mathbf{s}(n)$ contiene los símbolos transmitidos y $\mathbf{r}(n)$ es un vector de ruido gaussiano cuya matriz de autocorrelación depende de la técnica de diversidad.

La matriz de mezcla, \mathbf{H} , ha de ser de rango completo en las columnas para garantizar la separación entre las señales de los diferentes usuarios [64]- [9]. Si esta matriz no fuese de rango completo, sus columnas tendrían alguna dependencia lineal y el problema se diría que está mal condicionado. En ciertos entornos es posible aumentar el grado de diversidad permitiendo eliminar las dependencias en las columnas de \mathbf{H} , pero en otros casos, como los abordados en el capítulo 2, la matriz no tiene sus columnas linealmente independientes y plantea problemas en la separación de los diferentes usuarios.

1.5 Arquitecturas Receptoras

En la sección anterior se ha visto que en todo sistema de comunicaciones la señal recibida es una mezcla lineal y ruidosa de las transmitidas. El objetivo de todo sistema receptor es procesar las señales recibidas para extraer la información deseada. En un principio, las arquitecturas de los receptores pueden ser muy diversas pero esta tesis se restringe a aquellas que realizan un procesamiento lineal de las señales recibidas.

A grandes rasgos, los receptores se pueden clasificar en dos grandes grupos:

- **Receptores Monousuario:** en los que el objetivo es extraer la información transmitida por un único usuario. Estos son los receptores a emplear en termina-

les personales como por ejemplo un teléfono móvil individual. Estos receptores utilizan una arquitectura denominada MISO (Multiple Input Single Output) que tiene una única salida correspondiente a una combinación lineal de las L señales de entrada de la forma

$$y(n) = \sum_{k=0}^{L-1} w_k^* x_k(n) \quad (1.41)$$

donde el superíndice $*$ representa conjugado y w_i son los coeficientes complejos del receptor que definen la combinación realizada. Esta expresión puede ser representada en forma vectorial de manera más compacta por medio del producto escalar de dos vectores

$$y(n) = \mathbf{w}^H \mathbf{x}(n) \quad (1.42)$$

donde $\mathbf{w} = [w_0, \dots, w_{L-1}]^T$ es el vector de coeficientes complejos del receptor, el superíndice H denota transpuesto conjugado y $\mathbf{x}(n)$ es el vector de señales de entrada.

- **Receptores Multiusuario:** en los que el objetivo es extraer la información de todos los usuarios a la vez. Estos son los receptores a utilizar en los nodos de comunicaciones como las estaciones base de telefonía móvil. El receptor, en este caso, utiliza una arquitectura denominada MIMO (Multiple Input Multiple Output) que tiene varias salidas cada una de las cuales será una combinación lineal de las señales de entrada de la forma

$$y_i(n) = \mathbf{w}_i^H \mathbf{x}(n) \quad i = 1, \dots, N \quad (1.43)$$

Este también se puede escribir de forma vectorial como

$$\mathbf{y}(n) = \mathbf{W}^H \mathbf{x}(n) \quad (1.44)$$

donde $\mathbf{y}(n) = [y_1(n), \dots, y_N(n)]^T$ y $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_N]$ es una matriz de dimensión $L \times N$

Los receptores que se pueden encontrar según la aplicación considerada son los siguientes:

1.5.1 Receptores MISO para Sistemas con Diversidad Temporal

Las entradas del sistema MISO son obtenidas a partir de versiones retardadas de la señal recibida. Esta arquitectura podría considerarse como un sistema SISO (Single Input Single Output) pues realmente sólo se hace uso de una señal recibida. Sin

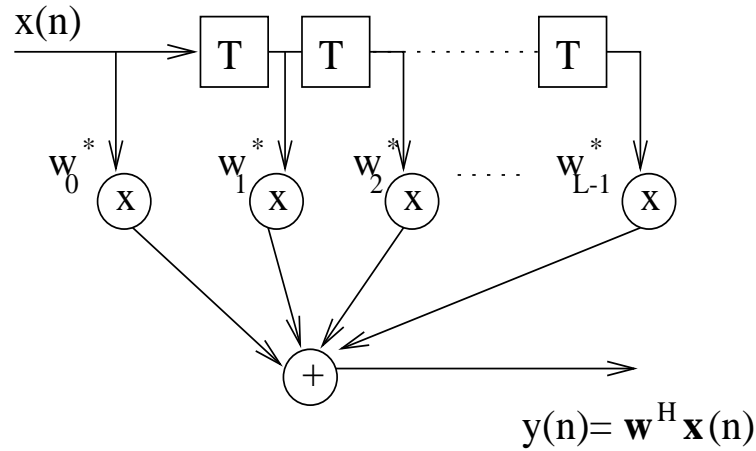


Figura 1.8: Arquitectura de receptor MISO para sistemas con diversidad temporal.

embargo, una línea de retardos proporciona versiones de señal que poseen diferentes mezclas de la secuencia transmitida. En la figura 1.8 se puede observar cómo se combinan las diferentes muestras proporcionadas por la línea de retardos. El número de señales de entrada, L , es igual a la longitud de la línea de retardos, es decir, el número de retardos más uno (la observación en el instante actual).

Este tipo de receptores son utilizados habitualmente para combatir los efectos de un canal dispersivo.

1.5.2 Receptores MISO para Sistemas con Diversidad Espacial o por Código

Las entradas del sistema MISO son proporcionadas por los sensores de una antena en array (diversidad espacial) o por las salidas de un banco de filtros adaptados a los códigos de usuario (diversidad por código). La figura 1.9 muestra el diagrama de bloques de este tipo de receptores. El número de señales de entrada es igual al número de versiones de señal, K , proporcionadas por la técnica de diversidad. Cuando la diversidad es obtenida a través de una técnica de diversidad espacial, el número de señales de entrada es igual al número de sensores de la antena. Sin embargo, cuando la diversidad se obtiene por código, este número es igual al de filtros adaptados, es decir, al número de usuarios.

Este tipo de receptores son utilizados habitualmente para combatir la interferencia de acceso múltiple (MAI) presente en comunicaciones multiusuario. Estos también se pueden utilizar para combatir los efectos dispersivos de los canales con multitrayecto, ya que permiten extraer uno de los trayectos que alcanzan el receptor.

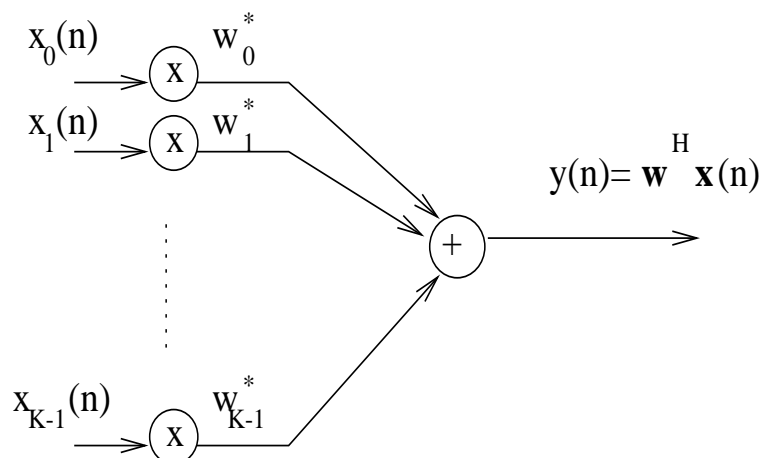


Figura 1.9: Arquitectura de receptor MISO para sistemas con diversidad espacial o por código.

1.5.3 Receptores MISO para Sistemas con Doble Diversidad

Las entradas del sistema MISO son el resultado de combinar dos tipos diferentes de diversidad. Considerando los tres tipos de diversidad más utilizados se pueden obtener tres arquitecturas de receptor:

- Diversidad por Código y Temporal. La figura 1.10 representa el diagrama de bloques de este tipo de receptores. Como se puede observar, cada una de las señales proporcionadas por los N filtros adaptados a los códigos de usuario son la entrada a una línea de retardos que proporciona la diversidad temporal. El número de señales de entrada al receptor es LN .

Este tipo de receptores se utilizan habitualmente para combatir los efectos de un canal dispersivo y la interferencia de acceso múltiple, simultáneamente.

- Diversidad Espacial y Temporal. Este tipo de receptor es idéntico al anterior con la única diferencia en que las señales de entrada a la línea de retardos son obtenidas de cada uno de los K sensores de la antena. El número de señales de entrada al receptor es KL . Además, el campo de aplicación de estos receptores es el mismo que en el caso anterior.
- Diversidad Espacial y por Código. La figura 1.11 representa el diagrama de bloques de este tipo de receptores. En este caso tenemos K bancos de N filtros adaptados, uno por cada uno de los K sensores de la antena. El número de señales de entrada al receptor es KN . Aunque como se mencionó anteriormente, cualquiera de las dos diversidades pueden ser utilizadas para combatir el efecto

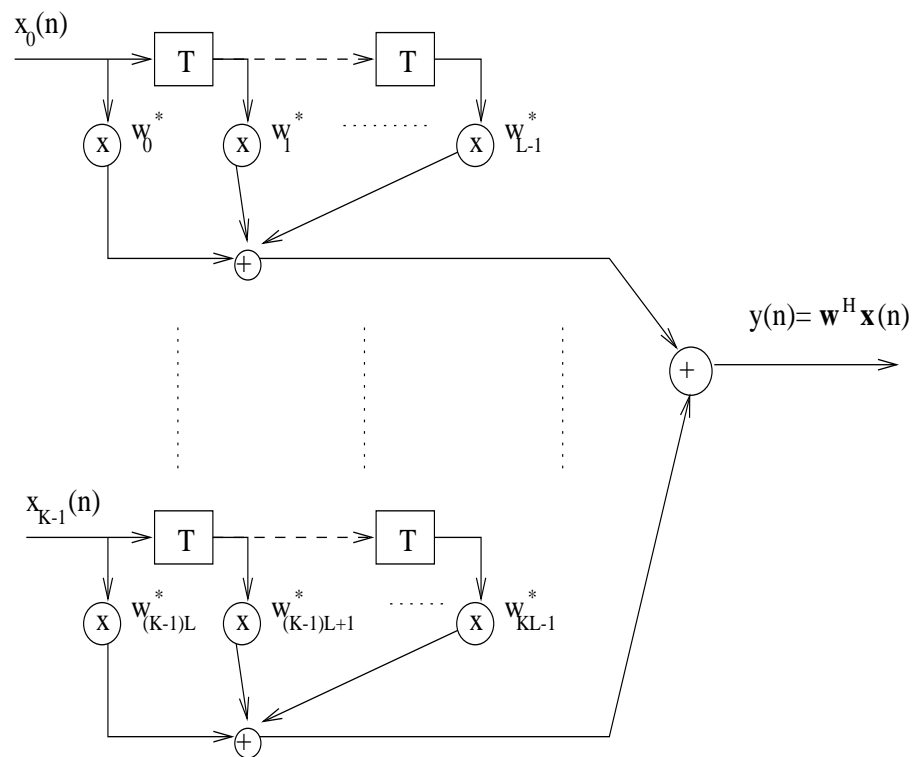


Figura 1.10: Arquitectura de receptores MISO para sistemas con doble diversidad: código o espacial y temporal.

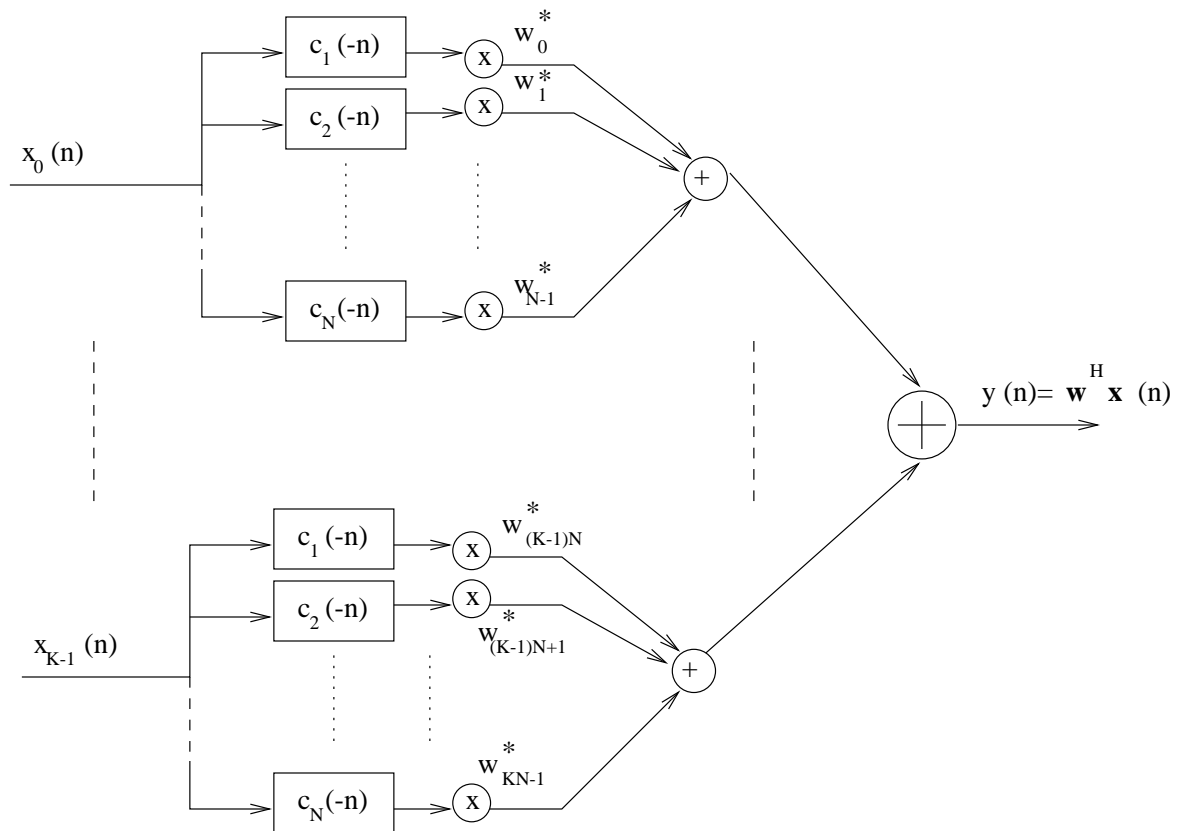


Figura 1.11: Arquitectura de receptores MISO para sistemas con doble diversidad: espacial y por código.

dispersivo de un canal, habitualmente este tipo de receptores son utilizados para aumentar la capacidad de un sistema de comunicaciones multiusuario.

1.5.4 Receptores MISO para Sistemas con Triple Diversidad

Las entradas del sistema MISO son el resultado de combinar tres tipos de diversidad: espacial, código y temporal. La figura 1.12 representa un diagrama de un receptor de este tipo en el que se ha considerado que cada una de las señales obtenidas por K sensores de una antena son la entrada a un banco de N filtros adaptados a los códigos de usuario cuyas salidas son aplicadas a una línea de retardos de longitud L . El número de señales de entrada al receptor es KNL . Este tipo de receptores podrían ser utilizados para aumentar la capacidad de un sistema de comunicaciones multiusuario y para combatir los efectos de un canal dispersivo, simultáneamente.

1.5.5 Receptores MIMO

Cuando el receptor ha de considerar múltiples salidas donde se extraen los diferentes usuarios de un sistema de comunicaciones multiusuario, se habla de sistemas MIMO (Multiple Input Multiple Output). Como muestra la figura 1.13, estos sistemas se pueden considerar como conexiones en paralelo de N sistemas MISO (uno por cada usuario a extraer) con las mismas entradas. Estos receptores requieren extraer los diferentes usuarios del sistema eliminando la interferencia de acceso múltiple. Además, según sea la arquitectura interna de los sistemas MISO conectados en paralelo, estos receptores podrán también combatir efectos dispersivos de los canales.

1.6 Planteamiento y Objetivos de la Tesis

Las arquitecturas de receptores que se han descrito en la sección anterior combinan linealmente las señales suministradas por un *front-end* con diversidad para producir la salida. Los coeficientes de estas combinaciones lineales deben ajustarse para que a la salida se obtenga la información deseada. Desgraciadamente, las características de los canales no son conocidas *a priori* y suelen variar con el tiempo, lo cual obliga a ajustar los parámetros de los receptores de acuerdo con un criterio estadístico. El criterio más utilizado en filtrado adaptativo es el de Error Cuadrático Mínimo (MMSE, Minimum Mean Square Error) en el que los coeficientes del sistema receptor se ajustan con el objetivo de minimizar el error cuadrático medio (MSE) entre la salida y los símbolos transmitidos. Este criterio, sin embargo, plantea el problema de que los algoritmos prácticos que llevan a cabo la minimización del MSE necesitan conocer los símbolos transmitidos lo cual en muchas aplicaciones de comunicaciones no

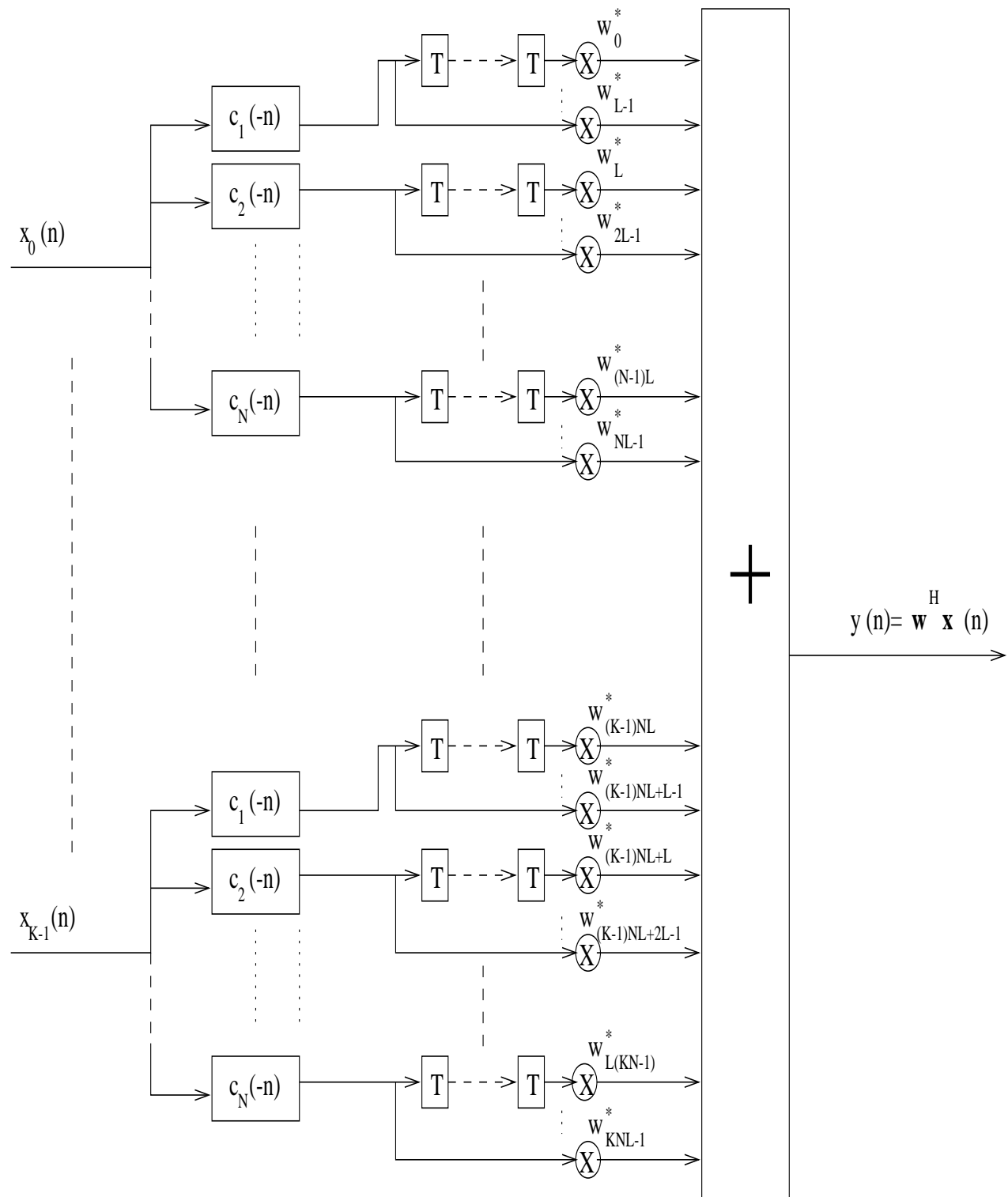


Figura 1.12: Arquitectura de receptores MISO para sistemas con triple diversidad.

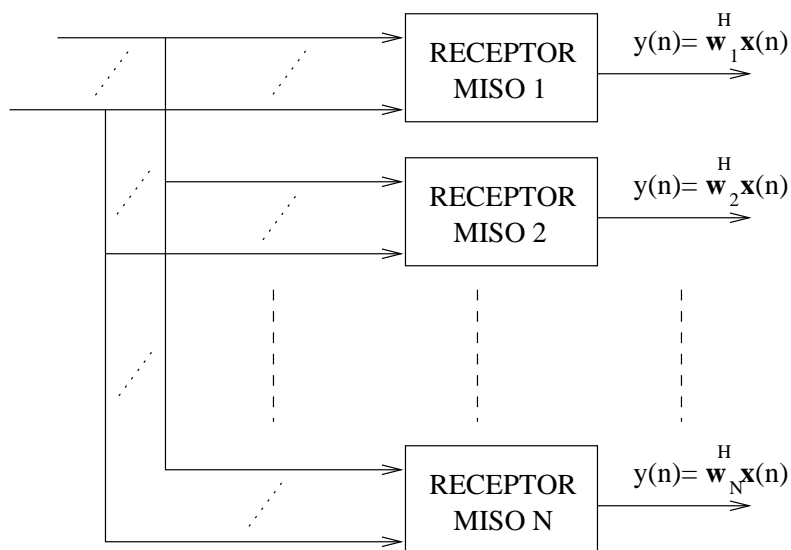


Figura 1.13: Arquitectura de receptores MIMO.

es posible. Para evitar estas limitaciones se han propuesto recientemente numerosos criterios, denominados ciegos, que ajustan los coeficientes de un receptor aprovechando propiedades estadísticas de las señales transmitidas y no necesitan conocer los símbolos transmitidos.

Esta tesis centra sus objetivos en el problema de igualación y en el de separación de usuarios en sistemas con diversidad. Se propondrán dos nuevos criterios estadísticos que se basan en el bien conocido criterio de Constant Modulus (CM). Este criterio fue inicialmente propuesto para igualación de canales y tiene por objetivo restaurar la propiedad de módulo constante que la señal tiene cuando es transmitida y que se pierde por el efecto del canal. Veremos en los siguientes capítulos cómo este criterio debe ser modificado para poder ser empleado satisfactoriamente en las aplicaciones que se consideran.

La estructura de la tesis es la siguiente. En el capítulo 2 se aborda la limitación que presenta el criterio CM de contener mínimos locales en los que la ISI no se elimina satisfactoriamente cuando se utiliza en igualación de canales con respuesta al impulso infinita (IIR, Infinite Impulse Response) de tipo auto-regresivo (AR, Auto-Regressive). Se presenta una sencilla modificación a la función de coste CM que no presenta este problema. El criterio CM también presenta dificultades cuando se utiliza para eliminar la MAI en sistemas de comunicación multiusuario con receptores MIMO ya que la misma señal se puede extraer en salidas distintas. Este es el problema que se aborda en el capítulo 3 en donde también se propondrá una modificación del criterio CM para resolverlo. Finalmente, en el capítulo 4 se hace un resumen de las aportaciones de la tesis y se presentan las líneas futuras de trabajo que quedan

abiertas.

Capítulo 2

Igualación de Canales AR

2.1 Introducción

En los capítulos anteriores hemos explicado como en todo sistema de comunicaciones las señales enviadas por un transmisor a través de un canal dispersivo, llegan al receptor de forma distorsionada. En estos casos, el canal introduce una dispersión temporal que alarga la duración de los pulsos transmitidos de forma que un símbolo interfiere con los transmitidos anterior y posteriormente. Este fenómeno se conoce con el nombre de Interferencia Intersimbólica (ISI, InterSymbol Interference). Ejemplos de este tipo de canales los podemos encontrar en el bucle de abonado telefónico que utiliza par trenzado, en televisión digital a través de cable coaxial o en comunicaciones móviles afectadas por propagación multitrayecto.

En el capítulo anterior también mostramos que todo canal dispersivo puede ser modelado por un filtro con unos coeficientes complejos elegidos para obtener un modelo preciso de la respuesta al impulso del canal. Si la deformación que introduce un canal puede ser modelada como un filtrado lineal e invariante en el tiempo parece lógico suponer que otro filtro lineal e invariante en el tiempo pueda ser diseñado en recepción para compensar (igualar) sus efectos. Este filtro no puede ser diseñado a priori debido al desconocimiento exacto de la respuesta al impulso del canal al comienzo de la transmisión o a las continuas variaciones temporales que puede sufrir durante una transmisión. La práctica habitual consiste en reajustar continuamente los coeficientes de este filtro para poder seguir los cambios del canal. Esto es lo que se denomina un igualador adaptativo, actualmente presentes en todos los receptores de comunicaciones digitales que aprovechan eficientemente el ancho de banda disponible.

Si exceptuamos los igualadores de máxima verosimilitud, los igualadores que se utilizan en la mayor parte de las situaciones prácticas pueden clasificarse en dos grandes grupos: lineales y con realimentación de decisiones (DFE, Decision Feedback

Equalizers)¹ [4]. Un igualador lineal es un filtro transversal convencional con una línea de retardos y unos coeficientes que se ajustan de acuerdo a un criterio estadístico. La figura 2.1 muestra el diagrama de bloques de un receptor monousuario que utiliza este tipo de igualadores. Un DFE, por el contrario, es un dispositivo no lineal que consta de dos filtros transversales y un dispositivo decisor dispuestos como muestra la figura 2.2. La señal recibida es la entrada a uno de los filtros transversales y la entrada al otro son las decisiones correspondientes a símbolos del pasado. La entrada al dispositivo decisor es la salida del primer filtro al que se le sustrae la salida del segundo. De esta forma se produce una realimentación capaz de eliminar la ISI causada por símbolos anteriores al que se está detectando. Los DFE son particularmente eficientes para igualar canales de fase no mínima con una fuerte componente de ISI anticausal. Sin embargo, su análisis es extremadamente difícil debido a su naturaleza no lineal y su comportamiento se degrada considerablemente cuando el decisor toma decisiones erróneas. Además, son más costosos de implementar que los igualadores lineales y pueden volverse inestables debido al camino de realimentación.

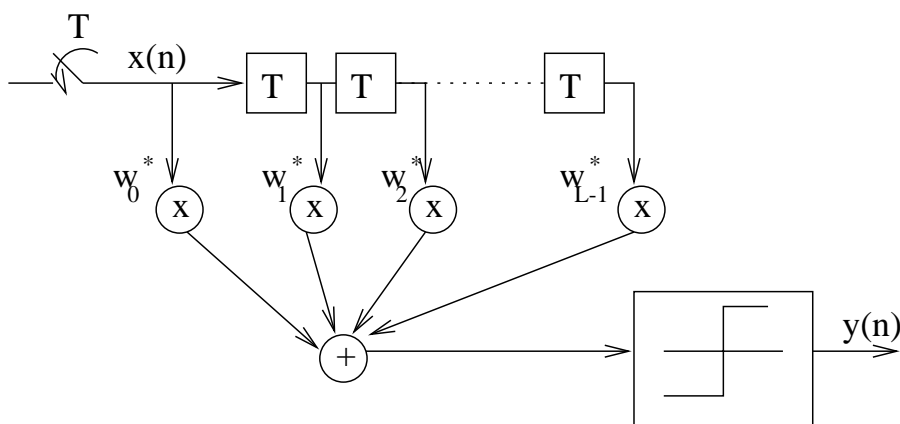


Figura 2.1: Diagrama de bloques de un receptor monousuario con igualador lineal.

Otra clasificación de los igualadores permite dividirlos en síncronos y en fraccionarios (FSE, Fractionally Spaced Equalizers) [28]. En los igualadores síncronos se utilizan líneas con retardos iguales al período de símbolo, mientras que en los fraccionarios los retardos son inferiores a dicho valor. Los FSE en principio tienen más coeficientes pero presentan una mayor robustez frente a errores de sincronismo.

Por razones de sencillez y brevedad en la exposición, en este capítulo sólo consi-

¹Los igualadores de máxima verosimilitud presentan una complejidad que crece exponencialmente con la longitud de la respuesta al impulso del canal, de ahí que sólo se utilicen en aplicaciones donde esta longitud sea de 2 ó 3 coeficientes, como es el caso de los sistemas de comunicaciones digitales celulares de baja velocidad (GSM). La complejidad de los igualadores lineales y DFE, sin embargo, crece linealmente con la longitud del canal que se pretende igualar y por ello son más utilizados en la práctica.

deraremos igualadores lineales y síncronos. No obstante, la extensión de las ideas a los otros tipos de igualadores es inmediata.

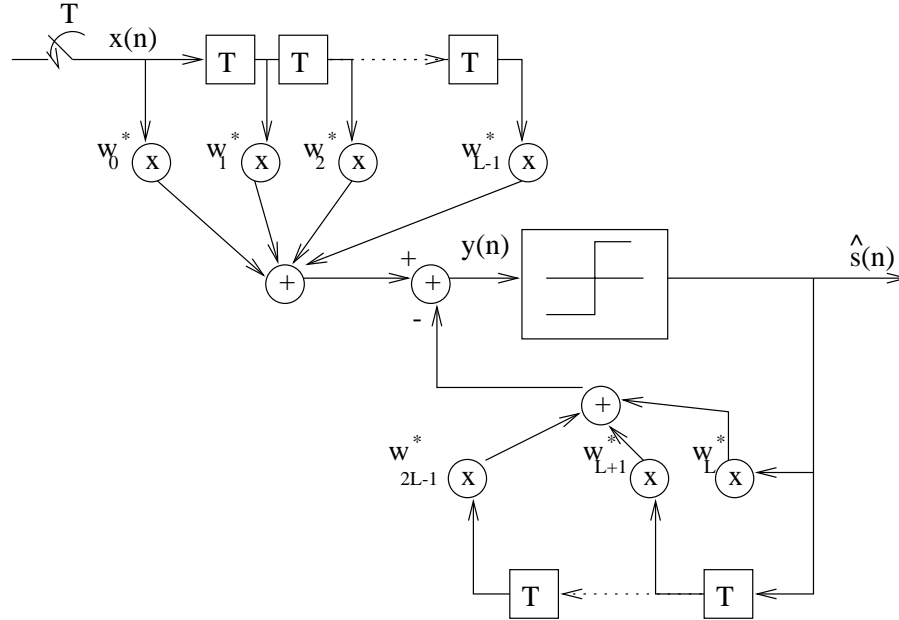


Figura 2.2: Diagrama de bloques de un receptor monousuario con igualador con realimentación de decisiones (DFE)

Cuando un igualador es adaptativo, sus coeficientes se ajustan automáticamente con un algoritmo recursivo de forma que puedan compensar el efecto del canal sin conocerlo. Las técnicas tradicionales de igualación usan habitualmente algoritmos supervisados que se basan en la minimización del error cuadrático medio MSE (Mean Square Error) entre la salida del igualador y los símbolos transmitidos. Es práctica habitual que el cálculo de los coeficientes óptimos del igualador se haga de forma recursiva mediante el denominado algoritmo LMS (Least Mean Square), que es un algoritmo de gradiente estocástico asociado al criterio MMSE (Minimum MSE). La principal limitación del algoritmo LMS es que su implementación obliga a la transmisión de secuencias de entrenamiento que deben ser conocidas a priori por el receptor [56] cosa que puede no ser posible o deseable en numerosas aplicaciones prácticas. En estos casos es preferible utilizar algoritmos ciegos que permiten ajustar los coeficientes del igualador utilizando exclusivamente las propiedades estadísticas de la señal de salida. De entre los muchos y muy diversos algoritmos ciegos que han sido propuestos, el más ampliamente conocido y utilizado en igualación es el CMA (Constant Modulus Algorithm) debido a su sencillez de implementación. Así por ejemplo, el módem de un radioenlace de microondas *MegaStarTM 2000* de *AT&T/Harris* [66] utiliza en recepción un igualador fraccionario cuyos coeficientes son inicialmente ajustados por medio del algoritmo CMA hasta obtener una sustancial reducción de la ISI, pasando

después a ser guiado por decisión, DD.

Una de las limitaciones de los igualadores CMA es que sus coeficientes se ajustan para minimizar una función de coste que involucra estadísticos de orden superior y que puede contener mínimos no deseados. Si bien para igualadores de longitud infinita se ha podido demostrar que la función de coste CMA no tiene mínimos locales [25], estudios más recientes que consideran igualadores de longitud finita han mostrado [41] que aparecen mínimos locales cuando el canal a igualar sigue un modelo auto-regresivo AR (Auto-Regressive).

Muchos canales prácticos se pueden aproximar bastante bien por un modelo AR. La figura 2.3 dibuja la respuesta al impulso correspondiente a un radioenlace de microondas [62]. Puede comprobarse fácilmente como se puede aproximar bastante bien por una exponencial unilateral, ejemplo paradigmático de respuesta al impulso de un sistema AR.

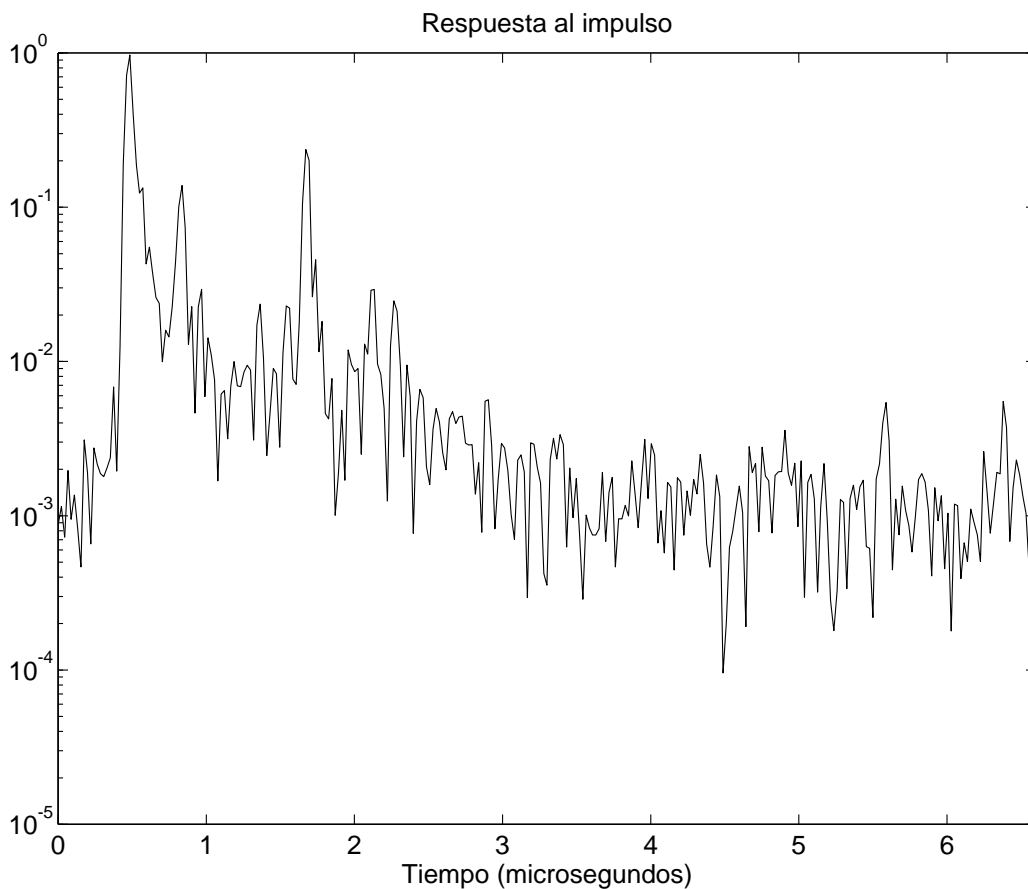


Figura 2.3: Respuesta al impulso de un canal de microondas.

El objetivo de este capítulo es proponer una sencilla modificación de la función de coste CMA de tal forma que se eliminen los mínimos locales que surgen cuando se igualan canales AR. Presentamos un análisis para una situación muy sencilla de un canal AR de orden 1 y un igualador de dos coeficientes. Este análisis nos permitirá poner claramente de manifiesto las ventajas de nuestra aproximación. Posteriormente extenderemos nuestras ideas a canales AR de mayor orden. También se presenta un análisis teniendo en cuenta el ruido gaussiano del canal. Aunque este análisis está limitado a situaciones en las que la SNR es alta, ilustra claramente el óptimo comportamiento del igualador propuesto. Finalmente, se presentan los resultados de varias simulaciones por ordenador llevadas a cabo para mostrar el comportamiento del método propuesto y así validar las aproximaciones realizadas en el análisis.

2.2 Modelo de Señal

La figura 2.4 representa el diagrama de bloques del modelo equivalente discreto de un sistema de transmisión de datos en banda base convencional. Vamos a considerar que la señal transmitida, $s(n)$, es una secuencia compleja correspondiente a una constelación QAM (Quadrature Amplitude Modulation) de media cero y con símbolos estadísticamente independientes e idénticamente distribuidos (i.i.d.). De acuerdo con la clasificación establecida en el capítulo anterior, un igualador lineal es un *front-end* de comunicaciones que suministra diversidad temporal por medio de una línea de retardos. La señal $x(n)$ existente en los distintos puntos de la línea de retardos puede escribirse como

$$x(n) = \sum_{k=0}^{\infty} h(k)s(n-k) + r(n) \quad (2.1)$$

donde $h(n)$ es la respuesta al impulso del canal y $r(n)$ es el ruido aditivo. El ruido lo modelamos como un proceso aleatorio gaussiano de media cero, complejo, con densidad espectral de potencia constante $\frac{N_0}{2}$ y estadísticamente independiente de los símbolos transmitidos.

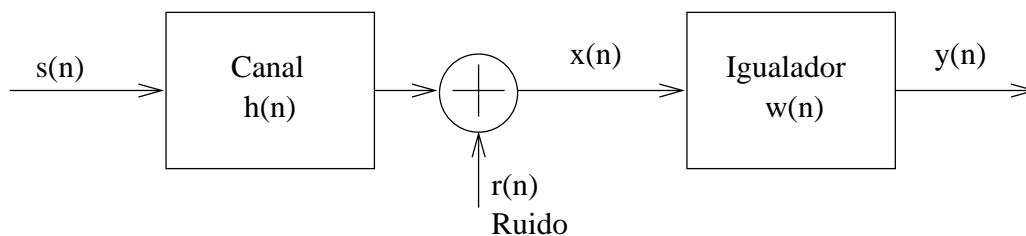


Figura 2.4: Diagrama de bloques de un sistema de igualación de canal.

Reescribiendo (2.1), la señal recibida se puede descomponer de la forma

$$x(n) = \underbrace{h(l)s(n-l)}_{\text{Deseado}} + \underbrace{\sum_{\substack{k=0 \\ k \neq l}}^{\infty} h(k)s(n-k)}_{\text{ISI}} + \underbrace{r(n)}_{\text{Ruido}} \quad (2.2)$$

En general, el canal no tiene por qué ser de fase mínima y el máximo de su respuesta al impulso ocurrirá en un instante que llamaremos l , $h(l) = \max_n h(n)$. Esto significa que el símbolo $s(n-l)$ es el que se recibe con mayor amplitud y, por lo tanto, lo consideraremos como el símbolo deseado. Como se deduce de (2.2), el símbolo deseado se ve perturbado por la presencia del ruido, $r(n)$, y de una componente que contiene contribuciones de los símbolos transmitidos anterior y posteriormente. Esta componente, que hemos denominado ISI, es distinta de cero cuando el canal no es ideal, es decir, cuando $h(n) \neq \delta(n-l)$. En transmisiones digitales como las realizadas a través de canales telefónicos, la componente de ISI es muy superior a la del ruido y constituye su principal limitación.

Para utilizar una notación más adecuada, vamos a agrupar la señal de entrada en un vector de datos de longitud L de la forma

$$\mathbf{x}(n) = [x(n), \dots, x(L-1)]^T \quad (2.3)$$

De la ecuación (2.1) se obtiene que este vector puede ser expresado por medio del siguiente producto

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{r}(n) \quad (2.4)$$

donde $\mathbf{s}(n) = [s(n), s(n-1), \dots]^T$ es el vector de símbolos de entrada, $\mathbf{r}(n) = [r(n), \dots, r(n-L+1)]^T$ es el vector de ruido y \mathbf{H} es una matriz de convolución de dimensión $L \times \infty$ correspondiente al canal

$$\mathbf{H} = \begin{bmatrix} h(0) & h(1) & \dots & h(L-1) & \dots \\ 0 & h(0) & \dots & h(L-2) & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & h(0) & \dots \end{bmatrix}_{L \times \infty} \quad (2.5)$$

La figura 2.1 representa el diagrama de bloques de un igualador lineal síncrono convencional que, recordemos, consiste en un filtro transversal con una línea de retardos separados entre sí por un intervalo T igual al período de símbolo. La salida (entrada al circuito decisor) es una combinación lineal de las señales proporcionadas por la línea de retardo y, por tanto, se puede escribir como

$$y(n) = \sum_{i=0}^{L-1} w_i^* x(n-i) \quad (2.6)$$

donde w_i son los L coeficientes complejos del igualador, $x(n)$ es la secuencia presente en la línea de retardos y $*$ denota complejo conjugado. Usando notación vectorial, $y(n)$ se puede escribir de una forma más compacta a través del producto escalar

$$y(n) = \mathbf{w}^H \mathbf{x}(n) \quad (2.7)$$

donde $\mathbf{w} = [w_0, \dots, w_{L-1}]^T$ es el vector de coeficientes del igualador. Los superíndices T y H denotan transpuesto y transpuesto conjugado, respectivamente.

Sustituyendo (2.4) en (2.7), la salida del igualador se puede expresar como

$$y(n) = \mathbf{w}^H \mathbf{H} \mathbf{s}(n) + \mathbf{w}^H \mathbf{r}(n) = \mathbf{g}^H \mathbf{s}(n) + \mathbf{w}^H \mathbf{r}(n) \quad (2.8)$$

donde

$$\mathbf{g} = [g_0, g_1, \dots]^T = \mathbf{H}^H \mathbf{w} \quad (2.9)$$

es un vector que representa la respuesta impulsional conjunta del canal y el igualador. Cada elemento g_i se puede interpretar como la amplitud que tiene el símbolo $s(n-i)$ a la salida del igualador.

2.3 Igualadores Zero-Forcing

Debido a que las componentes del vector de respuesta conjunta del canal e igualador, \mathbf{g} , son las amplitudes de diferentes símbolos a la salida, parece lógico pensar que el igualador que converja hacia una solución con una única componente de \mathbf{g} distinta de cero obtendrá a su salida, en ausencia de ruido ($\mathbf{r}(n) = \mathbf{0}$), un único símbolo eliminando por completo la ISI. Una posible solución será de la forma

$$\mathbf{g}_{opt}^l = [0, \dots, 0, g_l, 0, \dots, 0]^T \quad (2.10)$$

donde el índice l representa el retardo introducido por el efecto conjunto del canal y el igualador. Los igualadores que persiguen una solución de este estilo con $g_l = 1$ son los denominados *zero-forcing* [42]. Un algoritmo adaptativo que converja hacia la solución (2.10) ajusta sus coeficientes de la siguiente forma [27]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu e(n) \mathbf{s}(n) \quad (2.11)$$

donde μ es un parámetro que controla la velocidad de convergencia del algoritmo y $e(n) = y(n) - s(n-l)$ representa el error entre la salida del igualador y el símbolo deseado. Se puede comprobar que la correlación entre el error y los símbolos diferentes al l -ésimo es de la forma

$$E[e(n)s(n-m)] = \begin{cases} g_l - E[s^2(n)] & m = l \\ g_m & m \neq l \end{cases} \quad (2.12)$$

donde $E[\cdot]$ representa el operador esperanza. Cuando la potencia de la señal transmitida es normalizada a la unidad, entonces (2.12) es una medida de lo cerca que el vector de amplitudes, \mathbf{g} , está cerca de la solución (2.10).

La implementación de este algoritmo es muy sencilla y, por ello, adecuada para aplicaciones de alta velocidad de transmisión (comunicaciones por enlaces de microondas o por fibra óptica). Sin embargo, para implementar el algoritmo (2.11) es necesario disponer de los símbolos actualmente transmitidos. Como esto no es posible, este tipo de algoritmos utilizan un período inicial en el que se transmiten unas secuencias conocidas por el transmisor y el receptor denominadas de entrenamiento. Durante esta fase inicial el receptor ajusta los coeficientes de acuerdo con (2.11) de forma que se alcance una buena inicialización. Una vez superada la fase de entrenamiento, las salidas del circuito decisor, presente a la salida del igualador, serán las que sustituyan en (2.11) a los símbolos transmitidos y se dice que el algoritmo está guiado por decisión, DD (Decision Directed).

La respuesta frecuencial de un igualador de *zero-forcing* trata de ajustarse a la inversa del canal [56]. Esto puede tener grandes problemas cuando en el canal existe ruido, pues ante la presencia de un nulo en la respuesta del canal el igualador tratará de compensarlo amplificando el ruido presente en esta frecuencia. La no consideración del ruido limita las aplicaciones de este tipo de igualadores dando lugar a otros que tratarán de compensar la amplificación del ruido y la eliminación de la ISI.

2.4 Igualadores MMSE

En ausencia de ruido, el igualador de *zero-forcing* elimina completamente la ISI. Sin embargo, desde el punto de vista de minimizar la probabilidad de error puede ser conveniente permitir cierta ISI residual en la entrada del decisor si esto permite dotar al igualador de un mayor grado de libertad que le permita alcanzar una sustancial reducción de la potencia de ruido. Un igualador que minimice la probabilidad de error es demasiado complejo [1] y, por ello, se prefieren soluciones que obtengan buenas aproximaciones de esta minimización. Un ejemplo son los igualadores que tienen como objetivo la minimización del error cuadrático medio (MSE, Mean Square Error), entre la salida del igualador y el símbolo deseado, es decir,

$$MSE = E[(y(n) - s(n-l))^2] = E[e(n)^2] \quad (2.13)$$

Estos se denominan igualadores MMSE (Minimum Mean Square Error) [33]. El algoritmo adaptativo más utilizado por los igualadores MMSE es el LMS (Least Mean Square) [67]. Este algoritmo actualiza los coeficientes del igualador de la siguiente forma

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) \quad (2.14)$$

donde la señal de error es la misma que la del algoritmo adaptativo de *zero-forcing*, diferenciándose este algoritmo con respecto a (2.11) en que la señal que multiplica al error es la secuencia observada, $\mathbf{x}(n)$, en lugar de la transmitida, $\mathbf{s}(n)$. El error del algoritmo ha de ser calculado por medio de la secuencia transmitida. Sin embargo, ésta no estará disponible y, como ocurre con los igualadores *zero-forcing*, en su lugar se utilizan secuencias de entrenamiento que permiten ajustar inicialmente los coeficientes del igualador. Al igual que en la sección anterior, cuando finaliza la fase de entrenamiento el igualador pasa a ser guiado por decisión, DD.

En aplicaciones prácticas donde el transmisor opere independientemente del receptor (por ejemplo, en sistemas de multidifusión como la televisión digital) o donde los canales varíen rápidamente con el tiempo (por ejemplo, en comunicaciones móviles), un mecanismo que necesite secuencias de entrenamiento puede no ser posible (debido a la independencia entre el receptor y el transmisor) o deseable (debido a la reducción del *throughput* del sistema al intercalar secuencias de entrenamiento). Todas estas limitaciones han motivado el desarrollo de igualadores ciegos.

2.5 Igualadores Ciegos

Los igualadores descritos en las secciones anteriores se caracterizan por la necesidad de conocer los símbolos transmitidos. Como ya se ha comentado, esto puede suponer una gran limitación debido a la cooperación que el transmisor y el receptor han de mantener. Para superar esta limitación, se han desarrollado los denominados igualadores ciegos que permiten ajustar sus coeficientes a partir únicamente de los datos recibidos, sin necesidad de secuencias de entrenamiento.

El funcionamiento de los igualadores ciegos se basa en las propiedades estadísticas de la señal de salida. El problema de la igualación ciega ha sido objeto de estudio por varios autores dando lugar a múltiples algoritmos que, dependiendo de la información utilizada para ajustar los coeficientes, pueden agruparse en dos grandes familias [33].

Basados en estadísticos de orden superior Este tipo de algoritmos, denominados HOS (Higher Order Statistics), pueden ser divididos a su vez en dos categorías:

Algoritmos tipo Bussgang En estos la señal de salida del filtro ha de cumplir la condición de *Bussgang* [33]. Su estructura consiste en un filtro lineal seguido de algún dispositivo no lineal sin memoria.

Basados en Cumulantes o en Polyespectro Estos algoritmos utilizan los *cumulantes* [48] de orden superior o su transformada de Fourier conocida como *polyespectro*.

Basados en estadísticos cicloestacionarios Utilizan la información obtenida de la periodicidad de los momentos de las señales involucradas en la comunicación [26]. Estos algoritmos son utilizados cuando la amplitud, fase o frecuencia de una portadora senoidal es alterada de acuerdo con la información que se desea transmitir.

Debido a que en el resto del capítulo se presenta un igualador basado en la familia de los algoritmos tipo *Bussgang*, se describen estos más detalladamente.

Algoritmos tipo Bussgang

En este tipo de algoritmos la señal filtrada es obtenida de un estimador no lineal sin memoria que involucra estadísticos de orden superior de la señal recibida. Estos logran la convergencia cuando cumplen la denominada condición de Bussgang². Esta familia de algoritmos incluye al conocido algoritmo de Sato [58] que define una nueva función de coste cuya minimización se alcanza por medio de la adaptación de los coeficientes del filtro. El algoritmo que regula el ajuste de los coeficientes es muy similar al LMS (2.14) pero con una señal de error, $e(n)$, diferente a la utilizada para el criterio MMSE.

Del algoritmo de Sato han surgido muchas variantes que modifican su función de coste. Ejemplos de este tipo de algoritmos son los propuestos por Benveniste [7], Picchi [54] y Shynk [60], que combinan las funciones de coste del algoritmo de Sato y del MMSE.

Godard [29] generalizó el algoritmo de Sato en otra familia de algoritmos que, como demuestran los estudios comparativos presentados en [35] y [60], son los algoritmos tipo Bussgang más eficaces. Además, tal y como apunta Papadias en [49], este algoritmo es el más robusto con respecto a los errores de recuperación de fase, el que obtiene el menor MSE cuando el algoritmo ha convergido y el que permite eliminar la ISI cuando ésta es más severa. Este capítulo se basa en un caso particular de los algoritmos de Godard que son descritos a continuación.

Algoritmo de Godard

Godard [29] consideró un algoritmo ciego para modulaciones QAM (Quadrature Amplitude Modulation). De forma intuitiva, buscó una función de coste que fuese independiente de la fase de la señal obtenida a la salida del filtro para poder alcanzar la convergencia del algoritmo de forma independiente a la fase de adquisición de

²Un proceso discreto $y(n)$ se dice de Bussgang si satisface la condición $E[y(n)y(n+k)] = E[y(n)g(y(n+k))]$, donde $g(\cdot)$ es una función no lineal y sin memoria.

portadora. Para ello se define la siguiente función de coste

$$J = E[(|y(n)|^p - R_p)^2] \quad (2.15)$$

donde p es un entero y $R_p = \frac{E[|s(n)|^{2p}]}{E[|s(n)|^p]^2}$ es una constante de dispersión que controla la amplificación del filtro. Esta función persigue la extracción de una señal con módulo constante a la salida del filtro. La presencia de interferencias y/o ruido superpuestas a la señal transmitida provocará que en recepción la señal no posea la propiedad de módulo constante. Al minimizar esta función de coste, es decir, el error de la potencia p -ésima del módulo de la señal de salida y la constante de dispersión, cabe esperar que se recupere la propiedad de módulo constante eliminando las interferencias y el ruido que provocaba oscilaciones en el módulo de la señal.

Godard analizó el algoritmo para $p = 1$ y $p = 2$. Para $p = 2$ el algoritmo posee una sencilla implementación y se conoce con el nombre de CMA (Constant Modulus Algorithm). Este algoritmo fue empleado por otros autores para eliminar las interferencias y la ISI introducida por canales con multitrayecto [65].

Godard demostró que el mínimo absoluto de la función de coste (2.15) para CMA ($p = 2$) corresponde con la eliminación de la ISI introducida por un canal dispersivo siempre y cuando la kurtosis³ de los símbolos transmitidos sea negativa. Sin embargo, esto no significa que no existan otros mínimos locales. Foschini [25] completó el trabajo de Godard demostrando que si la kurtosis de los símbolos transmitidos es negativa la función (2.15) no posee más mínimos que los absolutos. La condición de que la señal transmitida ha de poseer una kurtosis negativa es muy interesante pues extiende el uso del algoritmo CMA a señales que no necesariamente han de tener estrictamente módulo constante y, por lo tanto, este algoritmo puede ser usado en prácticamente todas las señales utilizadas en comunicaciones digitales. Como ejemplo, la tabla 2.1 muestra la kurtosis de señales que usan diferentes modulaciones digitales.

Considerando un análisis similar al de Foschini para señales reales, la salida de un igualador síncrono de longitud doblemente infinita combina las muestras de la señal de entrada de la siguiente forma

$$y(n) = \sum_{l=-\infty}^{\infty} w_l x(n-l) = \mathbf{w}^T \mathbf{x}(n) \quad (2.16)$$

donde $\mathbf{w} = [w_{-\infty}, \dots, w_{-1}, w_0, w_1, \dots, w_{\infty}]^T$ es un vector de longitud infinita que contiene los coeficientes del igualador y $\mathbf{x}(n) = [\dots, x(n+1), x(n), x(n-1), \dots]^T$ es un vector que contiene las muestras de la señal de entrada. Teniendo en cuenta (2.4), la matriz de convolución del canal que obtiene el vector de entrada, $\mathbf{x}(n)$, es de la

³La kurtosis de una señal real se define como $\frac{E[|s(n)|^4]}{E^2[|s(n)|^2]} - 3$, y la de una señal compleja circularmente invariante como $\frac{E[|s(n)|^4]}{E^2[|s(n)|^2]} - 2$.

Modulación	Kurtosis
2-PAM	-2
4-PAM	-0.36
8-PAM	-0.2381
4-QAM	-2
16-QAM	-0.68
64-QAM	-0.6191
M-PSK	-2

Tabla 2.1: Kurtosis de algunas de las modulaciones empleadas en comunicaciones digitales.

forma

$$\mathbf{H} = \begin{bmatrix} h(0) & h(1) & \cdots & h(M-1) & 0 & 0 & \cdots & \cdots \\ 0 & h(0) & h(1) & \cdots & h(M-1) & 0 & 0 & \cdots \\ 0 & 0 & h(0) & \ddots & \cdots & h(M-1) & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}_{\infty \times \infty} \quad (2.17)$$

Considerando en (2.8) que no existe ruido, podemos reescribir la salida $y(n)$ de la siguiente forma

$$y(n) = \mathbf{g}^T \mathbf{s}(n) \quad (2.18)$$

donde $\mathbf{g} = \mathbf{H}^T \mathbf{w}$ es la respuesta al impulso conjunta del canal y el igualador.

Expandiendo los términos de la función de coste (2.15) para CMA ($p = 2$) y normalizando respecto a m^2 ($m = E[s^2(n)]$), obtenemos la siguiente expresión

$$J = \frac{1}{m^2} E[(y^2(n) - R_2)^2] = \frac{E[y^4(n)]}{m^2} - 2k \frac{E[y^2(n)]}{m} + k^2 \quad (2.19)$$

donde $k = \frac{E[s^4(n)]}{E^2[s^2(n)]}$ es el momento de cuarto orden normalizado de $s(n)$. Teniendo en cuenta (2.18) y que los símbolos $s(n)$ son independientes e idénticamente distribuidos (hipótesis i.i.d.), las esperanzas en (2.19) resultan ser

$$\begin{aligned} \frac{E[y^2(n)]}{m} &= \frac{E[\mathbf{g}^T \mathbf{s}(n) \mathbf{s}^T(n) \mathbf{g}]}{m} = \frac{\mathbf{g}^T E[\mathbf{s}(n) \mathbf{s}^T(n)] \mathbf{g}}{m} = \\ &= \mathbf{g}^T \mathbf{g} = \sum_{i=0}^{\infty} g_i^2 \\ \frac{E[y^4(n)]}{m^2} &= \frac{E[(\mathbf{g}^T \mathbf{s}(n))^4]}{m^2} = \frac{E[(\sum_{i=0}^{\infty} g_i s(n-i))^4]}{m^2} = \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{m^2} E \left[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{u=0}^{\infty} \sum_{l=0}^{\infty} g_i g_j g_u g_l s(n-i) s(n-j) s(n-u) s(n-l) \right] \\
&= \begin{cases} i = j = u = l & \Rightarrow \sum_{i=0}^{\infty} g_i^4 k \\ \left\{ \begin{array}{ll} i = j & u = l \\ i = u & j = l \\ i = l & u = j \end{array} \right\} & \Rightarrow 3 \sum_{j=0}^{\infty} \sum_{\substack{l=0 \\ l \neq j}}^{\infty} g_j^2 g_l^2 \end{cases} \\
&= k \sum_{i=0}^{\infty} g_i^4 + 3 \sum_{j=0}^{\infty} \sum_{\substack{l=0 \\ j \neq l}}^{\infty} g_j^2 g_l^2 \tag{2.20}
\end{aligned}$$

Sustituyendo (2.20) en (2.19) se obtiene

$$J = k \sum_{i=-\infty}^{\infty} g_i^4 + 3 \sum_{j=-\infty}^{\infty} \sum_{\substack{l=-\infty \\ j \neq l}}^{\infty} g_j^2 g_l^2 - 2k \sum_{i=-\infty}^{\infty} g_i^2 + k^2 \tag{2.21}$$

El siguiente paso es calcular los puntos estacionarios de la función de coste (2.21), que son aquellos en los que el gradiente de J con respecto a \mathbf{w} es igual a cero. Utilizando la regla de la cadena [63], este gradiente puede expresarse en términos del gradiente de J con respecto a \mathbf{g} , $\nabla_{\mathbf{g}} J$, de la siguiente forma

$$\nabla_{\mathbf{w}} J = \mathbf{H} \nabla_{\mathbf{g}} J \tag{2.22}$$

Es obvio que

$$\nabla_{\mathbf{g}} J = \mathbf{0} \Rightarrow \nabla_{\mathbf{w}} J = \mathbf{0} \tag{2.23}$$

Sin embargo, para que la relación inversa sea cierta, debe cumplirse que todas las columnas de la matriz \mathbf{H} sean linealmente independientes, o lo que es lo mismo, que el subespacio nulo de \mathbf{H} sea igual al trivial

$$\nabla_{\mathbf{w}} J = \mathbf{0} \Rightarrow \nabla_{\mathbf{g}} J = \mathbf{0} \tag{2.24}$$

El estudio de las condiciones bajo las cuales la matriz de convolución \mathbf{H} tiene un subespacio nulo igual al trivial ha sido abordado por Ding [17] demostrando que si el igualador es doblemente infinito, y la matriz del canal es de la forma (2.17), es posible asegurar que los puntos estacionarios de J están contenidos en las soluciones del sistema de ecuaciones $\nabla_{\mathbf{g}} J = \mathbf{0}$.

En el apéndice B se realiza el análisis de los puntos estacionarios de J con respecto a \mathbf{g} demostrándose que los únicos mínimos de esta función se corresponden con los que eliminan completamente la ISI si y sólo si la kurtosis de las señales transmitidas es negativa.

2.5.1 Mínimos Locales Dependientes de la Longitud

Los análisis realizados por Foschini [25] y Godard [29] consideran situaciones poco prácticas con igualadores de longitud doblemente infinita, es decir, $L = \infty$. Sin embargo, cuando se consideran igualadores de longitud finita, los igualadores ciegos, en particular los basados en el CMA, presentan mínimos locales no deseados denominados LDLM (*Length Dependent Local Minima*) [41, 40, 18, 17].

En particular cuando el canal de comunicaciones se puede representar por un modelo auto-regresivo de un solo polo (AR(1)), el filtro IIR que lo representa tiene una función de transferencia

$$H(z) = \frac{1}{1 - az^{-1}} \quad (2.25)$$

y una respuesta al impulso

$$h(n) = a^n u(n) \quad 0 < a < 1 \quad (2.26)$$

donde $u(n)$ es la función escalón unitario y a es una constante real. La matriz de convolución (2.5) asociada a este canal y un igualador de longitud finita, L , es

$$\mathbf{H} = \begin{bmatrix} 1 & a & \cdots & a^{L-1} & a^L & \cdots \\ 0 & 1 & \cdots & a^{L-2} & a^{L-1} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & \cdots & 1 & a & \cdots \end{bmatrix}_{L \times \infty} \quad (2.27)$$

Obsérvese que \mathbf{H} contiene columnas linealmente dependientes. Si \mathbf{h}_i es la i -ésima columna de \mathbf{H} , entonces $\mathbf{h}_i = a^{i-L} \mathbf{h}_L$ para $i \geq L$. En este caso, no se puede garantizar la relación (2.24) y esta dependencia lineal obtiene la siguiente relación entre las ganancias a la salida del igualador

$$g_i = a^{i-L+1} g_{L-1} \quad \forall i \geq L - 1 \quad (2.28)$$

Por ello, un igualador que converja a una solución en la que $g_i \neq 0$ para algún $i \geq L - 1$ obtendrá $g_i \neq 0 \quad \forall i \geq L - 1$ y su salida consistirá en una combinación lineal de los $s(n - i) \quad \forall i \geq L - 1$ impidiendo una completa eliminación de la ISI. Este tipo de soluciones corresponden con los LDLM. Sin embargo, si el igualador obtiene $g_i \neq 0$ para $i < L - 1$, el vector \mathbf{g} adopta la forma (2.10) y se consigue una eliminación completa de la ISI, extrayéndose el símbolo $s(n - i)$.

2.6 Modificación de los Igualadores CMA

Una de las principales contribuciones de este trabajo consiste en proponer una sencilla modificación de la función de coste CMA (2.15) que no posee LDLM cuando los

canales siguen un modelo AR. Esta modificación, presentada por primera vez en [21], consiste en añadir un término que penalice la convergencia a estos mínimos no deseados resultando una función de coste de la forma

$$J = J_{CMA} + \alpha J_{penaliz} \quad (2.29)$$

El nuevo término $J_{penaliz}$ es multiplicado por una constante α que le da un peso respecto a la función de coste CMA. Hemos escogido este término como el valor absoluto al cuadrado de la correlación entre la señal de salida del igualador $y(n)$ y la señal correspondiente al último elemento del vector de datos observados $x(n-L+1)$

$$J_{penaliz} = |E[y(n)x^*(n-L+1)]|^2 \quad (2.30)$$

por lo que la función de coste resultante es

$$J = E[(|y(n)|^2 - R_2)^2] + \alpha |E[y(n)x^*(n-L+1)]|^2 \quad (2.31)$$

Aunque en un principio la introducción de (2.30) puede parecer un tanto heurística, la razón es la siguiente. Supongamos que no existe ruido y que el igualador alcanza un punto de equilibrio donde $g_i \neq 0$, $i < L-1$. Este será un mínimo deseado en el que la salida del igualador $y(n) = s(n-i)$. De (2.1) se puede comprobar que el último elemento del vector de datos observados es

$$x(n-L+1) = \sum_{k=0}^{\infty} h(k)s(n-L+1-k) \quad (2.32)$$

que sólo depende de los símbolos transmitidos $s(n-j) \quad \forall j > L-1$. Estos símbolos son anteriores a $s(n-i)$ para $i < L-1$ y estadísticamente independientes de éste. Por lo tanto, la salida $y(n) = s(n-i)$ está incorrelada con el elemento del vector de observación $x(n-L+1)$, el término de penalización será igual a cero y no afectará a la convergencia del igualador hacia los mínimos deseados donde la ISI sea eliminada.

Supongamos ahora que el igualador converge a uno de los puntos denominados LDLM donde $g_i \neq 0$ para $i \geq L-1$ en cuyo caso la salida que se obtiene es

$$y(n) = \sum_{i=L-1}^{\infty} g_i^* s(n-i) \quad (2.33)$$

Como se había mencionado, esta salida es una combinación lineal de los símbolos $s(n-i) \quad \forall i \geq L-1$. De la ecuación (2.32) se obtiene que en el elemento del vector de entrada $x(n-L+1)$ también existe una combinación lineal de los mismos símbolos y por lo tanto existirá una correlación entre la salida $y(n)$ y $x(n-L+1)$. En este caso, el nuevo término $J_{penaliz}$ es no nulo y penalizará la convergencia del algoritmo hacia este tipo de soluciones.

Es interesante observar como la función (2.31) puede interpretarse como el Lagrangiano asociado al problema de optimización

$$\min_{\mathbf{w}} E[(|y(n)|^2 - R_2)^2] \quad \text{sujeeto a} \quad |E[y(n)x^*(n-L+1)]|^2 = 0 \quad (2.34)$$

La ventaja de la aproximación presentada en este capítulo es que se puede optimizar mediante algoritmos adaptativos muy sencillos. La resolución del problema con restricciones obliga a plantear un algoritmo en dos etapas: una que minimice J_{CMA} y otra que imponga la restricción.

2.6.1 Algoritmo Adaptativo

La minimización de la función de coste (2.29) puede llevarse a cabo de una forma sencilla utilizando un algoritmo de tipo gradiente estocástico donde el vector \mathbf{w} es actualizado de acuerdo a la siguiente fórmula recursiva

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \nabla_{\mathbf{w}} J(n) \quad (2.35)$$

donde μ es el parámetro que controla la velocidad de convergencia del algoritmo y $\nabla_{\mathbf{w}} J(n)$ es el vector gradiente de J con respecto a \mathbf{w} en el punto $\mathbf{w}(n)$. Considerando (2.7), el vector gradiente para cada componente de J es de la forma

$$\begin{aligned} \nabla_{\mathbf{w}} J_{CMA}(n) &= \nabla_{\mathbf{w}} E[(\mathbf{w}^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w} - R_2)(\mathbf{w}^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w} - R_2)] = \\ &= 2E[(\mathbf{w}^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w} - R_2)(2\mathbf{x}^H(n) \mathbf{w} \mathbf{x}(n))] = \\ &= 4E[(|y(n)|^2 - R_2)y^*(n)\mathbf{x}(n)] \\ \nabla_{\mathbf{w}} J_{penaliz}(n) &= \nabla_{\mathbf{w}} [E[\mathbf{x}^H(n) \mathbf{w} x(n-L+1)]E[\mathbf{w}^H \mathbf{x}(n) x^*(n-L+1)]] = \\ &= 2E[\mathbf{x}^H(n) \mathbf{w} x(n-L+1)]E[\mathbf{x}(n) x^*(n-L+1)] = \\ &= 2E[y^*(n)x(n-L+1)]E[\mathbf{x}(n) x^*(n-L+1)] \end{aligned} \quad (2.36)$$

resultando el vector gradiente de J de la siguiente forma

$$\nabla_{\mathbf{w}} J(n) = 2(2E[(|y(n)|^2 - R_2)y^*(n)\mathbf{x}(n)] + \alpha E[y^*(n)x(n-L+1)]E[\mathbf{x}(n) x^*(n-L+1)]) \quad (2.37)$$

En la práctica, las esperanzas estadísticas de (2.37) son desconocidas y el algoritmo adaptativo se implementa con estimaciones obtenidas a partir de la salida $y(n)$ y la entrada $\mathbf{x}(n)$. Así pues, el algoritmo de gradiente estocástico que ajusta los coeficientes del igualador es

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) - \mu \left\{ \sum_{i=n-1}^n (|y(i)|^2 - R_2)y^*(i)\mathbf{x}(i) \right. \\ &\quad \left. + \frac{\alpha}{4} \sum_{j=n-1}^n y^*(j)x(j-l+1) \sum_{k=n-1}^n \mathbf{x}(k)x^*(k-l+1) \right\} \end{aligned} \quad (2.38)$$

donde se han usado dos muestras para estimar las esperanzas estadísticas de (2.37) y un factor de 2 ha sido introducido en el parámetro μ . Con este número de muestras en los estimadores eliminamos resultados ambiguos que aparecen cuando se trata con estadísticos de orden superior y se usan estimadores con una única muestra⁴.

2.6.2 Generalización a Canales AR de Mayor Orden

En la sección anterior, se ha presentado y justificado la modificación de los igualadores suponiendo canales auto-regresivos de primer orden, AR(1). La consideración de canales de mayor orden, AR(r) es más compleja. Considerando un igualador de L coeficientes y un canal AR de r polos, AR(r), con función de transferencia

$$H(z) = \frac{1}{1 - \sum_{i=1}^r a_i z^{-i}} \quad (2.39)$$

la respuesta al impulso de este canal es de la forma

$$h(n) = \begin{cases} 0 & n < 0 \\ 1 & n = 0 \\ \sum_{i=1}^r a_i h(n-i) & n \geq 1 \end{cases} \quad (2.40)$$

Es sencillo comprobar que la matriz convolutiva del canal, \mathbf{H} , es

$$\mathbf{H} = \begin{bmatrix} h(0) & h(1) & h(2) & \cdots & h(L-1) & h(L) & \cdots \\ 0 & h(0) & h(1) & \cdots & h(L-2) & h(L-1) & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & \cdots & h(0) & h(1) & \cdots \end{bmatrix}_{L \times \infty}. \quad (2.41)$$

De (2.40) es fácil observar que cuando el número de coeficientes es mayor que el de polos del canal⁵, $L > r$, las columnas $L+1$ en adelante son linealmente dependientes de las columnas $L, L-1, \dots, L-r+1$.

Debido a las dependencias lineales en las columnas de la matriz \mathbf{H} , las dependencias de las ganancias del igualador para canales AR(1) se pueden generalizar de la siguiente forma

$$g_i = \sum_{j=1}^r \beta_{ij} g_{L-j} \quad \forall i \geq L \quad (2.42)$$

⁴Por ejemplo, $|y(n)|^4$ es un estimador de una muestra para $E[|y(n)|^4]$ y $E^2[|y(n)|^2]$.

⁵Esta condición es necesaria para garantizar que existen soluciones donde es posible eliminar completamente la ISI.

donde β_{ij} son constantes que dependen de los a_i . Obsérvese que una solución $g_j \neq 0$ para cualquier $L - r \leq j \leq L - 1$ no elimina completamente la ISI. El número de LDLM en un igualador de L coeficientes en un canal AR(r) es r y, por lo tanto, el número de mínimos deseados es $L - r$.

La extensión de la función (2.31) para igualar canales AR(r) es inmediata y únicamente necesita añadir r términos similares a $J_{penaliz}$ en (2.30). El i -ésimo término añadido contendrá la correlación entre la salida del igualador y el elemento del vector de observaciones $(L - i)$ -ésimo. La función de coste resultante es

$$J_r = E[(|y(n)|^2 - R_2)^2] + \alpha \sum_{l=L-r}^{L-1} |E[y(n)x^*(n-l)]|^2 \quad (2.43)$$

Esta inclusión de r términos es necesaria para penalizar la convergencia del algoritmo hacia soluciones en las que se obtengan combinaciones lineales de símbolos. El único inconveniente de la función de coste generalizada es la necesidad de conocer el orden del canal r . Sin embargo, esto no constituye una limitación importante pues una sobreestimación de este orden nos seguirá conduciendo hacia una solución correcta en la que se elimina la ISI totalmente.

De forma completamente análoga a la subsección 2.6.1 es posible derivar un algoritmo de gradiente estocástico que minimice (2.43). El resultado que se obtiene es

$$\begin{aligned} \mathbf{w}(n+1) = \mathbf{w}(n) - \mu \left\{ \sum_{i=n-1}^n (|y(i)|^2 - R_2) y^*(i) \mathbf{x}(i) \right. \\ \left. + \frac{\alpha}{4} \sum_{l=L-r}^{L-1} \sum_{j=n-1}^n y^*(j) x(j-L+l) \sum_{k=n-1}^n \mathbf{x}(k) x^*(k-L+l) \right\} \end{aligned} \quad (2.44)$$

donde de nuevo se han utilizado estimadores de dos muestras en lugar de las esperanzas estadísticas.

Con anterioridad a nuestro trabajo, otros muchos autores han propuesto diversas técnicas para superar las limitaciones de los igualadores CMA. Chen [13, 14] introduce una modificación del igualador convencional CMA con el objeto inicial de mejorar su velocidad de convergencia. Este algoritmo, que es denominado CRIMNO (CRiterion with Memory Nonlinearity), modifica la función de coste del CMA introduciendo un conjunto de términos consistente en la correlación cruzada entre diferentes instantes de la señal de salida de la siguiente forma

$$J_{CRIMNO} = \alpha_0 E[(|y(n)|^2 - R_2)^2] + \sum_{l=1}^r \alpha_l |E[y(n)y^*(n-l)]|^2 \quad (2.45)$$

Aunque no se menciona en [14], un razonamiento similar al expuesto anteriormente podría inducir a pensar que los términos añadidos en CRIMNO también tienen la

capacidad de eliminar los LDLM ya que su valor no se anula en dichos puntos y sí lo hace en los mínimos deseados. Esta afirmación, sin embargo, no ha podido probarse hasta la fecha. A diferencia de nuestra aproximación, un análisis de la función de coste CRIMNO es extremadamente difícil de llevar a cabo y, por el momento, sólo se ha ilustrado su comportamiento mediante simulaciones por ordenador.

Una aproximación muy diferente a la eliminación de los LDLM es la propuesta por Papadias y Slock [51] quienes sugieren la utilización de un algoritmo CMA normalizado (NCMA, Normalized Constant Modulus Algorithm). Realmente la función de coste NCMA no está libre de LDLM. Lo que se demuestra en [51] es que el algoritmo que minimiza NCMA admite un mayor ruido de desajuste sin volverse divergente lo que permite al algoritmo escapar de los LDLM y converger hacia los mínimos deseados. En cierto modo esta técnica recuerda al “simulated annealing” utilizado en el aprendizaje de redes neuronales [32].

Zeng [70] utiliza las propiedades cicloestacionarias de la señal para obtener un nuevo algoritmo denominado CM-CF (Constant Modulus - Correlation Fitting). Este algoritmo presenta un problema de optimización con múltiples objetivos y que ha de realizarse en varias etapas lo que dificulta el análisis de su convergencia. Por el momento su funcionamiento sólo se ha ilustrado mediante simulaciones realizadas por ordenador.

Finalmente, Zazo [69] considera la imposición de restricciones lineales para obtener un nuevo algoritmo denominado MCMA (Modified Constant Modulus Algorithm). La restricción lineal considerada consiste en fijar uno de los coeficientes del igualador para impedir que el algoritmo adaptativo converja hacia los mínimos LDLM en los que este coeficiente alcanza un valor nulo.

2.7 Análisis de los Puntos Estacionarios

El problema de optimización (2.43) se basa en la minimización de una superficie de error que no es una forma cuadrática del vector de coeficientes y que, por lo tanto, puede poseer varios mínimos locales (como es el caso de los Length-Dependent Local Minima). En esta sección se va a llevar a cabo un análisis de los puntos estacionarios de esta superficie de error para comprobar que la nueva función de coste no contiene este tipo de mínimos locales. El análisis consiste en el cálculo de los puntos estacionarios, que es donde el gradiente se hace igual a cero, y en el examen de la naturaleza de la matriz hessiana en estos puntos para determinar si son mínimos, máximos o puntos de ensilladura.

2.7.1 Análisis sin Ruido

Por razones de simplicidad, y debido a la complejidad de J , el análisis se realiza para un canal auto-regresivo de primer orden, AR(1), libre de ruido, $\mathbf{r}(n) = \mathbf{0}$ y un igualador de dos coeficientes. También asumiremos que los símbolos transmitidos son una secuencia real con una modulación binaria PAM (Pulse Amplitude Modulation) y que los coeficientes del igualador son reales.

Comencemos normalizando (2.29) respecto a m^2 y desarrollándola como sigue

$$\frac{J}{m^2} = \frac{1}{m^2}E[y^4(n)] - 2\frac{k}{m}E[y^2(n)] + k^2 + \frac{\alpha}{m^2}(E[y(n)x^*(n-1)])^2 \quad (2.46)$$

donde $m = E[s^2(n)]$ es el momento de segundo orden de los símbolos transmitidos y $k = \frac{E[s^4(n)]}{E^2[s^2(n)]}$ representa el momento de cuarto orden normalizado de $s(n)$. Considerando un igualador de 2 coeficientes ($L = 2$), la relación (2.28) se convierte en $g_i = a^{i-1}g_1 \quad \forall i \geq 1$ y, como se demuestra en el apéndice C, (2.46) puede ser expresada como sigue

$$J(g_0, g_1) = k[g_0^4 + g_1^4 S_2] + 6g_1^2[g_0^2 S_1 + g_1^2 S_3] - 2k[g_0^2 + g_1^2 S_1] + k^2 + \alpha[g_1^2 S_1^2] \quad (2.47)$$

donde hemos expresado J como función de g_0 y g_1 , los dos primeros elementos del vector de respuesta conjunta del canal y el igualador. Estos dos parámetros representan las amplitudes que los símbolos $s(n)$ y $s(n-1)$ tienen a la salida del igualador. Las constantes S_1, S_2 y S_3 son las sumas de unas series que dependen de la respuesta impulsional del canal como sigue

$$\begin{aligned} S_1 &= \sum_{i=1}^{\infty} (a^{i-1})^2 = \frac{1}{1-a^2} \\ S_2 &= \sum_{i=1}^{\infty} (a^{i-1})^4 = \frac{1}{1-a^4} \\ S_3 &= \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} (a^{i-1})^2 (a^{j-1})^2 = \frac{a^2}{(1+a^2)(1-a^2)^2} \end{aligned} \quad (2.48)$$

Para analizar cuales son los puntos estacionarios de J es necesario calcular los puntos donde el gradiente $\nabla_{\mathbf{w}} J$ se hace cero. Usando la regla de la cadena [63] este gradiente se puede representar de la siguiente forma

$$\nabla_{\mathbf{w}} J = \sum_{i=0}^{\infty} \frac{\partial J}{\partial g_i} \nabla_{\mathbf{w}} g_i \quad (2.49)$$

Para nuestro caso de un igualador de 2 coeficientes y un canal AR(1) se cumple la relación $\nabla_{\mathbf{w}} g_i = a^{i-1} \nabla_{\mathbf{w}} g_1 \quad \forall j \geq 1$ y el gradiente (2.49) puede expresarse como

$$\nabla_{\mathbf{w}} J = \frac{\partial J}{\partial g_0} \nabla_{\mathbf{w}} g_0 + \frac{\partial J}{\partial g_1} \nabla_{\mathbf{w}} g_1 \sum_{i=1}^{\infty} a^{i-1} \quad (2.50)$$

Debido a que $\mathbf{g} = \mathbf{H}^T \mathbf{w}$ (donde \mathbf{H} es la matriz convolucional del canal (2.27)) es evidente que

$$\nabla_{\mathbf{w}} g_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \nabla_{\mathbf{w}} g_1 = \begin{bmatrix} a \\ 1 \end{bmatrix} \quad (2.51)$$

Estos dos vectores son linealmente independientes y, por lo tanto, todos los puntos estacionarios de J se corresponden con las soluciones del siguiente sistema de ecuaciones

$$\begin{aligned} \frac{\partial J}{\partial g_0} &= 4g_0[(g_0^2 - 1)k + 3g_1^2 S_1] = 0 \\ \frac{\partial J}{\partial g_1} &= 4g_1[(g_1^2 S_2 - S_1)k + 3(g_0^2 S_1 + 2g_1^2 S_3)] + 2\alpha g_1 S_1^2 = 0 \end{aligned} \quad (2.52)$$

Para comprobar si un punto estacionario de J es un mínimo o no hay que examinar si la matriz hessiana es definida positiva o no. Teniendo en cuenta la independencia lineal entre $\nabla_{\mathbf{w}} g_0$ y $\nabla_{\mathbf{w}} g_1$ es suficiente con examinar el carácter definido positivo de la siguiente matriz

$$\mathcal{H} = \begin{pmatrix} \frac{\partial^2 J}{\partial g_0^2} & \frac{\partial^2 J}{\partial g_0 \partial g_1} \\ \frac{\partial^2 J}{\partial g_1 \partial g_0} & \frac{\partial^2 J}{\partial g_1^2} \end{pmatrix} \quad (2.53)$$

cuyos elementos se obtienen a partir de (2.52) como sigue

$$\begin{aligned} \frac{\partial^2 J}{\partial g_0^2} &= 12[g_0^2 k + g_1^2 S_1] - 4k \\ \frac{\partial^2 J}{\partial g_0 \partial g_1} &= \frac{\partial^2 J}{\partial g_1 \partial g_0} = 24g_0 g_1 S_1 \\ \frac{\partial^2 J}{\partial g_1^2} &= 12g_1^2[kS_2 + 6S_3] + 12g_0^2 S_1 - 4kS_1 + 2\alpha S_1^2 \end{aligned} \quad (2.54)$$

El sistema de ecuaciones de (2.52) es no lineal y su resolución da lugar a múltiples soluciones que se clasificarán en cuatro grupos:

Grupo 1 : $g_0 \neq 0$ y $g_1 = 0$. Es un punto estacionario deseado donde se alcanza una igualación perfecta. La salida coincide perfectamente con el símbolo $s(n)$. Resolviendo (2.52) se obtiene

$$g_0 = 1 \quad g_1 = 0 \quad (2.55)$$

La matriz hessiana correspondiente a este punto se convierte en una matriz diagonal de la forma

$$\mathcal{H} = \begin{pmatrix} 8k & 0 \\ 0 & 4S_1(3 - k) + 2\alpha \end{pmatrix} \quad (2.56)$$

El primer elemento de la diagonal es siempre positivo⁶. El segundo elemento de

⁶El momento de cuarto orden normalizado, k , es siempre una cantidad positiva.

la diagonal es positivo si $\alpha \geq 0$ y $k - 3 < 0$. Por lo tanto, si α posee un valor positivo y la kurtosis normalizada⁷ de los símbolos transmitidos es negativa, que es lo habitual en todas las señales empleadas en comunicaciones digitales, el hessiano es siempre definido positivo y este punto es un mínimo de J .

Grupo 2 : $g_0 = g_1 = 0$. Es un punto estacionario no deseado donde el igualador obtiene una salida nula y ningún símbolo es extraído. La matriz hessiana en este punto es una matriz diagonal de la forma

$$\mathcal{H} = \begin{pmatrix} -4k & 0 \\ 0 & -4kS_1 + 2\alpha S_1^2 \end{pmatrix} \quad (2.57)$$

El primer elemento de la diagonal es siempre negativo y por lo tanto \mathcal{H} no puede ser definida positiva. Por ello, este punto no es un mínimo de J .

Grupo 3 : $g_0 = 0$ y $g_1 \neq 0$. Es un punto estacionario no deseado donde el igualador extrae la solución correspondiente a los LDLM. En estos puntos la ISI no es eliminada ya que $g_i \neq 0 \forall i \geq 2$. Resolviendo el sistema de ecuaciones (2.52) para este caso, la solución es

$$g_0 = 0 \quad g_1^2 = \frac{S_1(2k - \alpha S_1)}{2(kS_2 + 6S_3)} \quad (2.58)$$

Para representar la matriz hessiana en este punto, podemos observar la siguiente relación entre (2.54) y (2.52)

$$g_1 \frac{\partial^2 J}{\partial g_1 \partial g_1} = \frac{\partial J}{\partial g_1} + 8g_1^3[kS_2 + 6S_3] \quad (2.59)$$

y por lo tanto la matriz hessiana, \mathcal{H} resulta ser diagonal con la siguiente forma

$$\mathcal{H} = \begin{pmatrix} 4(3S_1g_1^2 - k) & 0 \\ 0 & 8g_1^2(kS_2 + 6S_3) \end{pmatrix} \quad (2.60)$$

Debido a que $k > 0$ y $S_i > 0$, es evidente que el segundo elemento de la diagonal es siempre positivo. Para asegurar que la matriz hessiana no es definida positiva, y por lo tanto este punto no es un mínimo de J , es necesario asegurar que el primer elemento de la diagonal sea negativo. Sustituyendo el valor de g_1^2 obtenido en (2.58) obtenemos que el primer elemento de la diagonal de \mathcal{H} vale

$$4\left(3S_1 \frac{S_1(2k - \alpha S_1)}{2(kS_2 + 6S_3)} - k\right) = \frac{2}{kS_2 + 6S_3} (k(6S_1^2 - 12S_3 - 2kS_2) - 3\alpha S_1^3) \quad (2.61)$$

⁷Definimos la kurtosis normalizada de una señal real como $k - 3$.

Es evidente que, para que (2.61) sea negativo, es necesario que el término entre paréntesis sea negativo y por lo tanto

$$\alpha > \frac{k}{3S_1^3}(6S_1^2 - 12S_3 - 2kS_2) \quad (2.62)$$

Sustituyendo el valor de las series (2.48) obtenemos la siguiente condición

$$\alpha > \frac{k(1 - a^2)^2(6 - 2k)}{3(1 + a^2)} \quad (2.63)$$

Teniendo en cuenta que $a < 1$ podemos obtener un límite de α más conservador e independiente del parámetro del canal

$$\alpha > \frac{2k(3 - k)}{3} \quad (2.64)$$

De esta forma, si garantizamos un valor de α que cumpla la condición (2.64), aseguraremos que el hessiano es no definido positivo y por lo tanto este punto no será un mínimo de J . Obsérvese que cuando se usa la función de coste original de CMA, es decir $\alpha = 0$, la condición (2.64) no se cumple y por lo tanto este punto estacionario se corresponde con un mínimo. Este tipo de mínimos son los denominados *LDLM*.

Grupo 4 : $g_0 \neq 0$ y $g_1 \neq 0$. Este es otro punto estacionario no deseado donde se produce la extracción de una combinación lineal de símbolos transmitidos.

Para reducir la complejidad de la matriz hessiana en este punto, podemos encontrar las siguientes relaciones entre (2.54) y (2.52)

$$\begin{aligned} g_0 \frac{\partial^2 J}{\partial g_0 \partial g_0} &= \frac{\partial J}{\partial g_0} + 8g_0^3 k \\ g_1 \frac{\partial^2 J}{\partial g_1 \partial g_1} &= \frac{\partial J}{\partial g_1} + 8g_1^3 [kS_2 + 6S_3] \end{aligned} \quad (2.65)$$

resultando una matriz hessiana de la forma

$$\mathcal{H} = \begin{pmatrix} 8g_0^2 k & 24g_0 g_1 S_1 \\ 24g_0 g_1 S_1 & 8g_1^2 (kS_2 + 6S_3) \end{pmatrix} \quad (2.66)$$

Es inmediato comprobar que el primer término de la diagonal es siempre positivo, por lo que la naturaleza de \mathcal{H} está determinada por el signo de su determinante.

$$\det(\mathcal{H}) = 64g_0^2 g_1^2 [k(kS_2 + 6S_3) - 9S_1^2] \quad (2.67)$$

Para asegurar que este punto no es un mínimo de J , es necesario garantizar que (2.67) sea negativo, es decir, que el término entre paréntesis, que denominaremos $F(k, a)$, sea negativo

$$F(k, a) \triangleq [k(kS_2 + 6S_3) - 9S_1^2] < 0 \quad (2.68)$$

Sustituyendo (2.48) en (2.68) obtenemos

$$F(k, a) = \frac{k^2(1 - a^2) + 6ka^2 - 9(1 + a^2)}{(1 - a^4)(1 - a^2)} \quad (2.69)$$

que, como muestra la figura 2.5, es una función parabólica convexa en k cuyas raíces son $k = -3(1 + a^2)/(1 - a^2)$ y $k = 3$. Debido a que $k > 0$ y $a^2 < 1$, podemos asegurar que $F(k, a)$ tiene un valor negativo si y sólo si $k - 3 < 0$.

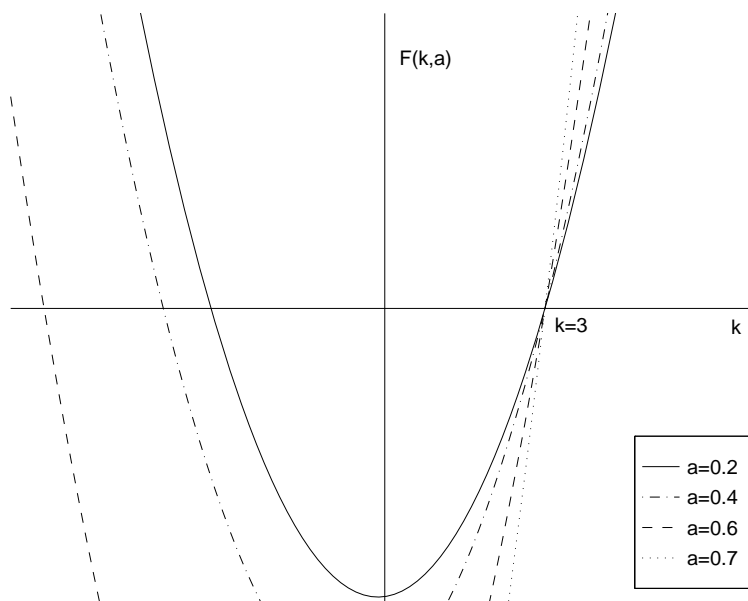


Figura 2.5: Representación de la función (2.69).

Por lo tanto, si la kurtosis de los símbolos transmitidos es negativa, podemos concluir que el hessiano no es definido positivo y, por lo tanto, este punto estacionario no es un mínimo de J . Obsérvese que esta condición es la misma que asegura que el punto estacionario deseado, donde se produce una igualación perfecta, sea un mínimo.

Del análisis de los puntos estacionarios que se acaba de presentar, y que se puede encontrar también en [20], se puede concluir que la nueva función de coste (2.31) no

posee los puntos de equilibrio no deseados presentes en CMA. La no existencia de estos puntos depende del parámetro α . Si este parámetro es demasiado pequeño, el algoritmo todavía posee estos puntos de equilibrio no deseados. Cuando α es superior a la cota (2.64) el igualador elimina la posibilidad de convergencia hacia este tipo de puntos.

La figura 2.6 representa las funciones de coste de un igualador de 2 coeficientes correspondientes al CMA convencional ($\alpha = 0$) y a la modificación propuesta ($\alpha = 2$). El canal considerado es AR(1) con un parámetro $a = 0.5$ y los símbolos transmitidos corresponden a una constelación PAM de dos niveles, $s(n) = \pm 1$ ($k = 1$). En la figura de la izquierda se observa la existencia de cuatro mínimos, dos deseados y dos no deseados (LDLM). En cambio, la figura de la derecha sólo posee los dos mínimos deseados. Puede observarse como los dos mínimos no deseados han desaparecido porque el valor de α escogido satisface la condición (2.64).

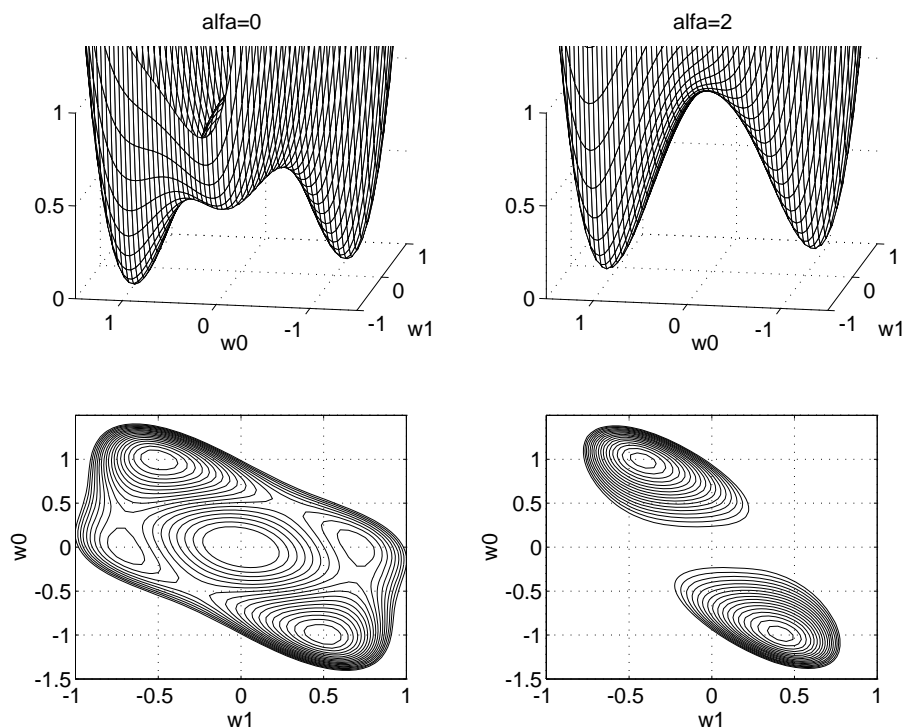


Figura 2.6: Representación de las funciones de coste correspondientes al nuevo algoritmo y al CMA para un igualador de dos coeficientes y un canal AR(1) con $a = 0.5$.

Un análisis para un canal AR de orden r , AR(r), es extremadamente difícil de realizar. Sin embargo, el límite de α obtenido en (2.64) es todavía válido para la función de coste general (2.43). La validez de este resultado es debido a que los conjuntos de dependencias lineales en las columnas de \mathbf{H} para un canal AR(r) son idénticas a

las de un canal AR(1) y, por ello, los LDLM de ambos canales poseen características similares. Simulaciones realizadas con canales AR(r) muestran la validez de esta conjetura.

2.7.2 Análisis con Ruido

El análisis de los puntos de equilibrio presentado en el apartado anterior se hizo suponiendo un canal libre de ruido. Un análisis similar para un canal con ruido aditivo es extremadamente difícil y no ha sido posible hacerlo. No obstante, sí es posible obtener algunos resultados aproximados si suponemos que la relación señal a ruido (SNR) es suficientemente alta. Bajo esta hipótesis podemos considerar que la función de coste J cuando hay ruido es una perturbación de cuando no lo hay en cuyo caso podemos suponer que el número y naturaleza de los puntos de equilibrio son los mismos. En presencia de ruido mostraremos que los puntos de equilibrio deseados son una versión perturbada de los mínimos obtenidos en la sección anterior. En esta sección se calcula la perturbación que el ruido introduce sobre los mínimos deseados y se estudia su efecto sobre el comportamiento del igualador. Este análisis también está recogido en [22].

Considerando que los símbolos transmitidos y los coeficientes del igualador son reales, la función de coste (2.31) normalizada respecto a m^2 para un igualador de L coeficientes es

$$\frac{J}{m^2} = \frac{1}{m^2} E[y^4(n)] - 2\frac{k}{m} E[y^2(n)] + k^2 + \frac{\alpha}{m^2} E[y(n)x(n-L+1)] \quad (2.70)$$

Al considerar la presencia de ruido en el canal y que la SNR es suficientemente alta, en el apéndice C se demuestra que (2.70) se puede escribir como

$$\frac{\tilde{J}}{m^2} \simeq \frac{J}{m^2} + 2\gamma[3\mathbf{w}^T \mathbf{w} (\sum_{i=0}^{L-2} g_i^2 + g_{L-1}^2 S_1) - k\mathbf{w}^T \mathbf{w} + \alpha(g_{L-1} S_1)w_{L-1}] \quad (2.71)$$

donde \tilde{J} y J representa la nueva función de coste con y sin ruido, respectivamente, y γ es la inversa de la relación señal a ruido.

Esta expresión de la función de coste depende de \mathbf{g} y de \mathbf{w} . A continuación vamos a expresarla en función únicamente de \mathbf{g} . Usando (2.9) podemos obtener que

$$\mathbf{w} = \mathbf{H}^\dagger \mathbf{g} \quad (2.72)$$

donde \mathbf{H}^\dagger representa la pseudoinversa de \mathbf{H} que satisface $\mathbf{H}\mathbf{H}^\dagger = \mathbf{I}$ (\mathbf{I} es la matriz identidad). Para una matriz \mathbf{H} de la forma (2.27), se puede comprobar que la

pseudoinversa resultante es la siguiente matriz de dimensión $\infty \times L$

$$\mathbf{H}^\dagger = \begin{pmatrix} 1 & -a & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -a & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & -a \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}_{\infty \times L} \quad (2.73)$$

Considerando (2.72) y $m = 1$ la función de coste (2.71) es igual a

$$\tilde{J} \simeq J + 2\gamma[3(\sum_{i=0}^{L-2} g_i^2 + g_{L-1}^2 S_1) \mathbf{g}^T \mathbf{D} \mathbf{g} - k \mathbf{g}^T \mathbf{D} \mathbf{g} + \alpha(g_{L-1} S_1)(g_{L-1} - a g_{L-2})] \quad (2.74)$$

donde

$$\mathbf{D} = \mathbf{H}^\dagger \mathbf{H}^{\dagger T} = \begin{pmatrix} \Phi_{L \times L} & \mathbf{0}_{L \times \infty} \\ \mathbf{0}_{\infty \times L} & \mathbf{0}_{\infty \times \infty} \end{pmatrix} \quad (2.75)$$

y

$$\Phi_{L \times L} = \begin{pmatrix} (1+a^2) & -a & 0 & 0 & 0 \\ -a & (1+a^2) & -a & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & 0 & -a & (1+a^2) & -a \\ 0 & 0 & 0 & -a & 1 \end{pmatrix} \quad (2.76)$$

Para la obtención de (2.74) hemos usado la relación $w_{L-1} = g_{L-1} - a g_{L-2}$ obtenida de (2.72) y de la forma de la matriz \mathbf{H}^\dagger (2.73).

Como ya vimos en el análisis sin ruido, los puntos estacionarios de \tilde{J}_{MCA} se corresponden con los puntos donde el gradiente se hace cero

$$\begin{aligned} \frac{\partial \tilde{J}}{\partial g_i} &= \frac{\partial J}{\partial g_i} + \\ &+ \gamma[12g_i \mathbf{g}^T \mathbf{D} \mathbf{g} + 12(\sum_{j=0}^{L-2} g_j^2 + g_{L-1}^2 S_1) \mathbf{D}_{(i+1,-)} \mathbf{g} - 4k \mathbf{D}_{(i+1,-)} \mathbf{g} - \\ &- 2a\alpha g_{L-1} S_1 \delta(i-L+2)] \quad i = 0, \dots, L-2 \\ \frac{\partial \tilde{J}}{\partial g_{L-1}} &= \frac{\partial J}{\partial g_{L-1}} + \\ &+ \gamma[12g_{L-1} S_1 \mathbf{g}^T \mathbf{D} \mathbf{g} + 12(\sum_{j=0}^{L-2} g_j^2 + g_{L-1}^2 S_1) \mathbf{D}_{(L,-)} \mathbf{g} - 4k \mathbf{D}_{(L,-)} \mathbf{g} + \\ &+ 2\alpha(S_1(g_{L-1} - a g_{L-2}) + g_{L-1} S_1)] \end{aligned} \quad (2.77)$$

donde $\delta(n)$ es el impulso unitario y $\mathbf{D}_{(i,-)}$ representa la i -ésima fila de la matriz \mathbf{D} . Además, el gradiente de la función de coste en una situación libre de ruido es

$$\begin{aligned}\frac{\partial J}{\partial g_i} &= 4(k-3)g_i^3 + 12\left(\sum_{j=0}^{L-2} g_j^2 + g_{L-1}^2 S_1\right)g_i - 4kg_i \quad i = 0, \dots, L-2 \\ \frac{\partial J}{\partial g_{L-1}} &= 4(k-3)g_{L-1}^3 S_2 + 12\left(\sum_{j=0}^{L-2} g_j^2 + g_{L-1}^2 S_1\right)g_{L-1} S_1 - 4kg_{L-1} S_1 + 2\alpha g_{L-1} S_1^2\end{aligned}\tag{2.78}$$

Generalizando el resultado de la sección anterior, es obvio que el gradiente (2.78) en situaciones sin ruido se hace cero en los mínimos

$$\mathbf{g}_{opt}^l = [\dots, 0, \underbrace{1}_l, 0, \dots]^T \quad l < L-1\tag{2.79}$$

También existen otros puntos en los que el gradiente (2.78) se anula, pero ya demostramos en la sección anterior que no son mínimos de la función de coste y por lo tanto ya no los vamos a considerar en ésta.

Supongamos que $\tilde{\mathbf{g}}^l$ es un mínimo de \tilde{J} . A continuación mostramos que para valores altos de la SNR, $\tilde{\mathbf{g}}^l$ puede ser aproximado por una versión perturbada de \mathbf{g}_{opt}^l de la siguiente forma

$$\tilde{\mathbf{g}}^l \simeq \mathbf{g}_{opt}^l + \gamma \Delta \mathbf{g} = [\dots, \gamma \Delta g_{l-1}, 1 + \gamma \Delta g_l, \gamma \Delta g_{l+1}, \dots]^T\tag{2.80}$$

Sustituyendo esta expresión en (2.77) se obtiene, como se demuestra en el apéndice C, el siguiente sistema de ecuaciones

$$\begin{aligned}\left. \frac{\partial \tilde{J}}{\partial \tilde{g}_i} \right|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} &\simeq \gamma[\Delta g_i(12-4k) + (12-4k)\mathbf{D}_{(i+1,l+1)}] = 0 \quad i < L-1 \text{ e } i \neq l \\ \left. \frac{\partial \tilde{J}}{\partial \tilde{g}_l} \right|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} &\simeq \gamma[\Delta g_l 8k + (24-4k)\mathbf{D}_{(l+1,l+1)}] = 0 \\ \left. \frac{\partial \tilde{J}}{\partial \tilde{g}_{L-1}} \right|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} &\simeq \gamma[\Delta g_{L-1} S_1(12-4k+2\alpha S_1) + (12-4k)\mathbf{D}_{(L,l+1)}] \\ &\quad - 2a\alpha S_1 g_{L-2} = 0\end{aligned}\tag{2.81}$$

donde $\mathbf{D}_{(i,j)}$ es el elemento (i,j) de la matriz \mathbf{D} . Finalmente, es inmediato comprobar que la solución de este sistema de ecuaciones es

$$\begin{aligned}\Delta g_i &= -\mathbf{D}_{(i+1,l+1)} \quad i < L-1 \text{ e } i \neq l \\ \Delta g_l &= \frac{(k-6)(1+a^2)}{2k} \\ \Delta g_{L-1} &= \delta(l-L+2) \frac{a}{S_1} = \delta(l-L+2)(a(1-a^2))\end{aligned}\tag{2.82}$$

Por lo tanto, los mínimos de la función de coste sufren una perturbación $\Delta \mathbf{g}$ debida al efecto del ruido. Este resultado es similar a los obtenidos en [41] y [24] para igualadores CMA. Esta perturbación tiene un doble efecto ya que se produce un cambio del valor de la l -ésima componente de \mathbf{g}_{opt}^l , es decir, de la amplitud con la que el igualador extrae $s(n-l)$ y, por otro lado, aparece una pequeña cantidad de ISI debido a que otras componentes de \mathbf{g}_{opt}^l diferentes de la l -ésima son distintas de cero. Esta cantidad de ISI es debida a que el igualador necesita llegar a una solución de compromiso en la cual se ha de buscar un equilibrio entre la minimización del ruido y de la ISI. Para probar el buen comportamiento del igualador a continuación se va a demostrar como el MSE en los mínimos de \tilde{J} es el mínimo que se puede alcanzar. Esto prueba que el igualador se comporta igual que un igualador MMSE cuando la SNR es alta.

La figura 2.7 representa el contorno de la nueva función de coste para $\alpha = 2$ y la compara con el del CMA convencional. Se ha considerado para esta representación una relación señal a ruido $SNR = 10$ dB, un igualador de dos coeficientes y un canal AR de primer orden ($h(n) = (0.5)^n u(n)$). Obsérvese que esta función es muy similar a la de la figura 2.6 donde se representa la función de coste en ausencia de ruido y para el mismo canal. El número de puntos estacionarios es idéntico en ambos casos, y su naturaleza permanece inalterada por la presencia del ruido. Las diferencias entre ambas figura radican en la posición de los mínimos que, como se puede apreciar, han sido movidos debido a la perturbación producida por el ruido.

2.7.3 Valor del MSE en los Mínimos con Ruido

El error cuadrático medio, normalizado respecto a m , entre la salida del igualador y un símbolo con un retardo l , se define como

$$\frac{MSE}{m} = \frac{E[(y(n) - s(n-l))^2]}{m} \quad (2.83)$$

Cuando estamos en una situación sin ruido en que $y(n) = \mathbf{g}^T \mathbf{s}(n)$ el MSE (para $m=1$) es

$$MSE = \mathbf{g}^T \mathbf{g} - 2g_l + 1 = (\mathbf{g} - \mathbf{g}_{opt}^l)^T (\mathbf{g} - \mathbf{g}_{opt}^l) \quad (2.84)$$

donde \mathbf{g}_{opt}^l es el valor de \mathbf{g} (2.79) para el que se minimiza la función de coste que proponemos en ausencia de ruido. Obsérvese como el MSE se hace cero en este punto, y como la modificación del CMA que proponemos minimiza el MSE.

En una situación con ruido la salida del igualador viene determinada por (2.8), es decir $y(n) = y_s(n) + y_r(n) = \mathbf{g}^T \mathbf{s}(n) + \mathbf{w}^T \mathbf{r}(n)$, y el MSE (2.84) se puede expresar como

$$M\tilde{S}E = MSE + \gamma \mathbf{w}^T \mathbf{w} = (\mathbf{g} - \mathbf{g}_{opt}^l)^T (\mathbf{g} - \mathbf{g}_{opt}^l) + \gamma \mathbf{g}^T \mathbf{D} \mathbf{g} \quad (2.85)$$

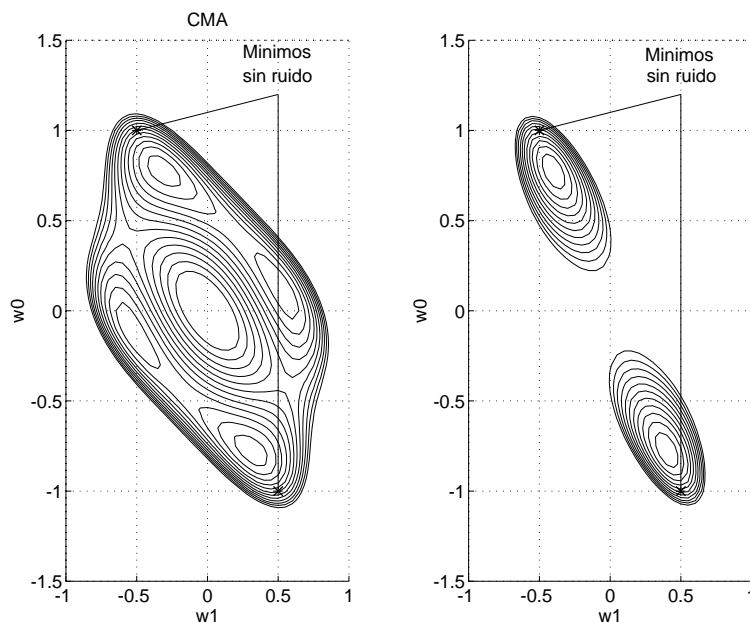


Figura 2.7: Representación del contorno de la nueva función de coste y de la de CMA para un igualador de dos coeficientes y un canal AR(1) con $a = 0.5$ y ruido aditivo con una $SNR = 10$ dB.

donde se ha usado la relación (2.72). Para obtener el punto donde $M\tilde{S}E$ alcanza su mínimo, es necesario calcular el punto donde el gradiente se hace cero, es decir

$$\nabla_{\mathbf{g}} M\tilde{S}E = 2(\mathbf{g} - \mathbf{g}_{opt}^l) + 2\gamma \mathbf{D}\mathbf{g} = 0 \quad (2.86)$$

Supondremos que la solución de esta ecuación es un punto de la forma

$$\tilde{\mathbf{g}}_{MSE}^l = \mathbf{g}_{opt}^l + \gamma \Delta \mathbf{g}_{MSE}, \quad (2.87)$$

Sustituyendo (2.87) en (2.86) y considerando de nuevo una SNR suficientemente alta, el gradiente se hace cero cuando

$$\begin{aligned} 2[\gamma \Delta \mathbf{g}_{MSE} + \gamma \mathbf{D}(\mathbf{g}_{opt}^l + \gamma \Delta \mathbf{g}_{MSE})] &\simeq \\ &\simeq 2\gamma[\Delta \mathbf{g}_{MSE} + \mathbf{D}\mathbf{g}_{opt}^l] = 2\gamma[\Delta \mathbf{g}_{MSE} + \mathbf{D}_{(-,l+1)}] = 0 \end{aligned} \quad (2.88)$$

y por lo tanto

$$\Delta \mathbf{g}_{MSE} = -\mathbf{D}_{(-,l+1)} \quad (2.89)$$

Como se puede observar, el mínimo obtenido es muy similar al obtenido en la minimización de la función de coste J en una situación con ruido (2.80), ya que el mínimo del MSE se alcanza para un valor $\tilde{\mathbf{g}}_{MSE}^l$ que es una versión perturbada del valor \mathbf{g}_{opt}^l obtenido en una situación sin ruido.

Sustituyendo $\tilde{\mathbf{g}}_{MSE}^l$ en la expresión (2.85), obtendremos la expresión del valor mínimo del MSE (MMSE)

$$\begin{aligned} MMSE &= M\tilde{S}E \Big|_{\mathbf{g}=\tilde{\mathbf{g}}_{MSE}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}_{MSE}} = \\ &= \gamma^2(\Delta\mathbf{g}_{MSE})^T(\Delta\mathbf{g}_{MSE}) + \gamma(\mathbf{g}_{opt}^l + \gamma\Delta\mathbf{g}_{MSE})^T\mathbf{D}(\mathbf{g}_{opt}^l + \gamma\Delta\mathbf{g}_{MSE}) \simeq \\ &\simeq \gamma(\mathbf{g}_{opt}^l)^T\mathbf{D}\mathbf{g}_{opt}^l = \gamma\mathbf{D}_{(l+1,l+1)} = \gamma(1+a^2) \end{aligned} \quad (2.90)$$

donde hemos despreciado los términos multiplicados por γ^2 y γ^3 debido a la suposición de una SNR suficientemente grande. Obsérvese que este valor del MMSE no depende del valor de la perturbación debida al ruido. Por ello el mismo valor será obtenido para el MSE en los mínimos de nuestro igualador

$$MSE \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} \simeq \gamma(\mathbf{g}_{opt}^l)^T\mathbf{D}\mathbf{g}_{opt}^l = MMSE \quad (2.91)$$

Este resultado demuestra que en presencia de ruido el igualador que estamos proponiendo se comporta de la misma forma que el igualador MMSE, obteniéndose un equilibrio entre la reducción del ruido y la eliminación de la ISI. Obsérvese también que el MSE alcanzado en los mínimos de J no depende del parámetro α y que por lo tanto el término de penalización incluido en la función de coste (2.29) no afecta al comportamiento global del igualador; este término lo único que hace es eliminar los puntos de equilibrio no deseados.

La figura 2.8 representa los valores del MSE correspondientes al criterio MMSE, al alcanzado en la minimización de J y a la expresión aproximada del MMSE dada por (2.90) para diferentes valores de la SNR. Excepto la curva correspondiente a (2.90), las curvas han sido obtenidos mediante simulaciones por ordenador. Se ha considerado un canal AR de primer orden con la siguiente función de transferencia

$$H(z) = \frac{1}{1 - 0.6z^{-1}} \quad (2.92)$$

y un igualador de 2 coeficientes. Como se puede comprobar en esta figura, la curva que representa esta última expresión es idéntica a las experimentales cuando los valores de la SNR están por encima de 10 dB.

La figura 2.9 dibuja las curvas de la probabilidad de error para un entorno idéntico al del experimento anterior. Como se puede observar, en este caso las curvas de ambos criterios son prácticamente idénticas para todos los valores de la SNR mostrando el correcto comportamiento del nuevo algoritmo.

2.8 Simulaciones

Esta sección presenta los resultados de varias simulaciones por ordenador que apoyan los resultados analíticos obtenidos en las secciones anteriores. En estas simulaciones

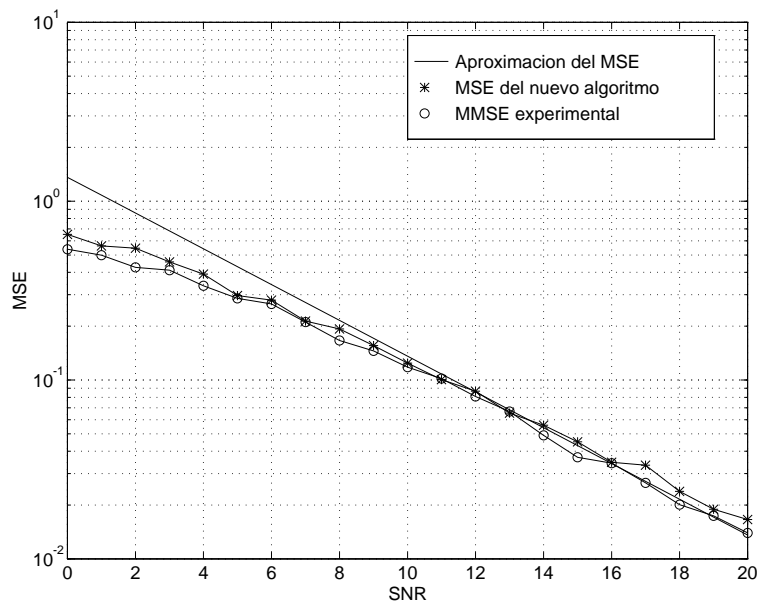


Figura 2.8: Curvas de comparación del MSE respecto a la SNR en recepción del nuevo algoritmo, MMSE y de la aproximación obtenida en la sección 2.7.3.

consideraremos que los símbolos transmitidos, $s(n)$, corresponden a una constelación binaria PAM. En este caso, la kurtosis de los símbolos es negativa ($k - 3 = -2$) y, por lo tanto, satisface una de las condiciones obtenidas para garantizar el correcto comportamiento del algoritmo. Los coeficientes del algoritmo son inicializados de forma aleatoria y serán adaptados de acuerdo con el algoritmo de gradiente (2.44) con un parámetro de control de la convergencia suficientemente pequeño para reducir el ruido de desajuste producido por el algoritmo del gradiente, $\mu = 0.002$.

El comportamiento del algoritmo es evaluado en términos de la ISI normalizada definida como

$$ISI = \frac{\sum_{i=0}^{\infty} |g_i|^2 - \max_i |g_i|^2}{\max_i |g_i|^2} \quad (2.93)$$

donde g_i es la amplitud del símbolo $s(n - i)$ a la salida del igualador.

El primer experimento considera una transmisión libre de ruido a través de un canal AR de primer orden con la siguiente función de transferencia

$$H(z) = \frac{1}{1 - 0.5z^{-1}} \quad (2.94)$$

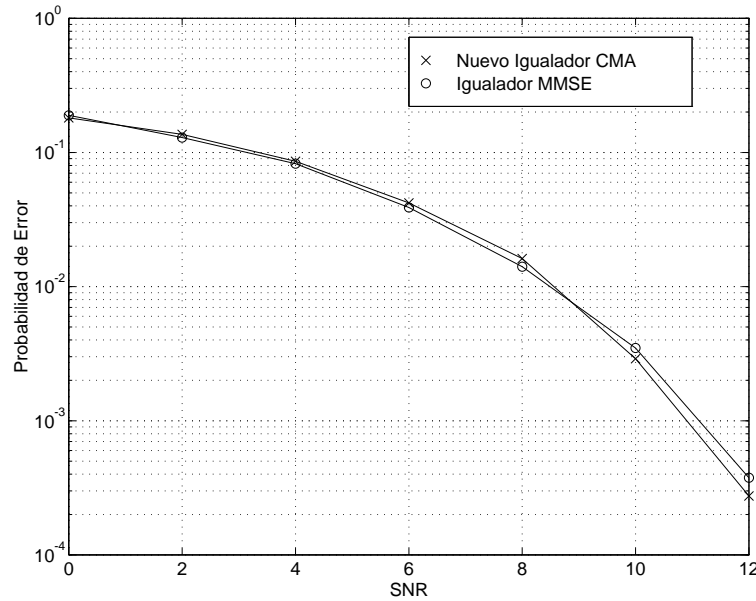


Figura 2.9: Curvas de comparación de las probabilidades de error respecto a la SNR en recepción del nuevo algoritmo y del MMSE

La evolución temporal de la ISI normalizada para un igualador de dos coeficientes está representada en la figura 2.10 para tres valores diferentes del parámetro α . Los valores $\alpha = 0$ (CMA convencional) y $\alpha = 0.3$ no cumplen la condición (2.64) para $k = 1$ y, como se aprecia en la figura, el algoritmo converge hacia un LDLM sin que se cancele completamente la ISI. Sin embargo, el valor $\alpha = 4/3$ satisface el límite obtenido en (2.64) y el algoritmo converge hacia los puntos de equilibrio deseados cancelando completamente la ISI. Un segundo experimento considera el mismo canal de comunicaciones para un igualador con cinco coeficientes. La figura 2.11 representa la ISI normalizada para este experimento, en el que se puede observar de nuevo como el algoritmo no elimina la ISI para $\alpha = 0$ (CMA) y $\alpha = 0.3$ debido a la convergencia hacia un mínimo no deseado, mientras que la eliminación completa de la ISI se produce para $\alpha = 4/3$.

Un tercer experimento muestra el comportamiento del algoritmo con la función de coste generalizada (2.43). Consideramos una transmisión libre de ruido en un canal AR de segundo orden con la siguiente función de transferencia

$$H(z) = \frac{1}{(1 - az^{-1})(1 - bz^{-1})} \quad (2.95)$$

Para este caso, y debido a que el número de coeficientes ha de ser mayor que el orden del canal, usaremos un igualador de tres coeficientes. De esta forma aparecerán dos

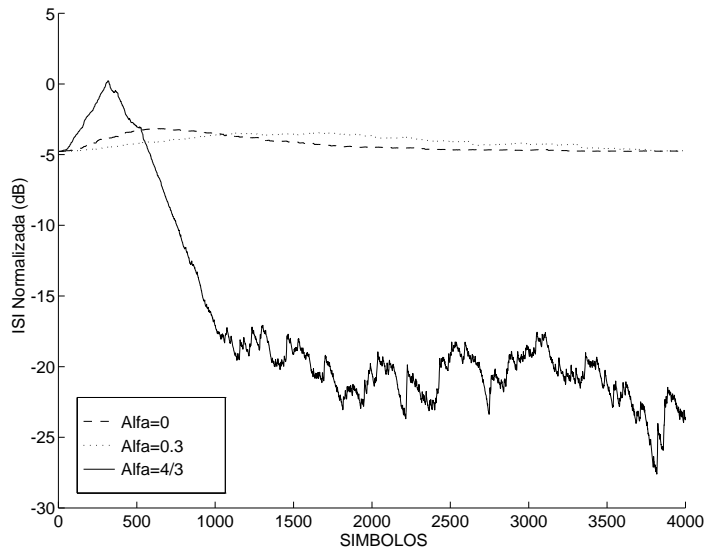


Figura 2.10: Evolución temporal de la ISI cuando se considera un igualador de dos coeficientes y un canal AR(1) con función de transferencia $H(z) = \frac{1}{1-0.5z^{-1}}$. ($\mu = 0.002$).

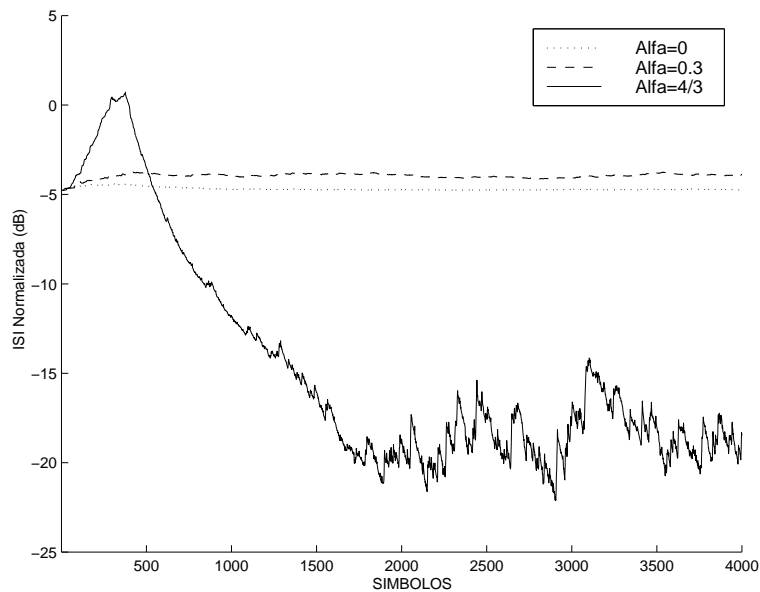


Figura 2.11: Evolución temporal de la ISI cuando se considera un igualador de cinco coeficientes y un canal AR(1) con función de transferencia $H(z) = \frac{1}{1-0.5z^{-1}}$. ($\mu = 0.002$).

términos de penalización en (2.43). La figura 2.12 dibuja la ISI normalizada para diferentes pares de polos, $\alpha = 4/3$ y $\mu = 0.004$. Se puede observar como el igualador alcanza un mínimo deseado y, por lo tanto, cancela la ISI.

Los resultados obtenidos de estos experimentos demuestran la correcta convergencia del algoritmo para valores de α que satisfacen (2.64). Esto ilustra que el límite derivado para el caso particular de un canal AR de primer orden y un igualador de dos coeficientes puede ser aplicado para un caso general de canales AR(r) con un igualador de L coeficientes.

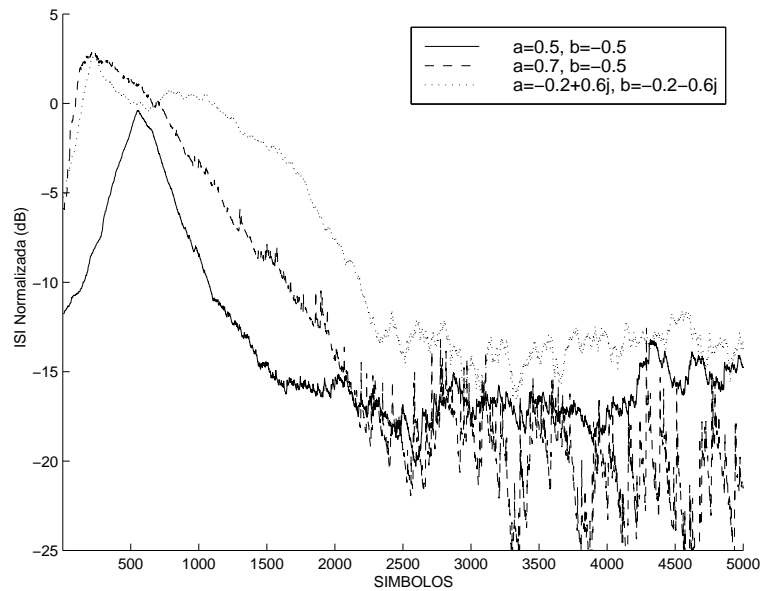


Figura 2.12: Evolución temporal de la ISI cuando se considera un igualador de tres coeficientes y un canal AR(2) con función de transferencia $H(z) = \frac{1}{(1-az^{-1})(1-bz^{-1})}$. ($\mu = 0.004$).

Finalmente, el último experimento muestra la convergencia del nuevo algoritmo en presencia de ruido aditivo. En esta ocasión se ha considerado un igualador de dos coeficientes y un canal AR de primer orden, con función de transferencia idéntica a la del primer y segundo experimento. La relación señal a ruido considerada es $SNR = 10$ dB. Como se puede observar en la figura 2.13, para $\alpha = 0$ (CMA) de nuevo el igualador converge a un mínimo no deseado y no se elimina la ISI. Sin embargo, para $\alpha = 4/3$ el algoritmo converge adecuadamente y se produce una reducción de la ISI. Es necesario apuntar que en este caso se obtiene una ISI residual superior a la de casos anteriores debido a que el igualador debe mantener un compromiso entre la amplificación del ruido y la eliminación de la ISI.

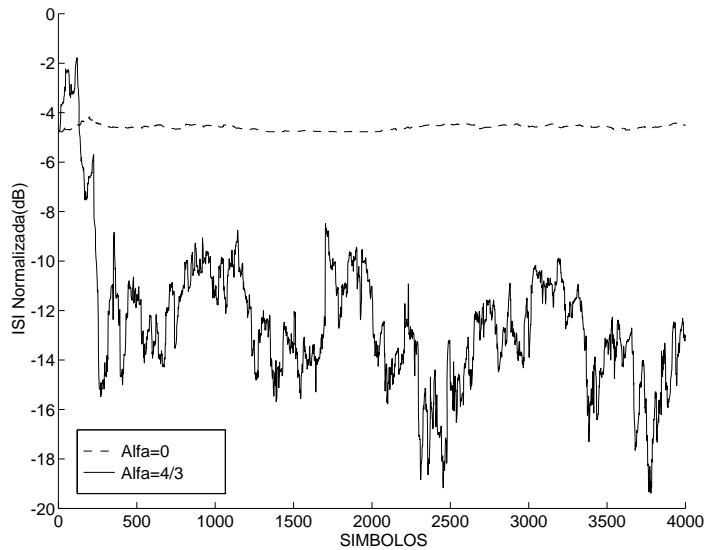


Figura 2.13: Evolución temporal de la ISI cuando se considera un igualador de dos coeficientes, un canal AR(1) con función de transferencia $H(z) = \frac{1}{1-0.5z^{-1}}$ y ruido aditivo con $SNR = 10$ dB. ($\mu = 0.002$).

2.9 Conclusiones

En este capítulo se ha discutido el problema de la igualación adaptativa ciega de canales de comunicaciones auto-regresivos. La matriz de convolución asociada a estos canales resulta tener columnas linealmente dependientes lo que provoca la existencia de mínimos no deseados en los que la ISI no es totalmente cancelada. Debido a que estos mínimos aparecen cuando se consideran igualadores de longitud finita y no cuando su longitud es infinita, se denominan *length dependent local minima* (LDLM).

Para eliminar la posibilidad de convergencia hacia los LDLM en los igualadores CMA hemos propuesto una modificación de la función de coste en la que se han añadido unos términos que penalizan la convergencia hacia los puntos no deseados. Se ha presentado un análisis completo de los puntos estacionarios de esta nueva función de coste para una sencilla situación de un canal AR de orden 1 y un igualador de dos coeficientes. Los resultados del análisis muestran que los LDLM desaparecen si se cumplen dos condiciones:

- La kurtosis de los símbolos transmitidos ha de ser negativa. Esta es la misma condición que garantiza la convergencia del igualador CMA convencional y es cumplida por la práctica totalidad de las señales moduladas digitalmente.
- Un parámetro presente en la nueva función de coste ha de ser superior al valor

(2.64) obtenido en el análisis. Este parámetro pondera los términos de penalización añadidos y, por ello, ha de ser suficientemente grande para evitar la convergencia a las soluciones no deseadas.

Los resultados anteriores se han generalizado a canales AR de mayor orden. Aunque no se ha demostrado analíticamente, las simulaciones dan a entender que los LDLM han desaparecido.

Posteriormente, se ha analizado el efecto del ruido en la convergencia del algoritmo. Por razones de sencillez, el análisis se ha efectuado bajo la hipótesis de que la SNR es alta. En este caso, los puntos estacionarios modifican su posición pero no su naturaleza. Se ha calculado la posición de los mínimos cuando existe ruido y se ha calculado el MSE residual en dichos puntos. Se ha demostrado que dicho valor es igual al mínimo MSE, lo cual indica que el igualador propuesto se comporta, en presencia de ruido, como el igualador MMSE mostrando un deseable equilibrio entre cancelación de ISI y amplificación del ruido.

Finalmente, se han presentado los resultados de algunas simulaciones por ordenador que ilustran el comportamiento del igualador propuesto y validan las aproximaciones realizadas durante el análisis.

Capítulo 3

Separación Ciega de Señales

3.1 Introducción

Los sistemas de comunicaciones celulares, las redes de ordenadores sin hilos (wireless LAN) y las comunicaciones vía satélite son ejemplos de sistemas de comunicaciones por radio que han experimentado un espectacular crecimiento a lo largo de los últimos años. La continua necesidad de incrementar la capacidad de estos sistemas obliga a compartir el espectro radio-eléctrico por un número de usuarios cada vez mayor y, por lo tanto, a buscar nuevas técnicas de procesamiento de señal que permitan separar y extraer los usuarios deseados en recepción.

Cuando varios usuarios comparten un mismo canal de comunicaciones es necesario dotar al receptor de algún sistema que permita diferenciarlos. Como ya se explicó en el capítulo de introducción, esto se consigue mediante las denominadas técnicas de diversidad, las cuales aprovechan diferentes propiedades de la señal recibida para obtener varias versiones de una misma señal en recepción. Como se vió también en ese capítulo, según sea la técnica de diversidad utilizada se hablará de diferentes técnicas de acceso múltiple: SDMA, TDMA, FDMA y/o CDMA.

Con independencia de cual sea el tipo de diversidad empleado, siempre nos encontraremos con que un receptor con diversidad dispondrá de un *front-end* de comunicaciones que proporciona diferentes versiones de una señal transmitida. Cuando el objetivo es extraer la información de un sólo usuario, estas señales constituyen la entrada de un sistema lineal MISO (Multiple Input Single Output) en el que la salida es una combinación lineal de las señales recibidas. La existencia de diversidad permite mejorar la calidad de la señal recibida aumentando la SNR y disminuir los efectos del desvanecimiento y las interferencias.

En este capítulo se va a abordar un problema más complejo en el que el objetivo no es extraer un único usuario sino varios simultáneamente. La arquitectura considerada

consiste en un sistema MIMO formado por múltiples receptores MISO para canales no dispersivos. Cuando el comportamiento es el correcto, cada una de las salidas del sistema MIMO se corresponden con una secuencia transmitida por un único y diferente usuario.

Este capítulo considera un canal no dispersivo que únicamente introduce una mezcla lineal instantánea de los símbolos transmitidos por diferentes usuarios. Un receptor de máxima verosimilitud extraería de forma óptima los diferentes usuarios. Sin embargo, este receptor presenta una complejidad que crece exponencialmente con el número de usuarios del sistema. Por ello se prefiere otro tipo de receptores que no se pueden considerar subóptimos pero de una complejidad reducida.

Los denominados receptores de decorrelación eliminan completamente la MAI en recepción cuando el canal considerado está libre de ruido. Sin embargo, en situaciones prácticas donde hay presencia de ruido aditivo, estos receptores pueden incrementar la potencia del ruido a la salida. Esto motiva la búsqueda de nuevos criterios para ajustar los coeficientes de un sistema MIMO, de forma que se separen de las señales de los diferentes usuarios reduciendo la MAI sin amplificar la potencia del ruido. El criterio MMSE persigue la minimización del error cuadrático medio a la salida del sistema. Sin embargo, estos criterios presentan el gran inconveniente de necesitar una secuencia de entrenamiento que puede reducir la efectividad y utilidad del sistema.

Los criterios ciegos surgen como alternativa a los MMSE evitando la necesidad de las secuencias de entrenamiento. Uno de los criterios ciegos utilizados en mezclas lineales de usuarios, es el denominado criterio de mínima varianza con restricciones lineales [34] ampliamente empleado en procesado en array [36]. Este criterio ajusta los coeficientes del sistema para minimizar la potencia de salida sujeta a una serie de restricciones que impiden que la señal deseada no sea cancelada en el proceso de minimización. Este criterio necesita un conocimiento previo de la matriz de mezclas.

Debido a que consideramos un sistema que tiene a su entrada una mezcla lineal de señales estadísticamente independientes entre sí (estas corresponden a usuarios aislados entre sí), otra alternativa es utilizar criterios desarrollados en el contexto de separación ciega de fuentes. Varios son los criterios propuestos para la separación de fuentes. En concreto, en [15] y [47] se presenta el concepto de función de contraste como aquella función estadística de los coeficientes del sistema MIMO cuyos máximos se corresponden con valores de los coeficientes que obtienen en las salidas los diferentes usuarios. El gran problema de este tipo de criterios es la existencia de múltiples puntos estacionarios no deseados.

El criterio del Módulo Constante es un criterio ciego muy conocido y empleado en la separación de usuarios. Basa su funcionamiento en la minimización de una función de coste que contiene estadísticos de orden superior de las señales de salida del sistema [29]. Como ya se vió en el capítulo anterior, este criterio fué desarrollado en el contexto de la igualación adaptativa ciega de canales, siendo posteriormente utilizado

en el campo de la conformación de haz [30] y [43]. Los análisis de la función de coste CM han demostrado que su minimización obtiene a la salida del sistema MIMO una señal correspondiente a un único usuario. Es decir, en un entorno multiusuario es posible extraer la señal correspondiente a un único usuario eliminando las interferencias correspondientes con las señales de otros usuarios. El gran problema del CM en aplicaciones multiusuario, es que no se puede determinar *a priori* cual será el usuario que extraiga el sistema. Por lo tanto, si cada una de las salidas del sistema MIMO están reguladas por distintos e independientes criterios CM no se puede garantizar cual será la señal obtenida en cada una de ellas y un mismo usuario podría ser extraído en varias salidas simultáneamente.

Este capítulo propone un nuevo criterio de optimización basado en CM para la adaptación de los coeficientes de un sistema MIMO. El criterio pretende una adaptación conjunta de los coeficientes correspondientes a las distintas salidas, sin necesidad de ninguna estrategia previa de inicialización. Para evitar la obtención de un mismo usuario en las salidas del sistema, se define una función de coste que involucra las diferentes salidas del sistema y penaliza la extracción de una misma señal por varias salidas. Además, en este capítulo se demuestra que la minimización de esta función de coste obtiene una perfecta extracción de los diferentes usuarios en las salidas del sistema MIMO.

3.2 Modelo de Señal

Como se vió en el capítulo 1, un receptor con diversidad proporciona unas observaciones que son mezclas de las señales transmitidas por los usuarios que comparten el medio. Un vector que agrupe las señales observadas puede expresarse por medio de una matriz de mezcla, \mathbf{H} , de la siguiente forma

$$\mathbf{x}(n) = [x_0(n), \dots, x_{L-1}(n)]^T = \mathbf{H}\mathbf{s}(n) + \mathbf{r}(n) \quad (3.1)$$

donde $\mathbf{s}(n) = [s_1(n), \dots, s_N(n)]^T$ es un vector de tamaño $N \times 1$ que contiene los símbolos transmitidos por cada usuario y el vector $\mathbf{r}(n)$ representa el ruido presente a la entrada del receptor que, dependiendo de la técnica de acceso, puede ser blanco o coloreado (por ejemplo, en el caso de CDMA donde los filtros adaptados provocan este coloreado). Como se vió en el primer capítulo, al considerar canales no dispersivos, las mezclas de los usuarios son instantáneas.

Un sistema MISO lineal combina las señales presentes a la entrada de forma que la salida i -ésima se puede escribir como

$$y_i(n) = \sum_{l=0}^{L-1} w_{il}^* x_l(n) = \mathbf{w}_i^H \mathbf{x}(n) \quad (3.2)$$

donde $\mathbf{w}_i = [w_{i0}, \dots, w_{i(L-1)}]^T$ es el vector de coeficientes que determina la combinación lineal correspondiente a la salida i -ésima del sistema MIMO. La columna k -ésima (\mathbf{h}_k) de la matriz de mezcla representa las contribuciones del usuario k en las entradas del sistema. El vector de entrada (3.1) se puede expresar por medio de una suma de vectores de la forma

$$\mathbf{x}(n) = \sum_{k=1}^N \mathbf{h}_k s_k(n) + \mathbf{r}(n) \quad (3.3)$$

Sustituyendo esta expresión en (3.2) obtenemos

$$\begin{aligned} y_i(n) &= \sum_{k=1}^N \mathbf{w}_i^H \mathbf{h}_k s_k(n) + \mathbf{w}_i^H \mathbf{r}(n) = \\ &= \sum_{k=1}^N g_{ik} s_k(n) + \mathbf{w}_i^H \mathbf{r}(n) = \mathbf{g}_i^T \mathbf{s}(n) + \mathbf{w}_i^H \mathbf{r}(n) \end{aligned} \quad (3.4)$$

donde $g_{ik} = \mathbf{w}_i^H \mathbf{h}_k$ representa la amplitud de la señal correspondiente al usuario k en la salida i -ésima $\mathbf{g}_i^T = \mathbf{w}_i^H \mathbf{H}$ es el vector de ganancias para esta salida. Agrupando las salidas del sistema en un vector, éste se puede expresar como

$$\mathbf{y}(n) = \mathbf{W}^H \mathbf{x}(n) = \mathbf{G}^T \mathbf{s}(n) + \mathbf{W}^H \mathbf{r}(n) \quad (3.5)$$

donde $\mathbf{W} = [\mathbf{w}_0, \dots, \mathbf{w}_{L-1}]$ representa una matriz que agrupa los vectores de coeficientes del sistema y $\mathbf{G}^T = \mathbf{W}^H \mathbf{H}$ representa la matriz de ganancias cuyo elemento (i, k) es g_{ik} . El objetivo del sistema MIMO consiste en ajustar sus coeficientes para que la matriz de ganancias sea diagonal, o una versión permutada de ésta, obteniendo en las salidas del sistema los símbolos correspondientes a los diferentes usuarios.

En una técnica de acceso CDMA el número de señales de entrada es igual al de salidas del sistema MIMO y a su vez idéntico al número de usuarios en el sistema. En otras técnicas de acceso múltiple podríamos obtener un mayor número de entradas que de salidas o usuarios (por ejemplo, aumentando el número de sensores de una antena en array). Para simplificar la notación, consideraremos que el número de observaciones es siempre igual al de salidas o usuarios en el sistema, es decir N . Obsérvese que en este caso la matriz de mezcla, \mathbf{H} , es cuadrada y de tamaño $N \times N$. Para simplificar la notación, y sin pérdida de generalidad, consideraremos en el resto del capítulo un sistema MIMO con el mismo número de entradas que de salidas, es decir, N .

3.3 Receptor de Decorrelación

La matriz de ganancias \mathbf{G} en (3.5) representa el efecto conjunto de la mezcla del canal y del sistema MIMO. Parece lógico pensar que un receptor que converja hacia una

solución con \mathbf{G}^T diagonal obtendrá en cada salida, en ausencia de ruido ($\mathbf{r}(n) = \mathbf{0}$), un único y diferente usuario¹.

En un entorno sin ruido, la salida del sistema MIMO es de la forma

$$\mathbf{y}(n) = \mathbf{W}^H \mathbf{x}(n) = \mathbf{W}^H \mathbf{H} \mathbf{s}(n) = \mathbf{G}^T \mathbf{s}(n) \quad (3.6)$$

La solución óptima consiste en obtener una matriz \mathbf{W}^T de forma que la matriz de ganancias sea diagonal y que, por cuestiones de claridad, supondremos que tiene sus elementos iguales a uno de forma que $\mathbf{G}^T = \mathbf{I}$, siendo \mathbf{I} la matriz identidad. Es decir

$$\mathbf{G}^T = \mathbf{W}^H \mathbf{H} = \mathbf{I} \Leftrightarrow \mathbf{W}^H = \mathbf{H}^{-1} \quad (3.7)$$

Un sistema que obtenga una solución de este estilo se denomina receptor de decorrelación [44, 45]. Este tipo de receptores, en el caso hipotético de ausencia de ruido, recupera los símbolos de los diferentes usuarios eliminando totalmente la MAI. La gran ventaja de este tipo de receptores es su robustez frente a situaciones de *near-far*. Sin embargo, cuando en el canal hay ruido, el receptor presenta a su salida una componente de ruido que, como se muestra en (3.5), es igual a $\mathbf{W}^H \mathbf{r}(n) = \mathbf{H}^{-1} \mathbf{r}(n)$ y que puede ser amplificada según sean las características de la matriz de mezcla inversa.

Obsérvese la gran similitud que existe entre el receptor de decorrelación para la separación de mezclas instantáneas y el igualador de *zero-forcing* presentado en el capítulo anterior. Al igual que el igualador de *zero-forcing*, el receptor de decorrelación ha de permitir una MAI residual para evitar un incremento del ruido en la salida reduciendo notablemente la SNR.

3.4 Receptor MMSE

En ausencia de ruido, el receptor de decorrelación elimina completamente la MAI. Sin embargo, desde el punto de vista de minimizar la probabilidad de error puede ser conveniente permitir cierta MAI residual en las salidas del sistema MIMO para dotar al sistema de una mayor libertad que le permita una notable reducción del ruido. Un tipo de receptores que aproximan este funcionamiento son los receptores que tienen como objetivo la minimización del error cuadrático medio entre las salidas y los símbolos deseados, es decir,

$$MSE = \sum_{i=1}^N E[|y_i(n) - s_i(n)|^2] = \text{Tr}\{(\mathbf{y}(n) - \mathbf{s}(n))(\mathbf{y}(n) - \mathbf{s}(n))^H\} \quad (3.8)$$

donde $\text{Tr}\{\cdot\}$ representa la traza de una matriz.

¹Una matriz diagonal con sus líneas permutadas obtiene una solución equivalente donde los usuarios aparecen en salidas distintas.

Estos receptores se denominan MMSE [68]. Al igual que en los igualadores MMSE, para ajustar los coeficientes de forma que minimicen (3.8) se utiliza el algoritmo adaptativo LMS que, para la salida i -ésima, tiene la siguiente forma

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mu e(n)\mathbf{x}(n) \quad (3.9)$$

donde μ es el parámetro que controla la velocidad de convergencia del algoritmo y $e(n) = y_i(n) - s_i(n)$ representa el error entre la salida del sistema y el símbolo deseado.

Los receptores MMSE, al igual que los igualadores basados en este criterio de optimización, requieren períodos de entrenamiento durante los cuales se inicializan los coeficientes del sistema. Después de este período de entrenamiento el receptor es guiado por decisión, DD.

Como se apuntó en el capítulo anterior, hay situaciones prácticas donde puede no ser posible o deseable la transmisión de secuencias de entrenamiento. Además, en comunicaciones multiusuario, las características del canal pueden variar bruscamente cada vez que un usuario interferente es activado en cuyo caso es necesario interrumpir la transmisión en curso e intercalar secuencias de entrenamiento para ajustar de nuevo los coeficientes del sistema. Esta limitación es superada por los receptores ciegos.

3.5 Receptores Ciegos

Los receptores descritos en las secciones anteriores necesitan el conocimiento de los símbolos transmitidos o de las características del canal para ajustar los coeficientes del sistema MIMO. Los receptores ciegos ajustan sus coeficientes haciendo uso únicamente de las señales observadas.

Existen múltiples tipos de receptores ciegos que han sido propuestos para otras aplicaciones como la igualación de canales o la separación de fuentes. A continuación se describen algunos de los receptores más conocidos.

3.5.1 Receptores LCMV

Este receptor está basado en la minimización de la varianza a la salida del sistema sujeta a un conjunto de restricciones lineales (LCMV, Linearly Constrained Minimum Variance). Las restricciones lineales son impuestas sobre el vector de coeficientes impidiendo que el proceso de minimización de la varianza afecte a la señal deseada. Este tipo de receptores fueron presentados para técnicas de diversidad espacial [36] o de diversidad por código [34].

El problema de optimización correspondiente a la salida i -ésima consiste en

$$\min_{\mathbf{w}_i} \mathbf{w}_i^H \mathbf{R}_x \mathbf{w}_i \quad \text{sujeto a} \quad \mathcal{R} \mathbf{w}_i = \mathbf{f} \quad (3.10)$$

donde $\mathbf{R}_x = E[\mathbf{x}(n)\mathbf{x}^H(n)]$ es la matriz de covarianza del vector de datos de entrada, \mathcal{R} es la matriz de restricciones y \mathbf{f} especifica los valores de las restricciones.

La forma de implementar este receptor de forma adaptativa se realiza mediante una estructura conocida como cancelador de lóbulos secundarios (GSC, Generalized Sidelobe Canceller) [31] que convierte el problema de minimización con restricciones (3.10) en otro equivalente sin restricciones. Gracias a este mecanismo, el problema se reduce a la minimización de una función de coste tipo MSE mediante un algoritmo LMS.

3.5.2 Receptores Basados en Técnicas de Separación Ciega de Fuentes

Del contexto de la separación ciega de fuentes se pueden extraer nuevos tipos de receptores cuya justificación se basa en el conocido corolario del teorema de Darmois-Skitovich: si \mathbf{s} es un vector de señales estadísticamente independientes con distribución no gaussiana, entonces $\mathbf{y} = \mathbf{G}\mathbf{s}$ es un vector de señales estadísticamente independientes si y sólo si $\mathbf{G} = \mathbf{P}$, donde \mathbf{P} es una matriz de permutación. Estos receptores se basan en las denominadas técnicas ICA (Independent Component Analysis).

Los primeros trabajos sobre este tipo de técnicas fueron desarrollados por Héroult y Jutten [37] que utilizaron una red neuronal recursiva cuyas salidas están directamente conectadas a las entradas. Otros criterios basados en las técnicas ICA [3] miden la independencia estadística entre las salidas por medio de la distancia existente entre la función de densidad de probabilidad conjunta del vector de salida y el producto de las funciones de densidad de cada una de las salidas. Bell y Sejnowski [5] ajustan los coeficientes de forma que se maximice la transferencia de información en una red neuronal de una sola capa.

Finalmente Comon [15], motivado por el teorema de Darmois-Skitovich, introduce el concepto de función de contraste para designar aquellas cuyos máximos coinciden con la extracción de señales estadísticamente independientes a la salida del sistema. Estas funciones de contraste se basan en la suma de cumulantes de segundo orden. Posteriormente, Moreau y Macchi [47] presentaron nuevas funciones de contraste basadas en cumulantes de cuarto orden. Para la maximización de las funciones de contraste utilizan sencillos algoritmos de gradiente ascendente.

Sin embargo, el gran problema de estos receptores radica en que hasta la fecha no existen trabajos exhaustivos que muestren las condiciones bajo las cuales estos criterios se comportan correctamente convergiendo hacia situaciones que extraigan los diferentes usuarios a la salida.

3.5.3 Receptores CMA

Del contexto de la igualación ciega surgieron múltiples criterios de optimización. Entre ellos el más conocido y utilizado en la práctica debido su sencillez y facilidad de implementación es el algoritmo Módulo Constante, CMA. Este algoritmo define una función de coste para la salida i -ésima cuya minimización extraerá en ésta uno de los usuarios del medio. Esta función es de la siguiente forma

$$J_i = E[(|y_i(n)|^2 - R_2)^2] \quad (3.11)$$

donde $R_2 = \frac{E[|s(n)|^4]}{E[|s(n)|^2]}$ es la constante de dispersión. Debido a la no linealidad $|\cdot|^2$, esta función no es una forma cuadrática respecto al vector \mathbf{w}_i y en principio puede contener varios mínimos.

Se han realizado diversos análisis que muestran que la minimización de esta función de coste corresponde con un comportamiento óptimo del receptor. Así por ejemplo, Agee [2] demostró que cuando el vector de observaciones es obtenido a partir de una señal deseada perturbada con la superposición de ruido gaussiano, los únicos mínimos existentes corresponden a puntos donde la SNR es maximizada. Esta circunstancia se produce siempre y cuando la kurtosis de los símbolos transmitidos sea negativa. Esta condición, que es la misma encontrada para igualación ciega de canales, es satisfecha por prácticamente todas las señales presentes en comunicaciones. Previamente, Lundell y Widrow [43] analizaron los mínimos de la función de coste de CM considerando una técnica de diversidad espacial. Suponiendo que la señal de entrada está compuesta de la suma de varias señales de banda estrecha, complejas, estadísticamente independientes e incidiendo a la antena por diferentes ángulos, la función J_i contiene un mínimo por cada señal del entorno que tenga kurtosis negativa. Cada mínimo se corresponde con la extracción de una única señal en el entorno. Aunque la función contiene otros puntos estacionarios, se demuestra que su hessiano no es definido negativo y que por lo tanto los únicos mínimos existentes son los que acabamos de describir.

Cuando se considera la adaptación simultánea de los coeficientes de cada una de las salidas del sistema MIMO, la función conjunta que se ha de minimizar bajo un criterio CM es

$$J_{CM} = \sum_{i=1}^N J_i = \sum_{i=1}^N E[(|y_i(n)|^2 - R_2)^2] \quad (3.12)$$

Esta función se minimiza cuando cada una de las funciones J_i alcanza el mínimo donde se extrae un único usuario. Sin embargo, y aunque la minimización de J_{CM} corresponde con la extracción de un único usuario en cada salida, no se garantiza que los coeficientes de más de una salida converjan hacia una misma combinación extrayendo el mismo usuario. En esta situación, el sistema MIMO obtendría un mismo usuario en más de una salida no extrayendo otros que están presentes en la

mezcla de la señal de entrada. Este problema de captura de una misma señal se puede superar considerando modificaciones del criterio como la presentada en la siguiente sección.

3.6 Modificación Receptores CMA

Para evitar la extracción de un mismo usuario en varias salidas del sistema MIMO es necesario establecer algún tipo de vínculo entre los diferentes criterios CM asociados a cada salida. Para ello, en esta sección se propone que los vectores de coeficientes \mathbf{w}_i , $i = 1, \dots, N$ sean ajustados de acuerdo a la siguiente función de coste

$$J = \sum_{i=1}^N J_i + \frac{\alpha}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N J_{ij} \quad (3.13)$$

donde

$$J_i = E[(|y_i(n)|^2 - R_2)^2] \quad (3.14)$$

es la función de coste del criterio CM correspondiente a la salida i -ésima del sistema MIMO y

$$J_{ij} = |E[y_i(n)y_j^*(n)]|^2, \quad i \neq j \quad (3.15)$$

son un conjunto de nuevos términos que aseguran que cada salida se corresponde con la extracción de un usuario diferente. El parámetro α es una constante real y positiva que pondera el valor de estos nuevos términos. Esta función fue presentada por primera vez en [19].

El motivo de la inclusión de los nuevos términos en la función de coste es el siguiente: al minimizar (3.13), se minimizan simultáneamente cada uno de los criterios CM correspondientes a las salidas del sistema MIMO, J_i , y los términos J_{ij} . Cuando las salidas i y j del sistema MIMO extraen el mismo usuario, $y_i(n)$ y $y_j(n)$ están altamente correladas entre sí, el término J_{ij} será no nulo y por lo tanto no se alcanzará la minimización de la función de coste conjunta J . Es decir, los nuevos términos penalizan la convergencia del algoritmo hacia soluciones no deseadas. Esta idea de la correlación cruzada entre diferentes salidas ha sido anteriormente utilizada por varios autores (ver [53] y [52] y sus referencias) en el ámbito de la separación de señales.

Es interesante observar como la función (3.13) puede interpretarse como el Lagrangiano asociado al problema de optimización

$$\min_{\mathbf{w}} \sum_{i=1}^N E[(|y_i(n)|^2 - R_2)^2] \quad \text{sujeto a} \quad \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N |E[y_i(n)y_j^*(n)]|^2 = 0 \quad (3.16)$$

La ventaja de la aproximación presentada en este capítulo es que se puede optimizar mediante algoritmos adaptativos muy sencillos. La resolución del problema con restricciones obliga a plantear un algoritmo en dos etapas: una que minimice $\sum_{i=1}^N J_i$ y otra que imponga la restricción.

Shynk y Gooch [61] presentan otro tipo de receptor que supera el problema de captura del CM. Este consiste en la conexión en cascada de diversos subsistemas MISO con criterios de adaptación CM. La conexión entre cada uno de estos subsistemas se realiza por medio de un cancelador de señal que elimina de la entrada de cada uno la señal obtenida en la salida del anterior, y así sucesivamente. Una vez obtenido un usuario a la salida de uno de los subsistemas, el cancelador de señal garantiza que no va ser extraído por los subsistemas posteriores. El principal inconveniente de esta aproximación es que, además incrementar notablemente la complejidad del hardware que regula el sistema MIMO, el algoritmo puede tener una velocidad de convergencia muy reducida debido a la conexión en cascada de múltiples criterios ciegos [19].

Otra solución para el problema de captura de CM consiste en utilizar sistemas MISO inconexos que dispongan de una estrategia de inicialización de coeficientes que garantice que la convergencia de los algoritmos adaptativos de cada uno de estos sistemas converja hacia la extracción del usuario adecuado. Mathur [46] plantea una inicialización de los coeficientes del sistema MIMO basada en el algoritmo MUSIC (Multiple Signal Classification) [59] que conduce al algoritmo hacia soluciones donde se extraen los diferentes usuarios en las salidas del sistema.

Finalmente, Cerquides [12] considera la conexión en paralelo de múltiples sistemas MISO, solucionando el problema de captura del CM por medio de la cancelación, en la entrada de cada uno de éstos, de las señales obtenidas en las salidas. Esta estructura se puede interpretar como que los sistemas MISO compiten por las distintas señales del entorno.

3.6.1 Algoritmo Adaptativo

La forma más sencilla de minimizar la función de coste (3.13) es por medio de un algoritmo de tipo gradiente estocástico donde el vector \mathbf{w}_i es actualizado mediante la siguiente fórmula recursiva

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) - \mu \nabla_{\mathbf{w}_i} J(n), \quad i = 1, \dots, N \quad (3.17)$$

donde μ es el parámetro que controla la convergencia del algoritmo y $\nabla_{\mathbf{w}_i} J(n)$ es el vector gradiente de J con respecto a \mathbf{w}_i en el punto $\mathbf{w}_i(n)$ de la forma

$$\nabla_{\mathbf{w}_i} J = \nabla_{\mathbf{w}_i} J_i + \alpha \sum_{\substack{j=1 \\ j \neq i}}^N \nabla_{\mathbf{w}_i} J_{ij} \quad (3.18)$$

Considerando (3.2) y las propiedades del operador gradiente complejo [8] las componentes del vector gradiente pueden ser expresadas como

$$\begin{aligned}
\nabla_{\mathbf{w}_i} J_i &= \nabla_{\mathbf{w}_i} E[(\mathbf{w}_i^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}_i - R_2)(\mathbf{w}_i^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}_i - R_2)] = \\
&= 2E[(\mathbf{w}_i^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}_i - R_2)(2\mathbf{x}^H(n) \mathbf{w}_i \mathbf{x}(n))] = \\
&= 4E[(|y_i(n)|^2 - R_2)y_i^*(n)\mathbf{x}(n)] \\
\nabla_{\mathbf{w}_i} J_{ij} &= \nabla_{\mathbf{w}_i} |E[\mathbf{w}_i^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}_j]|^2 = \\
&= \nabla_{\mathbf{w}_i} E[\mathbf{w}_i^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}_j] E[\mathbf{w}_j^H \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}_i] = \\
&= E[y_i^*(n)y_j(n)]E[2y_j^*(n)\mathbf{x}(n)]
\end{aligned} \tag{3.19}$$

resultando el vector gradiente de J de la siguiente forma

$$\nabla_{\mathbf{w}_i} J = 2(2E[(|y_i(n)|^2 - R_2)y_i^*(n)\mathbf{x}(n)] + \alpha E[y_i^*(n)y_j(n)]E[y_j^*(n)\mathbf{x}(n)]) \tag{3.20}$$

Las esperanzas de (3.20) son desconocidas en la práctica y, por lo tanto, la recursión (3.17) se convierte en

$$\begin{aligned}
\mathbf{w}_i(n+1) &= \mathbf{w}_i(n) - \mu \left\{ \sum_{k=n-1}^n (|y_i(k)|^2 - R_2)y_i^*(k)\mathbf{x}(k) \right. \\
&\quad \left. + \frac{\alpha}{4} \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k=n-1}^n y_i^*(k)y_j(k) \sum_{k=n-1}^n y_j^*(k)\mathbf{x}(k) \right\}
\end{aligned} \tag{3.21}$$

donde para estimar las esperanzas estadísticas de (3.18) se han utilizado dos muestras y un factor de 2 ha sido introducido en el parámetro μ . Con este número de muestras evitamos los resultados ambiguos en los estimadores de una sola muestra cuando se manejan estadísticos de orden superior.

3.7 Análisis de los Puntos Estacionarios

La función de coste J es una forma no cuadrática de \mathbf{w} . Esto puede provocar que existan múltiples mínimos hacia los que el algoritmo podría converger dando lugar a un mal comportamiento del sistema. En esta sección se demuestra que los mínimos de la función J se corresponden con valores del vector de coeficientes donde se obtiene un comportamiento óptimo siempre que se garantice que la kurtosis de todas las señales en el entorno es negativa (idéntica condición a la obtenida en [2] y [43]).

3.7.1 Análisis sin Ruido

Por razones de simplicidad, en este apartado reduciremos nuestro análisis a dos usuarios y un canal libre de ruido. Este análisis consiste en la obtención de los puntos

estacionarios que se corresponden con los puntos donde el gradiente de J se anula. Para determinar si estos puntos son mínimos o no, se realiza posteriormente un análisis de su naturaleza estudiando si la matriz hessiana es definida positiva o no.

Consideremos que las señales transmitidas por dos usuarios $s_1(n)$ y $s_2(n)$ tienen media cero, y son estadísticamente independientes. Supondremos además, que las señales son complejas y circularmente invariantes, es decir, que las componentes en fase y cuadratura están mutuamente incorreladas entre sí y tienen la misma varianza. Esto significa que verifican lo siguiente

$$E[s_1^2(n)] = E[s_2^2(n)] = 0 \quad (3.22)$$

De la ecuación (3.1) se obtiene que las señales observadas en los diferentes sistemas de diversidad se pueden expresar como

$$\mathbf{x}(n) = s_1(n)\mathbf{h}_1 + s_2(n)\mathbf{h}_2 \quad (3.23)$$

donde \mathbf{h}_1 y \mathbf{h}_2 representan respectivamente la primera y segunda columna de la matriz de mezcla. Para recuperar las señales transmitidas por ambos usuarios, es necesario considerar un sistema MIMO con dos salidas. Llamando \mathbf{w}_1 y \mathbf{w}_2 a los vectores de coeficientes correspondientes a las dos salidas del sistema, estas serán

$$\begin{aligned} y_1(n) &= \mathbf{w}_1^H \mathbf{x}(n) = g_{11}s_1(n) + g_{12}s_2(n) \\ y_2(n) &= \mathbf{w}_2^H \mathbf{x}(n) = g_{21}s_1(n) + g_{22}s_2(n) \end{aligned} \quad (3.24)$$

donde $g_{ij} = \mathbf{w}_i^H \mathbf{h}_j$ admite la interpretación física de ser la amplitud del usuario j en la salida i -ésima.

Considerando la función de coste del nuevo criterio de optimización (3.13) de adaptación conjunta presentado en la sección anterior podemos expresar esta en función de las dos salidas de la siguiente forma

$$\begin{aligned} J(\mathbf{w}_1, \mathbf{w}_2) &= E[|y_1(n)|^2 - 1|^2] + E[|y_2(n)|^2 - 1|^2] + \alpha |E[y_1^*(n)y_2(n)]|^2 = \\ &= E[|y_1(n)|^4] + E[|y_2(n)|^4] - 2E[|y_1(n)|^2] - 2E[|y_2(n)|^2] + \\ &\quad + \alpha |E[y_1(n)y_2^*(n)]|^2 + 2 \end{aligned} \quad (3.25)$$

donde se ha considerado, por simplicidad, que las constantes de dispersión de las señales transmitidas cumplen que $R_2 = 1$. Se puede observar que la función (3.25) depende de los momentos de cuarto orden de las salidas y que, por lo tanto, no es una función cuadrática de \mathbf{w}_i .

Considerando (3.24) y las hipótesis estadísticas que hemos establecido sobre $s_1(n)$ y $s_2(n)$, los operadores esperanza de (3.25) pueden ser expresados como

$$E[|y_1(n)|^4] = |g_{11}|^4 q_1 + |g_{12}|^4 q_2 + 4|g_{11}|^2 |g_{12}|^2 m_1 m_2$$

$$\begin{aligned}
E[|y_2(n)|^4] &= |g_{21}|^4 q_1 + |g_{22}|^4 q_2 + 4|g_{21}|^2 |g_{22}|^2 m_1 m_2 \\
E[|y_1(n)|^2] &= |g_{11}|^2 m_1 + |g_{12}|^2 m_2 \\
E[|y_2(n)|^2] &= |g_{21}|^2 m_1 + |g_{22}|^2 m_2 \\
E[y_1(n)y_2^*(n)] &= g_{11}g_{21}^* m_1 + g_{12}g_{22}^* m_2
\end{aligned} \tag{3.26}$$

donde $q_i = E[|s_i(n)|^4]$ y $m_i = E[|s_i(n)|^2]$ son los momentos de cuarto y segundo orden de $s_i(n)$, respectivamente.

Sustituyendo (3.26) en (3.25), la función J puede ser expresada en función de las amplitudes g_{ij} de la siguiente forma

$$\begin{aligned}
J(g_{11}, g_{12}, g_{21}, g_{22}) &= \\
&= |g_{11}|^4 q_1 + |g_{12}|^4 q_2 + 4|g_{11}|^2 |g_{12}|^2 m_1 m_2 + |g_{21}|^4 q_1 + |g_{22}|^4 q_2 + \\
&\quad + 4|g_{21}|^2 |g_{22}|^2 m_1 m_2 - 2|g_{11}|^2 m_1 - 2|g_{12}|^2 m_2 - 2|g_{21}|^2 m_1 - 2|g_{22}|^2 m_2 + 2 + \\
&\quad + \alpha (|g_{11}|^2 |g_{21}|^2 m_1^2 + |g_{12}|^2 |g_{22}|^2 m_2^2 + g_{11}g_{12}^* g_{21}^* g_{22} m_1 m_2 + g_{11}^* g_{12} g_{21} g_{22}^* m_1 m_2)
\end{aligned} \tag{3.27}$$

Como se puede observar, se ha obtenido una expresión de J más conveniente, pues esta depende ahora de cuatro variables complejas g_{ij} $i = 1, 2$ $j = 1, 2$ en lugar de los vectores de coeficientes \mathbf{w}_1 y \mathbf{w}_2 , cada uno de ellos con dimensión N . Esto simplifica notablemente el análisis de los puntos estacionarios como se comprobará a continuación.

Los puntos estacionarios se corresponden con aquellos donde los gradientes $\nabla_{\mathbf{w}_1} J$ y $\nabla_{\mathbf{w}_2} J$ se anulan. Usando la regla de la cadena y las propiedades del operador gradiente descritas en [8], $\nabla_{\mathbf{w}_1} J$ y $\nabla_{\mathbf{w}_2} J$ pueden ser expresados de la siguiente forma

$$\begin{aligned}
\nabla_{\mathbf{w}_1} J &= \frac{\partial J}{\partial g_{11}} \nabla_{\mathbf{w}_1} g_{11} + \frac{\partial J}{\partial g_{12}} \nabla_{\mathbf{w}_1} g_{12} = \frac{\partial J}{\partial g_{11}} \mathbf{h}_1 + \frac{\partial J}{\partial g_{12}} \mathbf{h}_2 = 0 \\
\nabla_{\mathbf{w}_2} J &= \frac{\partial J}{\partial g_{21}} \nabla_{\mathbf{w}_2} g_{21} + \frac{\partial J}{\partial g_{22}} \nabla_{\mathbf{w}_2} g_{22} = \frac{\partial J}{\partial g_{21}} \mathbf{h}_1 + \frac{\partial J}{\partial g_{22}} \mathbf{h}_2 = 0
\end{aligned} \tag{3.28}$$

Por lo tanto, los vectores gradiente pueden ser expresados como una combinación lineal de las columnas de la matriz de mezcla, \mathbf{h}_1 y \mathbf{h}_2 . Como ya hemos visto, las columnas de las matrices de mezcla son linealmente independientes y, por lo tanto, el gradiente se anulará en puntos donde las derivadas parciales respecto a g_{ij} se anulen

$$\nabla_{\mathbf{w}_1} J = \nabla_{\mathbf{w}_2} J = 0 \iff \frac{\partial J}{\partial g_{11}} = \frac{\partial J}{\partial g_{12}} = \frac{\partial J}{\partial g_{21}} = \frac{\partial J}{\partial g_{22}} = 0 \tag{3.29}$$

Calculando las derivadas parciales de la función J respecto a g_{ij} e igualando estas a cero, los puntos estacionarios se corresponderán con las soluciones del siguiente sistema de ecuaciones

$$\frac{\partial J}{\partial g_{11}} = g_{11}^* (2|g_{11}|^2 q_1 + 4|g_{12}|^2 m_1 m_2 - 2m_1 + \alpha |g_{21}|^2 m_1^2) + \alpha g_{21}^* g_{12}^* g_{22} m_1 m_2 = 0$$

$$\begin{aligned}
\frac{\partial J}{\partial g_{12}} &= g_{12}^*(2|g_{12}|^2 q_2 + 4|g_{11}|^2 m_1 m_2 - 2m_2 + \alpha|g_{22}|^2 m_2^2) + \alpha g_{11}^* g_{21} g_{22}^* m_1 m_2 = 0 \\
\frac{\partial J}{\partial g_{21}} &= g_{21}^*(2|g_{21}|^2 q_1 + 4|g_{22}|^2 m_1 m_2 - 2m_1 + \alpha|g_{11}|^2 m_1^2) + \alpha g_{11}^* g_{12} g_{22}^* m_1 m_2 = 0 \\
\frac{\partial J}{\partial g_{22}} &= g_{22}^*(2|g_{22}|^2 q_2 + 4|g_{21}|^2 m_1 m_2 - 2m_2 + \alpha|g_{12}|^2 m_2^2) + \alpha g_{11} g_{21}^* g_{12}^* m_1 m_2 = 0
\end{aligned} \tag{3.30}$$

Obsérvese que este sistema es no lineal y contiene numerosas soluciones que determinan los puntos estacionarios de la función J . Existen múltiples métodos para estudiar la naturaleza de estos puntos, es decir, si son mínimos o no. El análisis presentado en el apéndice D.1 realiza un test sobre la matriz hessiana para comprobar si esta es definida positiva (el punto estacionario sería un mínimo) o no. A continuación se resume el análisis del apéndice D.1 en donde se han encontrado seis grupos de puntos estacionarios.

Grupo 1 Ambas salidas cancelan las señales de los dos usuarios, es decir,

$$g_{11} = g_{12} = g_{21} = g_{22} = 0 \tag{3.31}$$

Este es un punto estacionario no deseado, pero en el apéndice D.1 se demuestra que siempre corresponde con un máximo de la función J .

Grupo 2 Una salida extrae la señal correspondiente a uno de los usuarios mientras que la otra cancela ambas señales. Uno de los cuatro puntos posibles pertenecientes a este grupo es

$$|g_{11}|^2 = \frac{m_1}{q_1}, \quad g_{12} = g_{21} = g_{22} = 0 \tag{3.32}$$

Estos puntos estacionarios son no deseados, pero en el apéndice D.1 se demuestra que la matriz hessiana nunca es definida positiva y por lo tanto estos puntos no pueden ser mínimos de J .

Grupo 3 Una de las salidas extrae una combinación lineal de las señales correspondientes a los dos usuarios y la otra cancela ambas. Uno de los dos puntos posibles pertenecientes a este grupo es

$$m_1 |g_{11}|^2 = \frac{k_2 - 2}{k_1 k_2 - 4}, \quad m_2 |g_{12}|^2 = \frac{k_1 - 2}{k_1 k_2 - 4}, \quad g_{21} = g_{22} = 0 \tag{3.33}$$

donde $k_i = q_i/m_i^2$ es el momento de cuarto orden normalizado de $s_i(n)$.

Este es un punto estacionario no deseado pero en el apéndice D.1 se demuestra que si la kurtosis $(k - 2)$ de las señales de ambos usuarios es negativa, la matriz hessiana no es definida positiva y, por lo tanto, los puntos estacionarios de este grupo no son mínimos de la función J .

Grupo 4 Ambas salidas extraen la misma señal y cancelan la otra. Uno de los dos puntos posibles pertenecientes a este grupo es

$$|g_{11}|^2 = |g_{21}|^2 = \frac{2m_1}{2q_1 + \alpha m_1^2}, \quad g_{12} = g_{22} = 0 \quad (3.34)$$

Este es de nuevo un punto estacionario no deseado. El análisis del apéndice D.1 demuestra que si la kurtosis de las señales transmitidas es negativa y el parámetro α es mayor que cero, el hessiano no es definido positivo y el punto no puede ser un mínimo.

Obsérvese que este punto estacionario se corresponde con la convergencia no deseada de varios criterios CM adaptados independientemente. Esta convergencia se puede producir si consideramos la minimización de la función de J cuando $\alpha = 0$, que no cumple la condición de α obtenida en el apéndice para garantizar que este punto no es un mínimo.

Grupo 5 Cada salida extrae a un único y diferente usuario. Uno de los dos puntos posibles pertenecientes a este grupo es

$$|g_{11}|^2 = \frac{m_1}{q_1}, \quad |g_{22}|^2 = \frac{m_2}{q_2}, \quad g_{12} = g_{21} = 0 \quad (3.35)$$

Este es un punto estacionario deseado, ya que las salidas extraen señales correspondientes a un único y distinto usuario. El apéndice D.1 demuestra que la condición de que las señales transmitidas tengan kurtosis negativas es suficiente para garantizar que la matriz es definida positiva y que por lo tanto estos puntos son mínimos de J . Es decir, una minimización de J ajustará el vector de coeficientes \mathbf{w} de forma que a la salida del sistema MIMO se obtengan las señales transmitidas por los diferentes usuarios

Grupo 6 Cada salida extrae una combinación lineal de las señales de ambos usuarios, es decir, las cuatro variables g_{ij} son diferentes de cero.

$$g_{11} \neq 0, \quad g_{12} \neq 0, \quad g_{21} \neq 0, \quad g_{22} \neq 0 \quad (3.36)$$

Estos puntos son no deseados, pero en el apéndice D.1 se demuestra que el hessiano no es definido positivo y, por lo tanto, estos puntos no son mínimos de J si las kurtosis de las señales transmitidas son negativas.

Como hemos visto, la función de coste perteneciente al nuevo criterio posee un gran número de puntos estacionarios. El estudio de la naturaleza de los puntos estacionarios mediante el análisis de la matriz hessiana ha mostrado que sólo los pertenecientes al grupo 5, grupo de los puntos deseados, son mínimos siempre que se cumpla que

$$k_i - 3 < 0 \quad i = 1, 2 \quad (3.37)$$

Al ser estos los únicos mínimos de la función, el algoritmo de gradiente que minimiza J siempre obtendrá una convergencia hacia los puntos estacionarios deseados, y por lo tanto se asegura una perfecta extracción de los diferentes usuarios a la salida del sistema MIMO.

Todo el análisis, presentado también en [10], se ha desarrollado para cualquier tipo de sistema MIMO con dos salidas independientemente de la diversidad empleada. Las únicas hipótesis que hemos considerado son que las componentes en fase y cuadratura de las señales transmitidas $s_i(n)$ han de estar mutuamente incorreladas y tener idéntica varianza. Es decir, la constelación de donde se obtienen los símbolos a transmitir ha de ser compleja y simétrica respecto al origen. No obstante un análisis para el caso de una constelación real obtendría resultados semejantes. Un análisis para un sistema MIMO con N salidas sería muy complejo y no ha podido ser desarrollado hasta ahora, aunque intuitivamente todo parece indicar que se obtendría un buen comportamiento del sistema en este caso. En la sección de simulaciones se muestran diferentes experimentos por ordenador que muestran el correcto comportamiento en escenarios más generales que no ha abarcado el análisis.

3.7.2 Análisis con Ruido

Un análisis completo de (3.13) que tenga en cuenta el ruido es extremadamente difícil de realizar. Para poder determinar cual es el comportamiento del nuevo algoritmo en presencia de ruido aditivo, consideraremos que la relación SNR obtenida a la entrada del sistema es suficientemente alta y así suponer que el número y naturaleza de los puntos estacionarios de la función de coste permanecen inalterados. La única consecuencia de la presencia del ruido será una perturbación en la posición de los mínimos de la función. Bajo estas hipótesis de la SNR, se comprueba que en estos mínimos el error cuadrático medio MSE entre la salida del sistema MIMO y los símbolos transmitidos es el mínimo valor posible, MMSE.

Consideramos el MSE (3.8) normalizado con respecto a la potencia de las señales recibidas $m_s = E[s_i^2(n)] \quad \forall i$ que, por simplicidad, se consideran iguales para los N usuarios.

$$\begin{aligned} MSE &= \frac{1}{m_s} \text{Tr} \{ E [(\mathbf{y}(n) - \mathbf{s}(n))(\mathbf{y}(n) - \mathbf{s}(n))^T] \} = \\ &= \frac{1}{m_s} \text{Tr} \{ E[\mathbf{y}(n)\mathbf{y}^T(n)] + E[\mathbf{s}(n)\mathbf{s}^T(n)] - E[\mathbf{y}(n)\mathbf{s}^T(n)] - E[\mathbf{s}(n)\mathbf{y}^T(n)] \} \end{aligned} \quad (3.38)$$

Cuando no hay ruido en el canal las salidas del sistema MIMO que, por razones de simplicidad, consideraremos con N entradas y N salidas agrupando estas últimas en un vector de la forma

$$\mathbf{y}(n) = \mathbf{G}^T \mathbf{s}(n) \quad (3.39)$$

donde $\mathbf{G}^T = \mathbf{W}^T \mathbf{H}$ es la matriz de amplitudes de las señales transmitidas a la salida del sistema y $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_N]$. Desarrollando (3.38) obtenemos

$$\begin{aligned} MSE &= \frac{1}{m_s} E [\text{Tr}\{\mathbf{G}^T \mathbf{s}(n) \mathbf{s}^T(n) \mathbf{G}\} + \text{Tr}\{\mathbf{s}(n) \mathbf{s}^T(n)\} - \\ &\quad - \text{Tr}\{\mathbf{G}^T \mathbf{s}(n) \mathbf{s}^T(n)\} - \text{Tr}\{\mathbf{s}(n) \mathbf{s}^T(n) \mathbf{G}\}] = \\ &= \frac{1}{m_s} [\text{Tr}\{\mathbf{G}^T \mathbf{R}_s \mathbf{G}\} + \text{Tr}\{\mathbf{R}_s\} - \text{Tr}\{\mathbf{G}^T \mathbf{R}_s\} - \text{Tr}\{\mathbf{R}_s \mathbf{G}\}] = \\ &= \text{Tr}\{\mathbf{G}^T \mathbf{G}\} - 2\text{Tr}\{\mathbf{G}\} + N \end{aligned} \quad (3.40)$$

donde se ha considerado que los símbolos de los diferentes usuarios son independientes entre sí y que por lo tanto $\mathbf{R}_s = E[\mathbf{s}(n) \mathbf{s}^T(n)] = m_s \mathbf{I}$, siendo \mathbf{I} la matriz identidad. Para obtener el valor de \mathbf{G} que minimiza el MSE, se ha de calcular el punto donde el gradiente del MSE con respecto a \mathbf{G} se anula, es decir

$$\frac{\partial MSE}{\partial \mathbf{G}} = 2\mathbf{G} - 2\mathbf{I} = \mathbf{0} \quad (3.41)$$

siendo la matriz que minimiza el MSE de la forma

$$\mathbf{G}_0 = \mathbf{I} \quad (3.42)$$

A partir de la expresión de \mathbf{G}_0 , el MSE (3.40) se puede expresar de la siguiente forma:

$$MSE = \text{Tr}[(\mathbf{G} - \mathbf{G}_0)^T (\mathbf{G} - \mathbf{G}_0)] \quad (3.43)$$

Obsérvese que cuando $\mathbf{G} = \mathbf{G}_0 = \mathbf{I}$ el MSE se anula. Este es el valor mínimo del MSE o MMSE. El hecho de obtener una matriz de amplitudes igual a la matriz idéntica indica la presencia de un único usuario en cada una de las salidas del sistema. Consideremos ahora la presencia de ruido aditivo en el canal, la señal de salida del sistema MIMO se puede descomponer de la forma

$$\mathbf{y}(n) = \mathbf{y}_s(n) + \mathbf{y}_r(n) = \mathbf{G}^T \mathbf{s}(n) + \mathbf{W}^T \mathbf{r}(n) \quad (3.44)$$

donde $\mathbf{y}_s(n) = \mathbf{G}^T \mathbf{s}(n)$ e $\mathbf{y}_r(n) = \mathbf{W}^T \mathbf{r}(n)$ son las componentes de señal y de ruido a la salida, respectivamente. Teniendo en cuenta la independencia estadística entre $\mathbf{s}(n)$ y $\mathbf{r}(n)$, la expresión del MSE con ruido es

$$\begin{aligned} \tilde{MSE} &= \frac{1}{m_s} \text{Tr} \{ E[(\mathbf{y}_s(n) + \mathbf{y}_r(n))(\mathbf{y}_s(n) + \mathbf{y}_r(n))^T] + E[\mathbf{s}(n) \mathbf{s}^T(n)] - \\ &\quad - E[(\mathbf{y}_s(n) + \mathbf{y}_r(n)) \mathbf{s}^T(n)] - E[\mathbf{s}(n)(\mathbf{y}_s(n) + \mathbf{y}_r(n))^T] \} \end{aligned} \quad (3.45)$$

donde las esperanzas estadísticas de esta nueva expresión que involucran alguna componente de $\mathbf{y}(n)$ son

$$\begin{aligned} E[(\mathbf{y}_s(n) + \mathbf{y}_r(n))(\mathbf{y}_s(n) + \mathbf{y}_r(n))^T] &= E[\mathbf{y}_s(n) \mathbf{y}_s^T(n)] + E[\mathbf{y}_r(n) \mathbf{y}_r^T(n)] \\ E[(\mathbf{y}_s(n) + \mathbf{y}_r(n)) \mathbf{s}^T(n)] &= E[\mathbf{y}_s(n) \mathbf{s}^T(n)] \\ E[\mathbf{s}(n)(\mathbf{y}_s(n) + \mathbf{y}_r(n))^T] &= E[\mathbf{s}(n) \mathbf{y}_s^T(n)] \end{aligned} \quad (3.46)$$

las cuales se han obtenido teniendo en cuenta la independencia estadística entre las señales y el ruido. Sustituyendo (3.46) y la expresión de \mathbf{y}_r en (3.45) obtenemos

$$\begin{aligned} M\tilde{S}E &= MSE + \frac{1}{m_s} \text{Tr}\{E[\mathbf{y}_r(n)\mathbf{y}_r^T(n)]\} = \\ &= MSE + \frac{1}{m_s} \text{Tr}\{\mathbf{W}^T \mathbf{R}_r \mathbf{W}\} = \\ &= MSE + \gamma \text{Tr}\{\mathbf{W}^T \mathbf{W}\} \end{aligned} \quad (3.47)$$

donde MSE representa el error cuadrático medio para el caso sin ruido (3.40). Para llegar a (3.47) hemos considerado que el ruido es blanco con una matriz de autocorrelación² $\mathbf{R}_r = m_r \mathbf{I}$ y varianza $m_r = E[r_i^2(n)]$, y que $\gamma = \frac{m_r}{m_s}$ es la inversa de la relación señal a ruido. Considerando la relación $\mathbf{W}^T = \mathbf{H}^{-1} \mathbf{G}^T$, el error cuadrático medio se puede expresar como

$$\begin{aligned} M\tilde{S}E &= MSE + \gamma \text{Tr}\{\mathbf{G}^T (\mathbf{H}^{-1} \mathbf{H}^{-1T}) \mathbf{G}\} = \\ &= (\mathbf{G} - \mathbf{G}_0)^T (\mathbf{G} - \mathbf{G}_0) + \gamma \text{Tr}\{\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}\} \end{aligned} \quad (3.48)$$

De nuevo, calculamos los puntos donde el gradiente de MSE con respecto a \mathbf{G} se anula, es decir

$$\frac{\partial M\tilde{S}E}{\partial \mathbf{G}} = 2\mathbf{G} - 2\mathbf{G}_0 + 2\gamma (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G} = \mathbf{0} \quad (3.49)$$

obteniendo la matriz $\tilde{\mathbf{G}}_0$ que satisface esta ecuación y que minimiza el MSE en presencia de ruido

$$\begin{aligned} \tilde{\mathbf{G}}_0 &= (\mathbf{I} + \gamma (\mathbf{H}^T \mathbf{H})^{-1})^{-1} \mathbf{G}_0 = \\ &= \mathbf{G}_0 - \left(\mathbf{I} + \frac{1}{\gamma} (\mathbf{H}^T \mathbf{H}) \right)^{-1} \mathbf{G}_0 \simeq \\ &\simeq \mathbf{G}_0 - \gamma (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}_0 = \\ &= \mathbf{G}_0 + \gamma \Delta \mathbf{G}_{MSE} \end{aligned} \quad (3.50)$$

Para obtener (3.50) hemos utilizado el lema de inversión³ y considerado que tenemos una elevada relación señal a ruido, SNR (es decir, un valor pequeño de γ). Obsérvese que el punto donde se alcanza el valor mínimo de $M\tilde{S}E$, $\tilde{\mathbf{G}}_0$, es una versión perturbada del obtenido para MSE , esta perturbación es una matriz, $\gamma \Delta \mathbf{G}_{MSE}$, que en general no será diagonal y que, por lo tanto, introduce una cierta cantidad de interferencia de acceso múltiple MAI (Multiple Access Interference). Esta MAI aparece

²En una técnica con diversidad por código el ruido que obtenemos a la entrada del receptor es coloreado debido al efecto del banco de filtros adaptados (ver capítulo de introducción) y por lo tanto la matriz de correlación del ruido es $\mathbf{R}_r = m_r \mathbf{H}$, donde \mathbf{H} es la matriz de mezcla del canal que en este caso corresponde con la matriz de autocorrelación de los códigos de usuario.

³El lema de inversión dice que $(\mathbf{I} + \gamma (\mathbf{H}^T \mathbf{H})^{-1})^{-1} = \mathbf{I} - (\mathbf{I} + \frac{1}{\gamma} (\mathbf{H}^T \mathbf{H}))^{-1}$.

como consecuencia del equilibrio que se ha de mantener entre la amplificación del ruido y la eliminación de interferencias a la salida del sistema MIMO.

Finalmente, obsérvese que el valor mínimo del MSE no es nulo sino que tiene un determinado valor residual. Para obtener este valor podemos considerar de nuevo una SNR suficientemente alta despreciando los términos multiplicados por γ^2 y γ^3 . Es decir, sustituyendo (3.50) en (3.48) obtenemos

$$\begin{aligned}
MMSE &= M\tilde{S}E(\tilde{\mathbf{G}}_0) = (\gamma\Delta\mathbf{G}_{MSE})^T(\gamma\Delta\mathbf{G}_{MSE}) + \\
&\quad + \gamma\text{Tr}\{(\mathbf{G}_0 + \gamma\Delta\mathbf{G}_{MSE})^T(\mathbf{H}^T\mathbf{H})^{-1}(\mathbf{G}_0 + \gamma\Delta\mathbf{G}_{MSE})\} = \\
&= \gamma^2(\Delta\mathbf{G}_{MSE})^T(\Delta\mathbf{G}_{MSE}) + \gamma\text{Tr}\{\mathbf{G}_0^T(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{G}_0\} + \\
&\quad + \gamma^2\text{Tr}\{\Delta\mathbf{G}_{MSE}^T(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{G}_0\} + \gamma^2\text{Tr}\{\mathbf{G}_0^T(\mathbf{H}^T\mathbf{H})^{-1}\Delta\mathbf{G}_{MSE}\} + \\
&\quad + \gamma^3\text{Tr}\{\Delta\mathbf{G}_{MSE}^T(\mathbf{H}^T\mathbf{H})^{-1}\Delta\mathbf{G}_{MSE}\} \simeq \\
&\simeq \gamma\text{Tr}\{\mathbf{G}_0^T(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{G}_0\} = \gamma\text{Tr}\{(\mathbf{H}^T\mathbf{H})^{-1}\}
\end{aligned} \tag{3.51}$$

Es importante observar que el valor residual del MMSE depende de γ , \mathbf{G}_0 y la matriz de mezcla \mathbf{R} pero no de la perturbación $\Delta\mathbf{G}_{MSE}$.

3.7.3 Valor del MSE en los Mínimos con Ruido

Una vez calculado el mínimo valor del MSE vamos a compararlo con el que se obtiene en los mínimos de la nueva función de coste (3.13). Para ello, consideramos la función de coste del nuevo algoritmo normalizada respecto a m_s^2 para un caso general de N usuarios y una transmisión de símbolos reales recibidos con la misma potencia (es decir, $E[s_i^2(n)] = m_s \quad \forall i$). De esta forma, podemos expresar la función de coste normalizada como

$$\begin{aligned}
\frac{J}{m_s^2} &= \frac{1}{m_s^2} \sum_{i=1}^N E\{y_i^4(n)\} - 2\frac{k}{m_s} \sum_{i=1}^N E\{y_i^2(n)\} + k^2N + \\
&\quad + \frac{1}{m_s^2} \frac{\alpha}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N (E\{y_i(n)y_j(n)\})^2
\end{aligned} \tag{3.52}$$

donde $k = \frac{E[s^4(n)]}{E^2[s^2(n)]}$ es el momento normalizado de cuarto orden de los símbolos transmitidos que, por razones de simplicidad, se supone idéntico para todos los usuarios. En una situación libre de ruido $r(n) = 0$. Como muestra el apéndice D.2, sustituyendo (3.44) en (3.52) obtenemos que

$$\begin{aligned}
J(\mathbf{G}) &= \frac{J}{m_s^2} = (k-3) \sum_{i=1}^N \sum_{m=1}^N g_{im}^4 + k^2N + (3 - \frac{\alpha}{2})\text{Tr}\{(\text{diag}(\mathbf{G}^T\mathbf{G}))^2\} - \\
&\quad - 2k\text{Tr}\{\mathbf{G}^T\mathbf{G}\} + \frac{\alpha}{2}\text{Tr}\{(\mathbf{G}^T\mathbf{G})^2\}
\end{aligned} \tag{3.53}$$

donde $\text{diag}(\mathbf{G}^T \mathbf{G})$ representa una matriz diagonal cuyos elementos corresponden con los de la diagonal de $\mathbf{G}^T \mathbf{G}$. Los puntos estacionarios de J son aquellos donde el gradiente de esta función se anula que, según se muestra en el apéndice D.2, son las soluciones al siguiente sistema de ecuaciones

$$\begin{aligned} \frac{\partial J(\mathbf{G})}{\partial \mathbf{G}} &= 4(k-3)\mathbf{G} \otimes \mathbf{G} \otimes \mathbf{G} + 4\left(3 - \frac{\alpha}{2}\right)\text{diag}(\mathbf{G}^T \mathbf{G})\mathbf{G} - 4k\mathbf{G} + \\ &+ 4\frac{\alpha}{2}(\mathbf{G}\mathbf{G}^T)\mathbf{G} = \mathbf{0} \end{aligned} \quad (3.54)$$

donde $\mathbf{A} \otimes \mathbf{B}$ representa el producto individual de los elementos de la matriz \mathbf{A} con los de la matriz \mathbf{B} . Es inmediato comprobar que la solución $\mathbf{G}_0 = \mathbf{I}$ satisface el sistema de ecuaciones (3.54). En este punto además de minimizarse la función de coste del nuevo algoritmo, como se demostró en la sección anterior, también se minimiza el MSE en ausencia de ruido pues la matriz \mathbf{G}_0 es la misma obtenida para minimizar el MSE en ausencia de ruido (3.42).

Consideraremos ahora la nueva función de coste cuando existe ruido en el canal. Como se demuestra en el apéndice D.2, sustituyendo (3.44) en (3.52) se obtiene la siguiente expresión de la función de coste normalizada

$$\begin{aligned} \tilde{J}(\mathbf{G}) &= J(\mathbf{G}) + \gamma(6 - \alpha)\text{Tr} \{ \text{diag}(\mathbf{G}^T \mathbf{G}) \text{diag}(\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}) \} - \\ &- 2\gamma k \text{Tr} \{ \mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G} \} + \gamma \alpha \text{Tr} \{ \mathbf{G}^T \mathbf{G} \mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G} \} \end{aligned} \quad (3.55)$$

donde $J(\mathbf{G})$ es 3.53), la función de coste cuando no hay ruido.

Los puntos estacionarios de $\tilde{J}(\mathbf{G})$ se corresponden con los que anulan el gradiente de esta función con respecto a \mathbf{G} , es decir, las raíces del siguiente sistema de ecuaciones, tal y como se muestra en el apéndice D.2

$$\begin{aligned} \frac{\partial \tilde{J}(\mathbf{G})}{\partial \mathbf{G}} &= \frac{\partial J(\mathbf{G})}{\partial \mathbf{G}} + 2\gamma \{ (6 - \alpha) [\text{diag}(\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}) + \text{diag}(\mathbf{G}^T \mathbf{G}) (\mathbf{H}^T \mathbf{H})^{-1}] - \\ &- 2k(\mathbf{H}^T \mathbf{H})^{-1} + \alpha [\mathbf{G}\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} + (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}\mathbf{G}^T] \} \mathbf{G} = \mathbf{0} \end{aligned} \quad (3.56)$$

Si suponemos que la solución de este sistema de ecuaciones es de la forma $\tilde{\mathbf{G}}_0 \simeq \mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM}$, donde $\mathbf{G}_0 = \mathbf{I}$ es la solución al sistema (3.54), en el apéndice D.2 se obtiene que

$$\begin{aligned} \tilde{\mathbf{G}}_0 &= \mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM} = \\ &\simeq \mathbf{G}_0 + \gamma \frac{(\alpha - 6)(k - 3)}{k(2(k - 3) + \alpha)} \text{diag}((\mathbf{H}^T \mathbf{H})^{-1}) + \\ &+ \gamma \frac{1}{2} \left(1 - \frac{2\alpha}{2(k - 3) + \alpha} \right) (\mathbf{H}^T \mathbf{H})^{-1} \end{aligned} \quad (3.57)$$

Esta expresión es similar a la obtenida para la minimización del MSE ya que el ruido perturba la posición del mínimo de la función de coste en una cantidad $\gamma \Delta \mathbf{G}_{CM}$.

Esta perturbación afecta de dos formas distintas a las salidas: por una parte, aparece una matriz diagonal que varía las ganancias con que se obtienen los usuarios en las salidas, y por otra, una matriz no diagonal que introduce una cierta cantidad de MAI. De nuevo esta MAI surge para evitar que el sistema de separación amplifique el ruido del canal.

Obsérvese el mínimo valor del MSE (MMSE) obtenido en presencia de ruido, para valores suficientemente altos de la SNR (3.51), no depende de la perturbación producida en la posición de los mínimos del caso sin ruido. La presencia del ruido también produce una perturbación similar en la posición de los mínimos de la nueva función de coste (3.57). Debido a que el mínimo de esta función (3.57) y el mínimo del MSE (3.50) se diferencian únicamente en el tipo de perturbación, el valor del MMSE es el mismo para ambos casos. Por lo tanto, la inclusión de los nuevos términos de penalización no afecta al rendimiento del nuevo algoritmo, ya que en la expresión del MMSE no aparece la influencia del parámetro α . Este resultado es similar al obtenido en el capítulo anterior para el caso de un igualador basado en el algoritmo CMA y fue presentado en [23].

3.8 Simulaciones

En esta sección se presentan los resultados de diferentes simulaciones por ordenador que muestran el comportamiento del nuevo criterio y la validez de las aproximaciones realizadas en los análisis. Para analizar el comportamiento del nuevo criterio se han usado dos índices: la interferencia de acceso múltiple, MAI (Multiple Access Interference), y la relación señal a interferencia más ruido, SINR (Signal to Interference plus Noise Ratio). El primer índice, la MAI obtenida en la salida i -ésima del sistema MIMO se define como

$$MAI_i = \frac{\sum_{j=1}^N |g_{ij}|^2 P_j}{\max_j (|g_{ij}|^2 P_j)} - 1 \quad (3.58)$$

donde N es el número de usuarios, $g_{ij} = \mathbf{w}_i^H \mathbf{h}_j$ es la amplitud del usuario j en la salida i -ésima y P_j es la potencia en recepción del usuario j . Como ya se había visto, \mathbf{h}_j es la columna j -ésima de la matriz de mezcla. Este índice lo utilizaremos para evaluar el funcionamiento del algoritmo en situaciones libres de ruido. Sin embargo, cuando en el canal exista ruido, es más adecuado utilizar la SINR. La SINR del usuario i a la salida k -ésima del sistema MIMO se define como

$$SINR_i[\mathbf{w}_k] = \frac{P_i |g_{ki}|^2}{\sum_{\substack{l=1 \\ l \neq i}}^N P_l |g_{kl}|^2 + \mathbf{w}_k^H \mathbf{R}_{ruido} \mathbf{w}_k} \quad (3.59)$$

donde $\mathbf{R}_{ruido} = E[\mathbf{r}(n)\mathbf{r}^H(n)]$ es la matriz de autocorrelación del ruido.

En un primer experimento se ha considerado un sistema con diversidad espacial. Se ha utilizado una antena en array con 10 sensores linealmente equiespaciados con una separación igual a la mitad de la longitud de onda. Todos los sensores se han supuesto ideales con un diagrama de radiación omnidireccional. El entorno considera la transmisión de dos usuarios con una modulación QPSK (Quaternary Phase Shift Keying) y el ruido del sistema es prácticamente inexistente. El primer usuario alcanza la antena desde un ángulo de 40° mientras que el otro llega por uno de -20° . El segundo usuario es recibido con una potencia 10 dB inferior a la del primero. La figura 3.1 representa la evolución temporal de la MAI cuando el parámetro α es igual a 1 y μ es igual a 2×10^{-4} (valor suficiente pequeño para reducir el error de desajuste en el algoritmo adaptativo). La curva superior muestra la MAI para la salida que extrae la señal del usuario más potente, siendo la curva inferior la correspondiente a la del usuario más débil. Obsérvese como una vez alcanzada la convergencia del algoritmo, la MAI correspondiente al usuario más potente es la que alcanza un valor más pequeño porque la señal que le interfiere es mucho más débil.

En el segundo experimento se ha escogido el mismo sistema de diversidad pero considerando la presencia de ruido. De nuevo transmiten dos usuarios con una modulación QPSK. El primero llega a la antena con un ángulo de -20° y una SNR de 6 dB, mientras que el otro llega por un ángulo de 40° y una SNR de 3 dB. La figura 3.2 muestra la evolución temporal de la SINR de salida para cada uno de los usuarios cuando se considera $\alpha = 1$ y $\mu = 5 \times 10^{-4}$. Claramente se puede observar que los usuarios han sido separados y que las SINR's de salida alcanzan sus valores óptimos. Como se puede observar, el efecto del ruido no afecta a la convergencia del algoritmo que es capaz de separar los usuarios y minimizar el ruido a la salida del sistema al mismo tiempo.

Un tercer experimento considera una diversidad por código, CDMA. En el entorno existen dos usuarios transmitiendo con una modulación DS/BPSK (Direct Sequence/Binary Phase Shift Keying). Las firmas o códigos de usuario utilizados son de tipo *Kasami* [57] con una longitud de 7 chips. El vector de observaciones ha sido generado por medio de dos filtros adaptados perfectamente sincronizados a los códigos de los usuarios. El ruido se ha considerado despreciable y los usuarios se reciben con una diferencia de 5 dB en sus potencias. La figura 3.3 representa la evolución temporal de la MAI para cada usuario cuando se considera $\alpha = 1$ y $\mu = 5 \times 10^{-5}$. Claramente, la MAI decrece con el tiempo indicando el adecuado comportamiento del algoritmo. La velocidad de convergencia es menor que en el primer experimento ya que en este caso el grado de diversidad empleado es menor que en los experimentos anteriores.

Para comparar el comportamiento del nuevo criterio en presencia de ruido con uno MMSE, se han realizado experimentos de una transmisión síncrona de dos usuarios con una modulación DS/BPSK que emplea códigos *Kasami* de longitud de 15 chips.

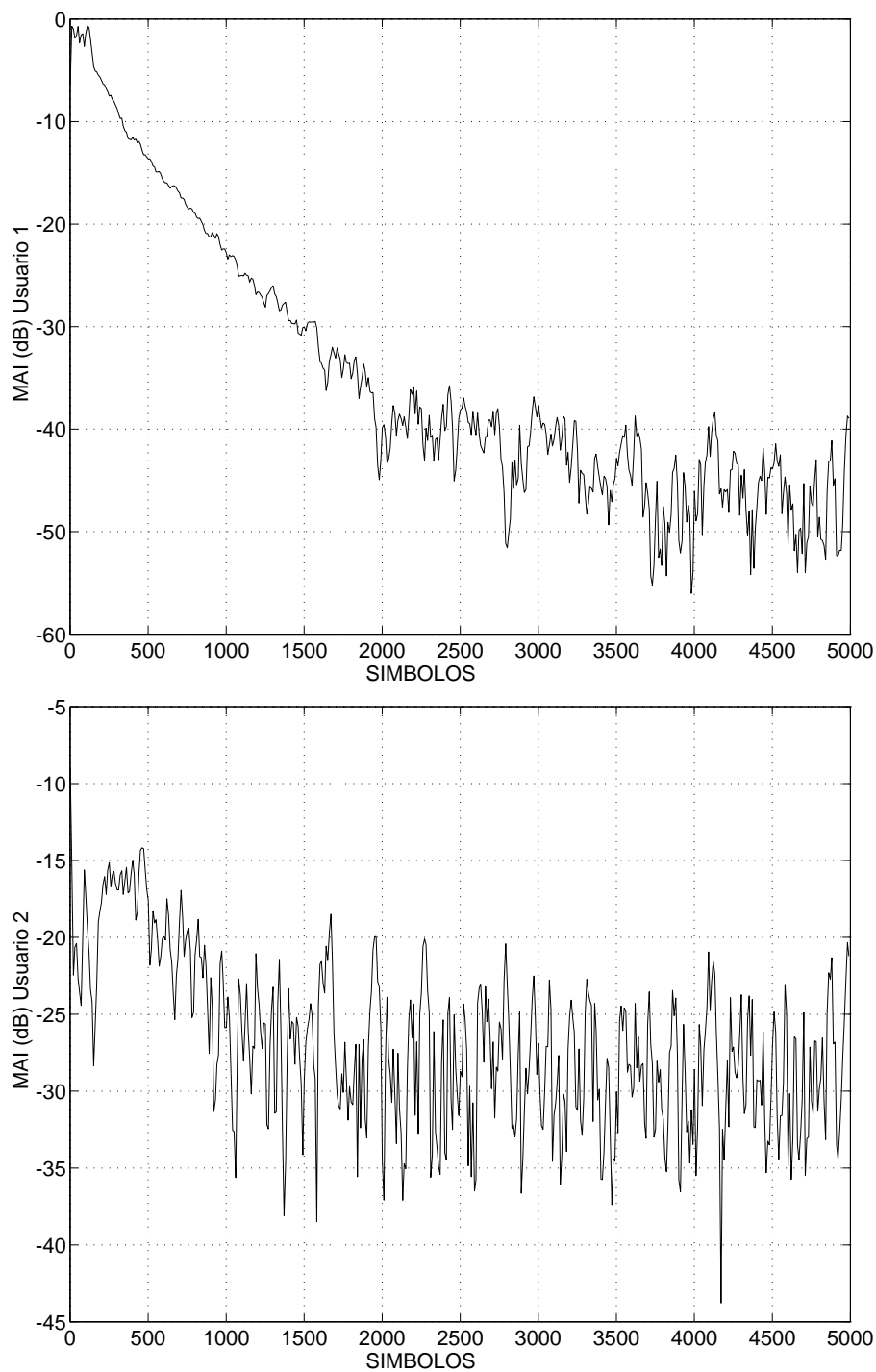


Figura 3.1: Evolución temporal de la MAI para cada una de las salidas del sistema con una antena en array de 10 sensores. La primera curva corresponde con la extracción de un usuario de potencia 10 dB superior a la del usuario de la curva inferior. ($\alpha = 1$ y $\mu = 2 \times 10^{-4}$).

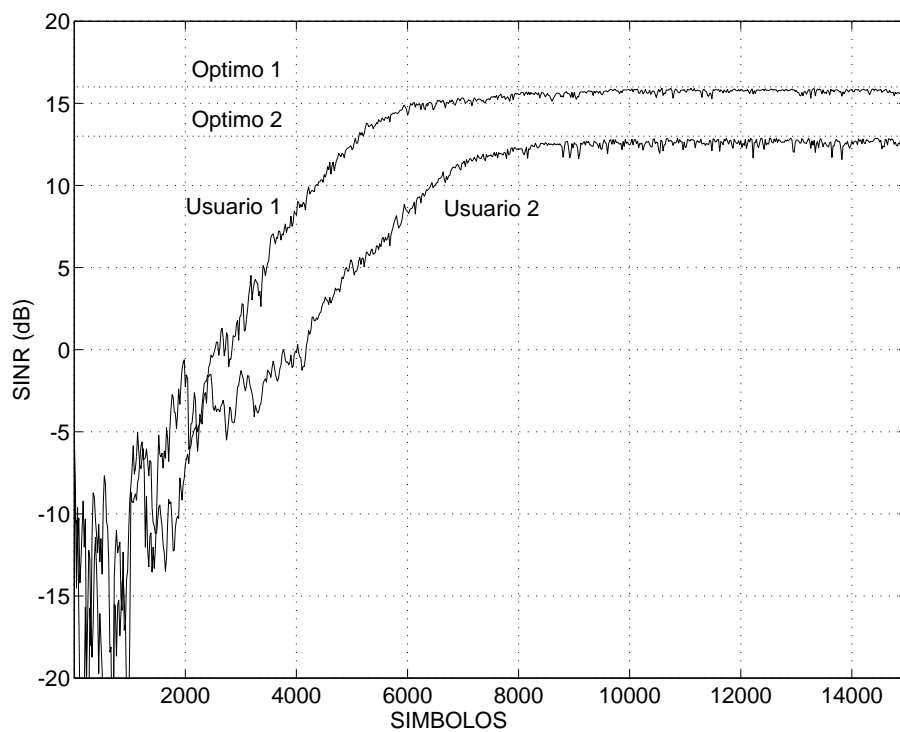


Figura 3.2: Evolución temporal de la SINR para cada una de las salidas del sistema con una antena en array de 10 sensores. SNR primer usuario 6 dB, SNR segundo usuario 3 dB. ($\alpha = 1$ y $\mu = 5 \times 10^{-4}$).

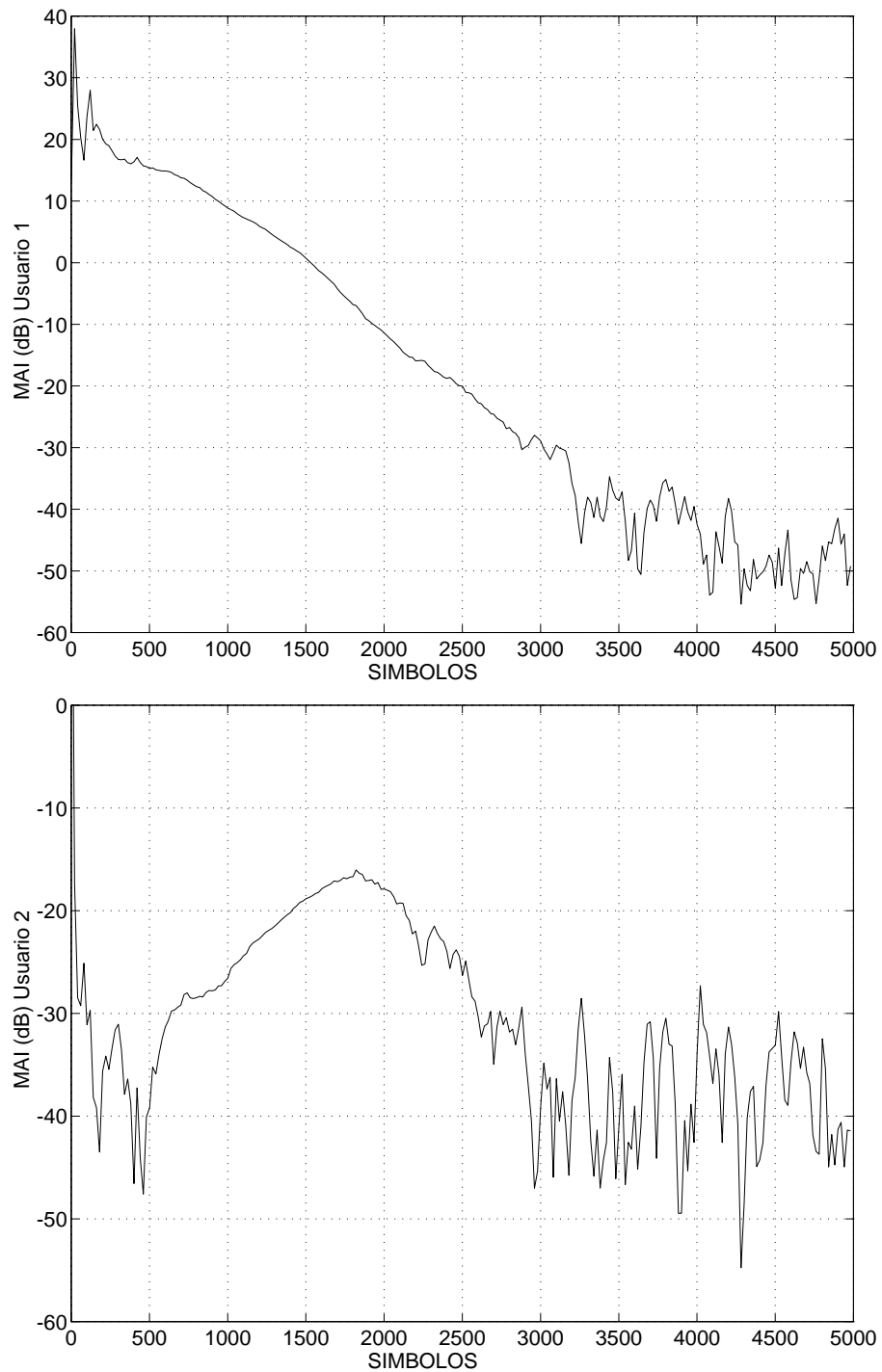


Figura 3.3: Evolución temporal de la MAI para cada uno de los usuarios de un sistema CDMA. La curva superior corresponde con la MAI de un usuario con una potencia de recepción 5 dB superior a la del usuario de la curva inferior. ($\alpha = 1$ y $\mu = 5 \times 10^{-5}$).

El criterio MMSE se ha implementado por medio de un sencillo algoritmo LMS. Los dos usuarios llegan al receptor con la misma amplitud. La figura 3.4 representa el MSE experimental frente a diferentes valores de la SNR, para el nuevo criterio y el MMSE. Como se puede observar, las curvas obtenidas son similares para valores de la SNR superiores a 5 dB. Por otro lado, la figura 3.5 presenta las probabilidades de error experimentales de estos dos criterios frente a diferentes valores de la SNR. Ambas curvas son prácticamente idénticas para cualquier valor de la SNR, lo que ilustra el buen comportamiento del nuevo criterio.

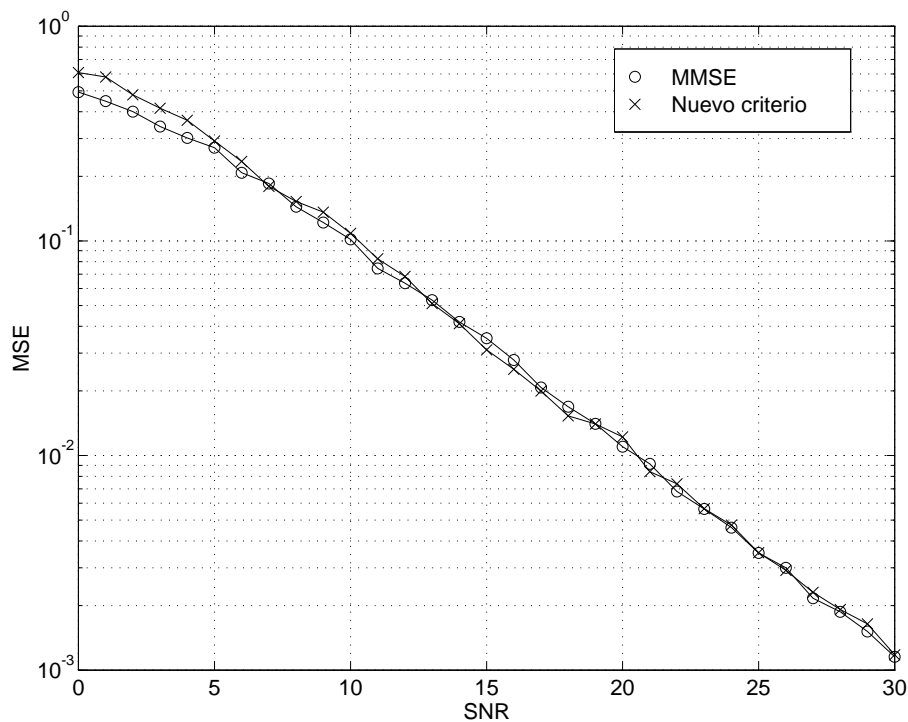


Figura 3.4: Error cuadrático medio (MSE) frente a la SNR del nuevo criterio CM y del MMSE obtenido en la recepción de uno de los dos usuarios que transmiten con una modulación DS/BPSK y un código *kasami* de 15 chips de longitud.

Finalmente, y considerando el mismo entorno de los dos últimos experimentos por ordenador, la figura 3.6 representa la evolución temporal de la SINR para uno de los dos usuarios extraído con el nuevo criterio y con el MMSE. Como se puede observar, el criterio MMSE es un poco más rápido que el basado en CM, pero ambos alcanzan el valor óptimo en un número razonable de símbolos.

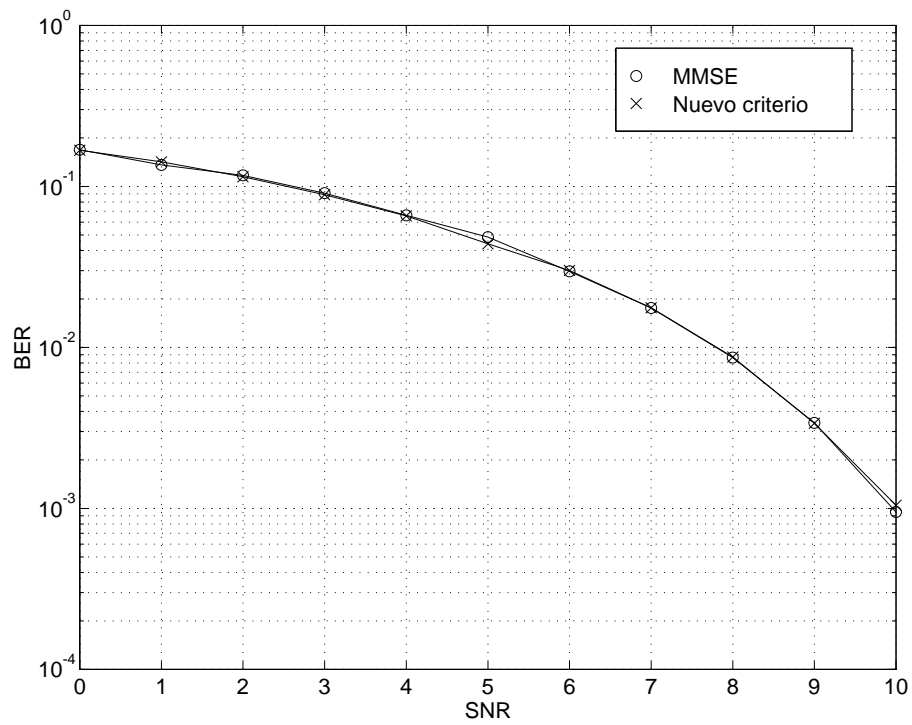


Figura 3.5: Probabilidad de error frente a la SNR alcanzada por el nuevo criterio CM y el MMSE en la recepción de uno de los dos usuarios que transmiten con una modulación DS/BPSK y códigos *kasami* de 15 chips de longitud.

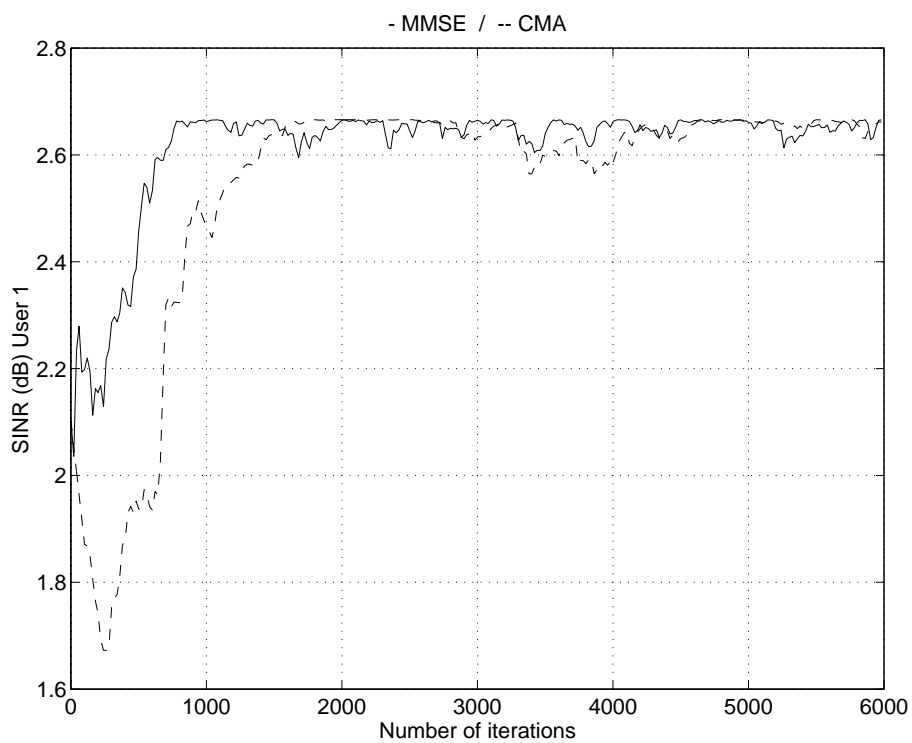


Figura 3.6: Evolución temporal de la SINR para el nuevo criterio y el MMSE para uno de los dos usuarios que transmiten con una modulación DS/BPSK y códigos *kasami* de 15 chips de longitud.

3.9 Conclusiones

En este capítulo hemos propuesto y estudiado un nuevo criterio estadístico para ajustar los parámetros de un sistema MIMO que efectúa la separación de diferentes señales en entornos de comunicaciones multiusuario. El criterio se puede considerar ciego en el sentido de que no necesita el conocimiento de las señales que están siendo transmitidas. Su planteamiento se basa en el criterio CM que es modificado añadiendo unos términos que contienen la correlación cruzada entre salidas distintas para penalizar que diferentes salidas extraigan al mismo usuario.

Para minimizar la función de coste propuesta se ha presentado un algoritmo de gradiente estocástico que ajusta los coeficientes del sistema de forma automática. Un análisis de los puntos estacionarios en un canal libre de ruido muestra que el algoritmo únicamente puede converger hacia mínimos que corresponden con la extracción de los diferentes usuarios siempre y cuando las kurtosis de las señales transmitidas sean negativas. Esta condición resulta ser poco restrictiva ya que es satisfecha por las señales comúnmente utilizadas en sistemas de comunicaciones.

También se ha realizado un análisis del comportamiento del nuevo criterio, en un canal con ruido, por medio del cálculo del MSE entre las salidas del sistema y las señales transmitidas. Bajo la suposición de disponer de una SNR suficientemente alta, 10 dB, se obtiene un valor aproximado del MSE en los mínimos de la nueva función de coste demostrando que resulta ser igual al mínimo valor del MSE (MMSE) que se puede obtener. El criterio propuesto tiene la ventaja de mantener un equilibrio entre la amplificación del ruido y la reducción de la MAI. Esta MAI residual es independiente del parámetro α que da peso a los términos cruzados y, por lo tanto, el nuevo criterio no modifica el comportamiento de un criterio CM en condiciones de ruido.

Todos los análisis de convergencia han sido apoyados por medio de simulaciones hechas por ordenador que muestran el correcto comportamiento del nuevo algoritmo en diferentes tipos de entornos.

Capítulo 4

Conclusiones y Líneas Futuras de Trabajo

4.1 Conclusiones

Esta tesis presenta nuevos criterios para ajustar los coeficientes de sistemas receptores en dos aplicaciones diferentes: en comunicaciones monousuario con canales dispersivos que introducen ISI y en comunicaciones multiusuario con canales no dispersivos que presentan MAI. Los nuevos criterios están basados en el conocido criterio Módulo Constante (CM). Este criterio emplea una función de coste no convexa que puede presentar, en determinadas situaciones, mínimos no deseados. Un algoritmo adaptativo que trate de minimizar esta función de coste podría converger hacia un mínimo no deseado causando un mal comportamiento del sistema receptor.

Los nuevos criterios utilizan esta función de coste como punto de partida y la modifican para eliminar los mínimos no deseados sin alterar aquéllos que se denominan globales o deseados. La modificación consiste en añadir un conjunto de términos, convenientemente ponderados mediante una constante, que hacen desaparecer los mínimos no deseados.

La inclusión de nuevos términos en la función de coste para penalizar la convergencia de los algoritmos hacia soluciones no deseadas es un mecanismo muy similar a la incorporación de restricciones sobre el vector de coeficientes. De hecho, las funciones de coste que se proponen pueden interpretarse como lagrangianos asociados a problemas de optimización con restricciones. La existencia de restricciones, sin embargo, complica notablemente los algoritmos adaptativos porque obliga a plantear dos etapas: una para la optimización de la función de coste y otra para imponer las restricciones. Nuestra aproximación presenta la ventaja de que da lugar a algoritmos adaptativos convencionales muy sencillos.

En el capítulo de introducción se presenta el modelo de señal utilizado durante toda la tesis mostrándose que es válido para diferentes tipos de aplicaciones. En todas ellas se considera un *front-end* de comunicaciones que proporciona un vector de observaciones que resulta ser una combinación lineal de los símbolos transmitidos por uno o más usuarios. Ese vector se puede escribir como

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{r}(n) \quad (4.1)$$

donde \mathbf{H} es la matriz de mezcla del canal. Las limitaciones de los algoritmos para extraer los símbolos transmitidos del vector de observaciones dependen en buena parte del condicionamiento de esta matriz. La estructura de \mathbf{H} y la composición del vector de símbolos transmitidos, $\mathbf{s}(n)$, depende del entorno de comunicaciones considerado y, por ello, habrá que plantear diferentes criterios de optimización para cada caso.

4.1.1 Igualación de Canales

El capítulo 2 está dedicado a la igualación de canales de respuesta al impulso infinita, IIR, representados por un modelo auto-regresivo, AR. En estos casos, los igualadores CMA pueden presentar un mal comportamiento cuando su número de coeficientes es finito ya que el igualador no elimina completamente la ISI. Este mal comportamiento se debe a que la función de coste CM presenta mínimos locales no deseados hacia los que el algoritmo adaptativo puede converger. Por ello, en este capítulo se presenta un nuevo criterio basado en el CM que elimina el mal comportamiento en este tipo de canales.

Los resultados más significativos presentados en este capítulo son los siguientes:

- Se propone ajustar los coeficientes de un igualador para minimizar la siguiente función de coste

$$J = E[||y(n)|^2 - R_2|^2] + \alpha \sum_{l=L-r}^{L-1} |E[y(n)x^*(n-l)]|^2 \quad (4.2)$$

donde α es el parámetro que regula el peso de los términos que penalizan la extracción de mínimos no deseados, L es el tamaño del igualador y r es el orden máximo que posee el canal AR. Obsérvese que la función de coste resultante se diferencia de la del algoritmo CMA en los términos multiplicados por el parámetro α . Estos nuevos términos penalizan la extracción de señales que presentan alguna correlación con las señales de entrada en instantes pasados ya que éstas presentan el mismo tipo de interferencia intersimbólica que la ISI residual obtenida en los mínimos no deseados.

- Se deriva un algoritmo de gradiente para minimizar (4.2). Este algoritmo es muy sencillo, permitiendo el ajuste adaptativo de los coeficientes del igualador por medio de hardware poco complejo.

- Se demuestra la capacidad del nuevo criterio para eliminar completamente la ISI introducida por canales AR. El análisis de los puntos estacionarios muestra que la función (4.2) posee únicamente mínimos globales deseados que corresponden con la completa eliminación de la ISI, si y sólo si se garantiza que el parámetro α es mayor que cierto límite inferior y que la kurtosis de los símbolos transmitidos es negativa. Esta última imposición no es muy restrictiva y es la misma que se obtiene para garantizar el buen comportamiento de un igualador CMA de longitud doblemente infinita.
- Se demuestra que el error cuadrático medio (MSE) entre la salida del igualador y los símbolos deseados cuando la SNR es suficientemente alta ($SNR \geq 10$ dB) coincide con la mínima alcanzable (MMSE). Esta aproximación permite suponer que el ruido sólo introduce una perturbación en la posición de los mínimos de la función de coste. Esta perturbación introduce una ISI residual consecuencia del compromiso existente entre la reducción de la ISI y la amplificación del ruido.
- Finalmente, los resultados presentados en este capítulo están apoyados con simulaciones por ordenador que permiten ilustrar el comportamiento del algoritmo adaptativo obtenido y de las aproximaciones realizadas en los análisis. Estas simulaciones también han servido para comparar los algoritmos resultantes con otros ya existentes en la bibliografía.

4.1.2 Separación de Señales

El capítulo 3 considera un receptor MIMO para sistemas de comunicaciones multiusuario en canales no dispersivos. El sistema MIMO ha de separar las señales de los usuarios que están inmersas en una mezcla instantánea producida por el canal de comunicaciones. El criterio de optimización en el que se basa el capítulo es, al igual que en el anterior, el de módulo constante, CM.

Un sistema MIMO que se interprete como múltiples sistemas MISO cada uno de los cuales se adapta con un algoritmo CMA independiente puede presentar un problema de captura en el que se obtenga la misma señal en distintas salidas. Para evitar este problema se propone un nuevo problema de optimización que introduce cierto grado de acoplamiento entre los algoritmos que ajustan cada una de las salidas.

Los resultados más significativos presentados en este capítulo son los siguientes:

- Se define una función de coste cuya minimización corresponde con la extracción de los diferentes usuarios en las salidas del sistema.

$$J = \sum_{i=1}^N E[||y_i(n)|^2 - R_2|^2] + \frac{\alpha}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N |E[y_i(n)y_j^*(n)]|^2 \quad (4.3)$$

donde α es el parámetro que regula el peso de los términos añadidos. Estos términos penalizan la extracción en diferentes salidas de señales correladas entre sí y, debido a la independencia estadística de las señales de los diferentes usuarios, la extracción de un mismo usuario en varias salidas.

- Se deriva un algoritmo de gradiente para minimizar (4.3). Este algoritmo es muy sencillo, permitiendo el ajuste adaptativo de los coeficientes del sistema por medio de un hardware poco complejo.
- Se demuestra que el nuevo criterio no presenta el problema de captura. Para ello, se realiza un análisis de los puntos estacionarios de la función (4.3) con sólo dos usuarios. Este análisis muestra que los únicos mínimos de la función de coste corresponden con la extracción de diferentes usuarios a la salida del sistema MIMO siempre que se garantice que $\alpha > 0$ y la kurtosis de las señales transmitidas sea negativa. Esta última condición no es muy restrictiva pues la mayoría de las modulaciones empleadas en comunicaciones digitales la cumplen y es la obtenida por otros autores para garantizar el correcto funcionamiento de un criterio CMA para la extracción de un usuario en un sistema MISO.
- Se realiza un análisis en presencia de ruido para entornos con relaciones SNR relativamente altas ($SNR \geq 10$ dB). Este análisis revela que los mínimos de la función de coste son movidos debido a una perturbación introducida por el ruido. Esta perturbación permite alcanzar un compromiso entre la reducción de la MAI y la amplificación del ruido. Además, un estudio del MSE en estos puntos muestra que éste coincide con el obtenido para un criterio MMSE.
- Finalmente, simulaciones por ordenador apoyan la validez del nuevo criterio y de las aproximaciones realizadas en el análisis en diferentes tipos de entorno. Además, estas simulaciones permiten comparar la convergencia del algoritmo presentado y su buen funcionamiento respecto a otras aproximaciones de la bibliografía.

4.2 Líneas Futuras de Trabajo

Las posibles líneas de investigación que pueden surgir a partir de los resultados de esta tesis son las siguientes:

- **Extensión del Nuevo Criterio de Igualación para Canales más Generales:** El criterio presentado en el capítulo 2 considera canales de tipo auto-regresivo, AR. Aunque este tipo de canales pueden modelar el comportamiento de canales de comunicaciones utilizados en la práctica (como los de microondas), la consideración de tipos de canales más generales extendería la utilidad de este nuevo criterio.

Los canales a considerar serán MA (Moving Average) y ARMA (Auto-Regressive Moving Average). Es necesario estudiar el efecto de los nuevos términos de la función de coste en este tipo de canales y cómo se podrían modificar éstos para eliminar o reducir los efectos de los mínimos locales dependientes de la longitud (LDLM).

- **Nuevos Criterios de Separación de Señales por Medio de la Maximización de la Información:** El capítulo 3 considera un sistema MIMO para la separación de usuarios. Este es un problema de separación de fuentes en el que las señales de los usuarios son estadísticamente independientes y están involucradas en una mezcla desconocida. Interpretando el sistema MIMO como la parte lineal de una red neuronal no lineal de una sola capa, la maximización de la transferencia de información para determinadas no linealidades a la salida del sistema MIMO es equivalente a la minimización de la función de coste CM con nuevos términos cruzados similares a los del capítulo 3 que garantizan la extracción de usuarios diferentes a la salida del sistema. Por ejemplo, Castedo [11] utilizando la no linealidad

$$g(x) = \int_{-\infty}^x e^{-(t^2-1)^2} dt \quad (4.4)$$

a la salida del sistema MIMO, obtiene una nueva función de coste que maximiza la transferencia de información de la red neuronal y que es de la siguiente forma

$$J = \sum_{i=1}^N E[|y_i(n)|^2 - 1|^2] - \ln|\det(\mathbf{W})| \quad (4.5)$$

donde \mathbf{W} es la matriz de coeficientes del sistema MIMO. Obsérvese cómo el término $-\ln|\det(\mathbf{W})|$ juega el mismo papel que los términos cruzados en (4.3): cuando varias salidas extraen el mismo usuario, la matriz \mathbf{W} es singular y el término $-\ln|\det(\mathbf{W})|$ tiende a infinito. Esto impide la convergencia a este tipo de soluciones.

- **Diseño de Receptores MIMO para Sistemas de Comunicaciones Multiusuario en Canales Dispersivos:** Básicamente, esta idea consiste en unir los conceptos de los capítulo 2 y 3. Como se vió en el primer capítulo de la tesis, cuando se considera un sistema multiusuario en canales dispersivos, el vector de entrada al sistema consiste en la mezcla de varios símbolos de cada uno de los usuarios del sistema. El criterio presentado en el capítulo 3 es válido únicamente para mezclas instantáneas y al aplicarlo a este caso el sistema podría extraer la misma secuencia de símbolos de un usuario con diferentes retardos, debido a la hipótesis de independencia estadística entre los símbolos de un mismo usuario. La extensión de la función de coste del capítulo 3 tiene que considerar que los términos de correlación entre las salidas no deben limitarse únicamente a

los instantes actuales sino considerar también las correlaciones entre diferentes instantes de las salidas del sistema. Una posible función de coste propuesta por Papadias [50] sería de la forma

$$J = \sum_{i=1}^N E[|y_i(n)|^2 - R_2]^2 + \frac{\alpha}{2} \sum_{l=0}^{P-1} \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N |E[y_i(n)y_j^*(n-l)]|^2 \quad (4.6)$$

donde $P-1$ sería el máximo retardo posible entre las secuencias de salida. Esta nueva función de coste ha añadido unos términos que impiden la extracción de la misma señal con diferentes retardos en múltiples salidas del sistema MIMO. Obsérvese que cuando este retardo es nulo ($P-1 = 0$) la función (4.6) es idéntica a la nueva función de coste correspondiente al criterio CRIMNO presentado en la sección 2.6.2. Esta función, sin embargo, no está libre de mínimos locales LDLM.

- **Mejora de la velocidad de convergencia:** Los algoritmos que hemos escogido para implementar los nuevos criterios propuestos en la tesis son de tipo gradiente. Estos algoritmos tienen la ventaja de ser extremadamente simples y robustos pero su velocidad de convergencia puede resultar insuficiente en determinadas aplicaciones. En estos casos será necesario desarrollar otros algoritmos basados en técnicas de optimización distintas como gradiente conjugado, *Least-Squares*, etc. No obstante, es conveniente destacar que el aumento de velocidad de convergencia se conseguirá a costa de estabilidad numérica.

Apéndice A

Diferenciación con Respecto a Vectores y Matrices

El cálculo de los puntos estacionarios de una función de coste, J , como los presentadas en esta tesis, lleva asociado la diferenciación de esta función con respecto a un escalar, un vector o una matriz de parámetros. Este apéndice presenta las bases de la diferenciación de una función con respecto a este tipo de variables (que inicialmente consideraremos complejas).

A.1 Diferenciación con Respecto a un Escalar

Sea $f(z)$ una función real de la variable compleja $z = x + jy$. La derivada de $f(z)$ respecto a z y z^* se define como

$$\begin{aligned}\frac{\partial f(z)}{\partial z} &= \frac{\partial f(z)}{\partial x} - j \frac{\partial f(z)}{\partial y} \\ \frac{\partial f(z)}{\partial z^*} &= \frac{\partial f(z)}{\partial x} + j \frac{\partial f(z)}{\partial y}\end{aligned}\tag{A.1}$$

Obsérvese que de acuerdo con estas definiciones se obtiene que

$$\begin{aligned}\frac{\partial z}{\partial z} &= \frac{\partial z^*}{\partial z^*} = 2 \\ \frac{\partial z}{\partial z^*} &= \frac{\partial z^*}{\partial z} = 0\end{aligned}\tag{A.2}$$

A.2 Diferenciación con Respecto a un Vector

Suponiendo que $f(\mathbf{z})$ es una función real dependiente de una variable consistente en un vector complejo de tamaño $M \times 1$, $\mathbf{z} = [x_1 + jy_1, \dots, x_M + jy_M]^T$, se define la derivada de la función con respecto al vector como

$$\frac{\partial f(\mathbf{z})}{\partial \mathbf{z}} = \begin{pmatrix} \frac{\partial f(\mathbf{z})}{\partial x_1} - j \frac{\partial f(\mathbf{z})}{\partial y_1} \\ \vdots \\ \frac{\partial f(\mathbf{z})}{\partial x_M} - j \frac{\partial f(\mathbf{z})}{\partial y_M} \end{pmatrix} \quad (\text{A.3})$$

y de forma similar su derivada conjugada es

$$\frac{\partial f(\mathbf{z})}{\partial \mathbf{z}^*} = \begin{pmatrix} \frac{\partial f(\mathbf{z})}{\partial x_1} + j \frac{\partial f(\mathbf{z})}{\partial y_1} \\ \vdots \\ \frac{\partial f(\mathbf{z})}{\partial x_M} + j \frac{\partial f(\mathbf{z})}{\partial y_M} \end{pmatrix} \quad (\text{A.4})$$

Obsérvese que debido a estas definiciones, se obtiene que

$$\frac{\partial \mathbf{z}^T}{\partial \mathbf{z}} = \frac{\partial \mathbf{z}^H}{\partial \mathbf{z}^*} = 2\mathbf{I} \quad (\text{A.5})$$

$$\frac{\partial \mathbf{z}^T}{\partial \mathbf{z}^*} = \frac{\partial \mathbf{z}^H}{\partial \mathbf{z}} = \mathbf{0} \quad (\text{A.6})$$

donde \mathbf{I} y $\mathbf{0}$ son las matrices identidad y nula, respectivamente. De acuerdo con estas definiciones, el operador gradiente complejo con respecto al vector $\mathbf{z} = [z_1, \dots, z_M]^T$ se define como $\nabla_{\mathbf{z}} = [\partial/\partial z_1^*, \dots, \partial/\partial z_M^*]^T$.

A.2.1 Ejemplos

- $f = \mathbf{c}^T \mathbf{z}$ donde \mathbf{c} es un vector constante de tamaño $M \times 1$

$$\frac{\partial f}{\partial \mathbf{z}^*} = \frac{\partial(\mathbf{c}^T \mathbf{z})}{\partial \mathbf{z}^*} = \frac{\partial(\mathbf{z}^T \mathbf{c})}{\partial \mathbf{z}^*} = \mathbf{0} \quad (\text{A.7})$$

- $f = \mathbf{z}^H \mathbf{c}$ donde \mathbf{c} es un vector constante de tamaño $M \times 1$

$$\frac{\partial f}{\partial \mathbf{z}^*} = \frac{\partial(\mathbf{z}^H \mathbf{c})}{\partial \mathbf{z}^*} = 2\mathbf{c} \quad (\text{A.8})$$

- $f = \mathbf{z}^H \mathbf{R} \mathbf{z}$ donde \mathbf{R} es una matriz hermítica de tamaño $M \times M$.

$$\frac{\partial f}{\partial \mathbf{z}^*} = \frac{\partial(\mathbf{z}^H \mathbf{R} \mathbf{z})}{\partial \mathbf{z}^*} = 2\mathbf{R} \mathbf{z} \quad (\text{A.9})$$

Caso Real. Si \mathbf{z} es real, entonces se cumple que

$$\begin{aligned}\frac{\partial \mathbf{z}^T}{\partial \mathbf{z}} &= \mathbf{I} \\ \frac{\partial \mathbf{z}^T \mathbf{c}}{\partial \mathbf{z}} &= \mathbf{c} \\ \frac{\partial \mathbf{z}^T \mathbf{R} \mathbf{z}}{\partial \mathbf{z}} &= 2\mathbf{R} \mathbf{z}\end{aligned}\tag{A.10}$$

donde \mathbf{R} es una matriz simétrica.

A.3 Diferenciación con respecto a una matriz

En este apartado se considera la derivada de una función escalar, $f(\mathbf{Z})$ de una matriz \mathbf{Z} . La derivada de dicha función con respecto a \mathbf{Z} es

$$\frac{\partial f(\mathbf{Z})}{\partial \mathbf{Z}} = \begin{pmatrix} \frac{\partial f(\mathbf{Z})}{\partial z_{11}} & \cdots & \frac{\partial f(\mathbf{Z})}{\partial z_{1M}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{Z})}{\partial z_{M1}} & \cdots & \frac{\partial f(\mathbf{Z})}{\partial z_{MM}} \end{pmatrix}\tag{A.11}$$

donde z_{ij} es el elemento (i, j) de la matriz \mathbf{Z} que puede ser real o complejo. Esta derivada también puede ser expresada con respecto a vectores de la siguiente forma

$$\frac{\partial f(\mathbf{Z})}{\partial \mathbf{Z}} = \left[\frac{\partial f(\mathbf{Z})}{\partial \mathbf{z}_1}, \dots, \frac{\partial f(\mathbf{Z})}{\partial \mathbf{z}_M} \right]\tag{A.12}$$

donde \mathbf{z}_i es la i -ésima columna de \mathbf{Z} .

Debido a que en los capítulos de la tesis este tipo de derivadas son siempre con respecto a matrices reales, en los siguientes ejemplos se considera sólo el caso de \mathbf{Z} como una matriz de elementos reales.

A.3.1 Ejemplos

- $$f(\mathbf{Z}) = \text{Tr}\{\mathbf{Z}^T \mathbf{Z}\} = \sum_{i=1}^M \mathbf{z}_i^T \mathbf{z}_i$$

$$\begin{aligned}\frac{\partial f(\mathbf{Z})}{\partial \mathbf{Z}} &= \left[\frac{\partial f(\mathbf{Z})}{\partial \mathbf{z}_1}, \dots, \frac{\partial f(\mathbf{Z})}{\partial \mathbf{z}_M} \right] = \\ &= [2\mathbf{z}_1, \dots, 2\mathbf{z}_M] = 2\mathbf{Z}\end{aligned}\tag{A.13}$$

- $f(\mathbf{Z}) = \text{Tr}\{\mathbf{Z}^T \mathbf{R} \mathbf{Z}\} = \sum_{i=1}^M \mathbf{z}_i^T \mathbf{R} \mathbf{z}_i$ donde \mathbf{R} es una matriz simétrica.

$$\begin{aligned} \frac{\partial f(\mathbf{Z})}{\partial \mathbf{Z}} &= \left[\frac{\partial f(\mathbf{Z})}{\partial \mathbf{z}_1}, \dots, \frac{\partial f(\mathbf{Z})}{\partial \mathbf{z}_M} \right] = \\ &= [2\mathbf{R}\mathbf{z}_1, \dots, 2\mathbf{R}\mathbf{z}_M] = 2\mathbf{R}\mathbf{Z} \end{aligned} \quad (\text{A.14})$$

Apéndice B

Puntos Estacionarios de un Igualador Infinito CMA

Como se vió en la sección (2.5), la función de coste de un igualador CMA doblemente infinito puede ser expresado en función de las ganancias g_i , de la siguiente forma

$$J = k \sum_{i=-\infty}^{\infty} g_i^4 + 3 \sum_{j=-\infty}^{\infty} \sum_{\substack{l=-\infty \\ j \neq l}}^{\infty} g_j^2 g_l^2 - 2k \sum_{i=-\infty}^{\infty} g_i^2 + k^2 \quad (\text{B.1})$$

En el caso de este tipo de igualadores y como se vió también en la sección (2.5), los puntos estacionarios de J son los que satisfacen el siguiente sistema de ecuaciones

$$\nabla_{\mathbf{g}} J = \mathbf{0} \iff \frac{\partial J}{\partial g_i} = 0 \quad \forall i \quad (\text{B.2})$$

y por lo tanto

$$\begin{aligned} \frac{\partial J}{\partial g_i} &= 4k g_i^3 + 12g_i \sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} g_j^2 - 4k g_i = \\ &= 4g_i (k g_i^2 + 3 \sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} g_j^2 - k) = 0 \end{aligned} \quad (\text{B.3})$$

Para analizar la naturaleza de los puntos estacionarios que satisfacen la anterior ecuación, es necesario estudiar el carácter definido positivo de la matriz hessiana de

J que es de la forma

$$\mathcal{H} = \begin{pmatrix} \frac{\partial^2 J}{\partial g_{-\infty} \partial g_{-\infty}} & \cdots & \frac{\partial^2 J}{\partial g_{-\infty} \partial g_0} & \cdots & \frac{\partial^2 J}{\partial g_{-\infty} \partial g_{\infty}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 J}{\partial g_0 \partial g_{-\infty}} & \cdots & \frac{\partial^2 J}{\partial g_0 \partial g_0} & \cdots & \frac{\partial^2 J}{\partial g_0 \partial g_{\infty}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 J}{\partial g_{\infty} \partial g_{-\infty}} & \cdots & \frac{\partial^2 J}{\partial g_{\infty} \partial g_0} & \cdots & \frac{\partial^2 J}{\partial g_{\infty} \partial g_{\infty}} \end{pmatrix} \quad (\text{B.4})$$

donde

$$\begin{aligned} \frac{\partial^2 J}{\partial g_i \partial g_i} &= 12k g_i^2 + 12 \sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} g_j^2 - 4k \\ \frac{\partial^2 J}{\partial g_i \partial g_j} &= 24g_i g_j \quad \forall j \neq i \end{aligned} \quad (\text{B.5})$$

Debido a que el sistema de ecuaciones (B.3) no es lineal, surgen múltiples soluciones que pueden ser agrupadas en tres grupos:

Grupo 1 : $g_i \neq 0$ y $g_j = 0 \quad \forall j \neq i$. Este es un punto deseado de J que extrae el símbolo $s(n-i)$. De (B.3) se obtiene que este punto es

$$g_i^2 = 1 \quad (\text{B.6})$$

La matriz hessiana resultante es diagonal con el i -ésimo elemento positivo e igual a $8k$, y con los demás elementos iguales a $4(3-k)$. Por lo tanto, para garantizar que la matriz hessiana sea definida positiva y que este punto sea un mínimo, se ha de cumplir que $k < 3$, es decir, que los símbolos transmitidos tengan kurtosis negativa (la kurtosis de una señal real es igual a $k-3$).

Grupo 2 : $g_i = 0 \quad \forall i$. Este es un punto donde el igualador no extrae ningún símbolo.

En este caso el hessiano es una matriz diagonal con todos sus elementos igual a $-4k$ y, por tanto, definida negativa. Este punto es siempre un máximo de la función J .

Grupo 3 : $g_i \neq 0$ y $g_j \neq 0$ para uno o varios valores de $j \neq i$. Considerando que

$$g_i \frac{\partial^2 J}{\partial g_i \partial g_i} = \frac{\partial J}{\partial g_i} + 8k g_i^3 \quad (\text{B.7})$$

Una submatriz del hessiano de tamaño 2×2 será de la forma

$$\mathcal{H} = \begin{pmatrix} 8k g_i^2 & 24g_i g_j \\ 24g_i g_j & 8k g_j^2 \end{pmatrix} \quad (\text{B.8})$$

que posee un determinante $64(g_i^2 g_j^2 (k^2 - 9))$. Para que este determinante no sea positivo y este punto no sea mínimo de J , es suficiente con garantizar que $k - 3 < 0$, es decir que la kurtosis de los símbolos transmitidos sea negativa.

La conclusión que se extrae al análisis anterior es que un igualador CMA doblemente infinito siempre converge a soluciones donde se elimina la ISI completamente si y sólo si la kurtosis de los símbolos transmitidos es negativa [25].

Apéndice C

Desarrollos Correspondientes al Capítulo 2

C.1 Desarrollo de la Ecuación (2.47)

Considerando la suposición de la no existencia de ruido en el canal, la salida del igualador obtenida en (2.8) es

$$y(n) = \mathbf{g}^T \mathbf{s}(n) = \sum_{i=0}^{\infty} g_i s(n-i) \quad (\text{C.1})$$

Con la hipótesis de que los símbolos transmitidos son i.i.d., las esperanzas estadísticas de (2.46) pueden ser expresadas de la siguiente forma

$$\begin{aligned} E[y^2(n)] &= E[\mathbf{g}^T \mathbf{s}(n) \mathbf{s}^T(n) \mathbf{g}] = \mathbf{g}^T E[\mathbf{s}(n) \mathbf{s}^T(n)] \mathbf{g} = \\ &= m \mathbf{g}^T \mathbf{I} \mathbf{g} = m \sum_{i=0}^{\infty} g_i^2 \\ E[y^4(n)] &= E[(\mathbf{g}^T \mathbf{s}(n))^4] = E[(\sum_{i=0}^{\infty} g_i s(n-i))^4] = \\ &= E[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{u=0}^{\infty} \sum_{l=0}^{\infty} g_i g_j g_u g_l s(n-i) s(n-j) s(n-u) s(n-l)] \\ &= \begin{cases} i=j=u=l & \Rightarrow \sum_{i=0}^{\infty} g_i^4 q \\ \left\{ \begin{array}{ll} i=j & u=l \\ i=u & j=l \\ i=l & u=j \end{array} \right\} & \Rightarrow 3 \sum_{j=0}^{\infty} \sum_{\substack{l=0 \\ l \neq j}}^{\infty} g_j^2 g_l^2 m^2 \end{cases} \end{aligned}$$

$$\begin{aligned}
&= q \sum_{i=0}^{\infty} g_i^4 + 3m^2 \sum_{j=0}^{\infty} \sum_{\substack{l=0 \\ j \neq l}}^{\infty} g_j^2 g_l^2 = \\
&= m^2 \left(k \sum_{i=0}^{\infty} g_i^4 + 6 \sum_{j=0}^{\infty} \sum_{l=j+1}^{\infty} g_j^2 g_l^2 \right) \\
(E[y(n)x(n-L+1)])^2 &= (E[\mathbf{g}^T \mathbf{s}(n)x(n-L+1)])^2 \\
&= \left(E \left[\sum_{i=0}^{\infty} g_i s(n-i) \sum_{j=0}^{\infty} h(j) s(n-j-L+1) \right] \right)^2 = \\
&= \left(\sum_{i=0}^{\infty} m g_{i+L-1} h(i) \right)^2 \tag{C.2}
\end{aligned}$$

donde \mathbf{I} representa la matriz identidad y $q = E[s^4(n)]$ es el momento de cuarto orden de $s(n)$. Particularizando (2.28) para un igualador de 2 coeficientes ($L = 2$) se obtiene la relación $g_i = a^{i-1} g_1 \forall i \geq 1$. Por ello, los promedios de (C.2) se pueden expresar como sigue

$$\begin{aligned}
\frac{1}{m} E[y^2(n)] &= g_0^2 + g_1^2 \sum_{i=1}^{\infty} (a^{i-1})^2 \\
\frac{1}{m^2} E[y^4(n)] &= k \left(g_0^4 + g_1^4 \sum_{i=1}^{\infty} (a^{i-1})^4 \right) \\
&\quad + 6g_1^2 \left(g_0^2 \sum_{i=1}^{\infty} (a^{i-1})^2 + g_1^2 \sum_{j=1}^{\infty} \sum_{l=j+1}^{\infty} (a^{j-1})^2 (a^{l-1})^2 \right) \\
\frac{1}{m^2} (E[y(n)x(n-1)])^2 &= g_1^2 \left(\sum_{i=1}^{\infty} (a^{i-1})^2 \right)^2 \tag{C.3}
\end{aligned}$$

Sustituyendo (C.3) en (2.46) obtenemos la ecuación (2.47)

$$J(g_0, g_1) = k[g_0^4 + g_1^4 S_2] + 6g_1^2[g_0^2 S_1 + g_1^2 S_3] - 2k[g_0^2 + g_1^2 S_1] + k^2 + \alpha[g_1^2 S_1^2] \tag{C.4}$$

Las constantes S_1, S_2 y S_3 son las sumas de unas series que dependen de la respuesta impulsional del canal como sigue

$$\begin{aligned}
S_1 &= \sum_{i=1}^{\infty} (a^{i-1})^2 = \frac{1}{1-a^2} \\
S_2 &= \sum_{i=1}^{\infty} (a^{i-1})^4 = \frac{1}{1-a^4} \\
S_3 &= \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} (a^{i-1})^2 (a^{j-1})^2 = \frac{a^2}{(1+a^2)(1-a^2)^2} \tag{C.5}
\end{aligned}$$

C.2 Desarrollo de la ecuación (2.77)

Cuando existe un ruido aditivo en el canal, la salida del igualador (2.8) se puede representar por la suma de una componente de señal $y_s(n)$ y una de ruido $y_r(n)$, es decir, $y(n) = y_s(n) + y_r(n)$ donde

$$\begin{aligned} y_s(n) &= \mathbf{g}^T \mathbf{s}(n) = \sum_{i=0}^{\infty} g_i s(n-i) \\ y_r(n) &= \mathbf{w}^T \mathbf{r}(n) = \sum_{i=0}^{L-1} w_i r(n-i) \end{aligned} \quad (\text{C.6})$$

El ruido es considerado gaussiano, real, de media cero, con una matriz de correlación $E[\mathbf{r}(n)\mathbf{r}^T(n)] = \frac{N_0}{2}\mathbf{I}$ (\mathbf{I} es la matriz identidad) y estadísticamente independiente de los símbolos transmitidos. Dado que los símbolos transmitidos y el ruido son estadísticamente independientes entre sí, también lo son $y_s(n)$ e $y_r(n)$ y podemos desarrollar las esperanzas estadísticas de (2.70) como

$$\begin{aligned} E[y^2(n)] &= E[y_s^2(n)] + E[y_r^2(n)] \\ E[y^4(n)] &= E[y_s^4(n)] + E[y_r^4(n)] + 6E[y_s^2(n)]E[y_r^2(n)] \\ (E[y(n)x(n-L+1)])^2 &= (E[y_s(n)x(n-L+1)])^2 + (E[y_r(n)x(n-L+1)])^2 + \\ &\quad + 2E[y_s(n)x(n-L+1)]E[y_r(n)x(n-L+1)] \end{aligned} \quad (\text{C.7})$$

Sustituyendo (C.7) en (2.70), la función de coste normalizada puede ser expresada como

$$\begin{aligned} \frac{\tilde{J}}{m^2} &\triangleq \frac{J}{m^2} + \frac{1}{m^2}E[y_r^4(n)] + \frac{6}{m^2}E[y_s^2(n)]E[y_r^2(n)] - \frac{2k}{m}E[y_r^2(n)] + \\ &\quad + \frac{\alpha}{m^2}((E[y_r(n)x(n-L+1)])^2 + 2E[y_s(n)x(n-L+1)]E[y_r(n)x(n-L+1)]) \end{aligned} \quad (\text{C.8})$$

donde $\frac{J}{m^2}$ representa la función de coste normalizada en ausencia de ruido. Esta nueva expresión de la función de coste incluye nuevos términos que representan la distorsión sufrida debido al ruido aditivo. Las esperanzas estadísticas de (C.8) que contienen a la señal $y_r(n)$ se pueden expresar como

$$\begin{aligned} \frac{1}{m}E[y_r^2(n)] &= \frac{1}{m}\mathbf{w}^T\left(\frac{N_0}{2}\mathbf{I}\right)\mathbf{w} = \gamma\mathbf{w}^T\mathbf{w} \\ \frac{1}{m^2}E[y_r^4(n)] &= \frac{1}{m^2}\left(\sum_{i=0}^{L-1} w_i^4 E[r^4(n)] + \sum_{j=0}^{L-1} \sum_{l=0}^{L-1} w_j^2 w_l^2 E^2[r^2(n)]\right) = \\ &= 3\gamma^2\left(\sum_{i=0}^{L-1} w_i^4 + (\mathbf{w}^T\mathbf{w})^2\right) \end{aligned}$$

$$\frac{1}{m}(E[y_r(n)x(n-L+1)])^2 = \frac{1}{m}E\left[\left(\sum_{i=0}^{L-1} w_i r(n-i)\right)r(n-L+1)\right] = \gamma w_{L-1} \quad (\text{C.9})$$

donde $\gamma = \frac{N_0}{2m}$ representa la inversa de la relación señal a ruido (SNR^{-1}). Obsérvese que se ha usado la relación $E[r^4(n)] = 3E^2[r^2(n)] = 3\left(\frac{N_0}{2}\right)^2$ válida para todo proceso gaussiano de media cero. Considerando (2.28) y (C.2) las esperanzas conteniendo a la señal $y_s(n)$ son

$$\begin{aligned} \frac{1}{m}E[y_s^2(n)] &= \sum_{i=0}^{\infty} g_i^2 = \sum_{i=0}^{L-2} g_i^2 + g_{L-1}^2 \sum_{i=L-1}^{\infty} (a^{i-L+1})^2 = \sum_{i=0}^{L-2} g_i^2 + g_{L-1}^2 S_1 \\ \frac{1}{m}(E[y_s(n)x(n-L+1)])^2 &= \sum_{i=L-1}^{\infty} g_i a^{i-L+1} = g_{L-1} \sum_{i=L-1}^{\infty} (a^{i-L+1})^2 = g_{L-1} S_1 \end{aligned} \quad (\text{C.10})$$

donde S_1 está definida en (2.48). Sustituyendo (C.9) y (C.10) en (C.8)

$$\begin{aligned} \frac{\tilde{J}}{m^2} &= \frac{J}{m^2} + 2\gamma[3\mathbf{w}^T \mathbf{w} \left(\sum_{i=0}^{L-2} g_i^2 + g_{L-1}^2 S_1\right) - k\mathbf{w}^T \mathbf{w} + \alpha(g_{L-1} S_1)w_{L-1}] + \\ &+ \gamma^2[3\left(\sum_{i=0}^{L-1} w_i^4 + (\mathbf{w}^T \mathbf{w})^2\right) + \alpha w_{L-1}^2] \end{aligned} \quad (\text{C.11})$$

Ahora bien, dado que la relación SNR es lo suficientemente alta, podemos aproximar $\gamma^2 \simeq 0$ y, por lo tanto, reducir la expresión anterior obteniendo la expresión (2.71)

$$\frac{\tilde{J}}{m^2} \simeq \frac{J}{m^2} + 2\gamma[3\mathbf{w}^T \mathbf{w} \left(\sum_{i=0}^{L-2} g_i^2 + g_{L-1}^2 S_1\right) - k\mathbf{w}^T \mathbf{w} + \alpha(g_{L-1} S_1)w_{L-1}] \quad (\text{C.12})$$

C.3 Desarrollo del sistema de ecuaciones (2.81)

Si la SNR es alta, $\gamma^2 \simeq \gamma^3 \simeq 0$ y por lo tanto se cumplen las siguientes aproximaciones en las componentes de (2.80)

$$\begin{aligned} \tilde{g}_i^l &= g_i + \gamma \Delta g_i \\ (\tilde{g}_i^l)^2 &= g_i^2 + 2\gamma g_i \Delta g_i + \gamma^2 \Delta^2 g_i \simeq g_i^2 + 2\gamma g_i \Delta g_i \\ (\tilde{g}_i^l)^3 &= g_i^3 + 3\gamma g_i^2 \Delta g_i + 3\gamma^2 g_i \Delta^2 g_i + \gamma^3 \Delta^3 g_i \simeq g_i^3 + 3\gamma g_i^2 \Delta g_i \end{aligned} \quad (\text{C.13})$$

Teniendo en cuenta estas expresiones y (2.80) podemos expresar el gradiente (2.77) de la forma

$$\left. \frac{\partial \tilde{J}}{\partial \tilde{g}_i} \right|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} = \left. \frac{\partial J}{\partial g_i} \right|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} +$$

$$\begin{aligned}
& +\gamma[12\gamma\Delta g_i(\tilde{\mathbf{g}}^l)^T \mathbf{D}\tilde{\mathbf{g}}^l + 12(1 + 2\gamma\Delta g_l)\mathbf{D}_{(i+1,-)}\tilde{\mathbf{g}}^l - \\
& - 4k\mathbf{D}_{(i+1,-)}\tilde{\mathbf{g}}^l - 2a\alpha\gamma\Delta g_{L-1}S_1\delta(i - L + 2)] \simeq \\
& \simeq \frac{\partial J}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} + \gamma[12\mathbf{D}_{(i+1,l+1)} - 4k\mathbf{D}_{(i+1,l+1)}] \\
& \hspace{15em} i < L - 1 \text{ e } i \neq l \\
\frac{\partial \tilde{J}}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & = \frac{\partial J}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} + \\
& +\gamma[12\gamma(1 + \Delta g_l)(\tilde{\mathbf{g}}^l)^T \mathbf{D}\tilde{\mathbf{g}}^l + 12(1 + 2\gamma\Delta g_l)\mathbf{D}_{(l+1,-)}\tilde{\mathbf{g}}^l - \\
& - 4k\mathbf{D}_{(l+1,-)}\tilde{\mathbf{g}}^l - 2a\alpha\gamma\Delta g_{L-1}S_1\delta(i - L + 2)] \simeq \\
& \simeq \frac{\partial J}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} + \gamma[24\mathbf{D}_{(l+1,l+1)} - 4k\mathbf{D}_{(l+1,l+1)}] \\
\frac{\partial \tilde{J}}{\partial \tilde{g}_{L-1}} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & = \frac{\partial J}{\partial \tilde{g}_{L-1}} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} + \\
& +\gamma[12\gamma\Delta g_{L-1}S_1(\tilde{\mathbf{g}}^l)^T \mathbf{D}\tilde{\mathbf{g}}^l + 12(1 + 2\gamma\Delta g_l)\mathbf{D}_{(L,-)}\tilde{\mathbf{g}}^l - \\
& - 4k\mathbf{D}_{(L,-)}\tilde{\mathbf{g}}^l - 2\alpha S_1(2\gamma\Delta g_{L-1} - a(g_{(L-2)} + \gamma\Delta g_{L-2}))] \simeq \\
& \simeq \frac{\partial J}{\partial \tilde{g}_{L-1}} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} + \gamma[12\mathbf{D}_{(L,l+1)} - 4k\mathbf{D}_{(L,l+1)} - 2a\alpha S_1 g_{(L-2)}]
\end{aligned} \tag{C.14}$$

donde

$$\begin{aligned}
\frac{\partial J}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & = 12(1 + 2\gamma\Delta g_l)\gamma\Delta g_i - 4k\gamma\Delta g_i \simeq \gamma\Delta g_i[12 - 4k] \\
\frac{\partial J}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & = 4(k - 3)(1 - 3\gamma\Delta g_l) + 12(1 + 2\gamma\Delta g_l)(1 + \gamma\Delta g_l) - 4k(1 + \gamma\Delta g_l) \simeq \\
& \simeq \gamma\Delta g_l[12(k - 3) + 36 - 4k] \\
\frac{\partial J}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & = 12(1 + 2\gamma\Delta g_l)\gamma\Delta g_{(L-1)}S_1 - 4k\gamma\Delta g_{(L-1)}S_1 + 2\alpha\gamma\Delta g_{(L-1)}S_1^2 \simeq \\
& \simeq \gamma\Delta g_{(L-1)}[12 - 4kS_1 + 2\alpha S_1^2]
\end{aligned} \tag{C.15}$$

Sustituyendo (C.15) en (C.14) el gradiente se hace cero cuando

$$\begin{aligned}
\frac{\partial \tilde{J}}{\partial \tilde{g}_i} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & \simeq \gamma[\Delta g_i(12 - 4k) + (12 - 4k)\mathbf{D}_{(i+1,l+1)}] = 0 \quad i < L - 1 \text{ e } i \neq l \\
\frac{\partial \tilde{J}}{\partial \tilde{g}_l} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & \simeq \gamma[\Delta g_l 8k + (24 - 4k)\mathbf{D}_{(l+1,l+1)}] = 0 \\
\frac{\partial \tilde{J}}{\partial \tilde{g}_{L-1}} \Big|_{\mathbf{g}=\tilde{\mathbf{g}}^l=\mathbf{g}_{opt}^l+\gamma\Delta\mathbf{g}} & \simeq \gamma[\Delta g_{L-1}S_1(12 - 4k + 2\alpha S_1) + (12 - 4k)\mathbf{D}_{(L,l+1)} \\
& - 2a\alpha S_1 g_{L-2}] = 0
\end{aligned} \tag{C.16}$$

siendo $\mathbf{D}_{(i,j)}$ es el elemento (i, j) de la matriz \mathbf{D} .

Apéndice D

Desarrollos Correspondientes al Capítulo 3

D.1 Naturaleza de los Puntos Estacionarios de la Función de coste

En este apéndice se calculan los puntos estacionarios de la función (3.25) de la sección 3.7 y se analiza su naturaleza (es decir, se determina si son mínimos o no). Para simplificar la notación agruparemos los vectores de dimensión N , \mathbf{w}_1 y \mathbf{w}_2 , en uno de dimensión $2N$, $\mathbf{w}^H = [\mathbf{w}_1^H \ \mathbf{w}_2^H]$. Por claridad en la exposición presentamos de nuevo la función de coste

$$\begin{aligned} J(\mathbf{w}) &= E[||y_1(n)|^2 - 1|^2] + E[||y_2(n)|^2 - 1|^2] + \alpha |E[y_1^*(n)y_2(n)]|^2 = \\ &= E[|y_1(n)|^4] + E[|y_2(n)|^4] - 2E[|y_1(n)|^2] - 2E[|y_2(n)|^2] + \\ &\quad + \alpha |E[y_1(n)y_2^*(n)]|^2 + 2 \end{aligned} \quad (\text{D.1})$$

Evidentemente, (D.1) es una función de $2N$ variables complejas cuyos puntos estacionarios se corresponden con los puntos donde el vector gradiente

$$\begin{bmatrix} \nabla_{\mathbf{w}} J \\ \nabla_{\mathbf{w}}^* J \end{bmatrix} = \begin{bmatrix} \nabla_{\mathbf{w}_1} J \\ \nabla_{\mathbf{w}_1}^* J \\ \nabla_{\mathbf{w}_2} J \\ \nabla_{\mathbf{w}_2}^* J \end{bmatrix} \quad (\text{D.2})$$

se anula. En [8] se demuestra que si J es una función real de un vector complejo, como es nuestro caso, se cumple que

$$\nabla_{\mathbf{w}} J = \mathbf{0} \Leftrightarrow \nabla_{\mathbf{w}}^* J = \mathbf{0} \quad (\text{D.3})$$

por lo que los puntos estacionarios se corresponden con aquellos que cumplen

$$\nabla_{\mathbf{w}} J = \begin{bmatrix} \nabla_{\mathbf{w}_1} J \\ \nabla_{\mathbf{w}_2} J \end{bmatrix} = \mathbf{0} \quad (\text{D.4})$$

Asimismo, es bien conocido del análisis matemático que la naturaleza de los puntos estacionarios se determina a partir del carácter definido positivo de la matriz hessiana compleja de dimensión $4N \times 4N$

$$\mathbf{H}_{\mathbf{w}} J = \begin{bmatrix} \left. \frac{\partial^2 J}{\partial w_i w_j^*} \right|_{i,j=1}^N & \left. \frac{\partial^2 J}{\partial w_i w_j} \right|_{i,j=1}^N \\ \left. \frac{\partial^2 J}{\partial w_i^* w_j^*} \right|_{i,j=1}^N & \left. \frac{\partial^2 J}{\partial w_i^* w_j} \right|_{i,j=1}^N \end{bmatrix} \quad (\text{D.5})$$

Usando el operador de gradiente complejo, esta matriz hessiana compleja puede ser expresada de una forma más compacta como

$$\mathbf{H}_{\mathbf{w}} J = \begin{bmatrix} \nabla_{\mathbf{w}}^H (\nabla_{\mathbf{w}} J) & \nabla_{\mathbf{w}}^T (\nabla_{\mathbf{w}} J) \\ (\nabla_{\mathbf{w}}^T (\nabla_{\mathbf{w}} J))^* & (\nabla_{\mathbf{w}}^H (\nabla_{\mathbf{w}} J))^* \end{bmatrix} \quad (\text{D.6})$$

Utilizando el cambio de variables empleado en la ecuación (3.24), la función J puede ser expresada en términos de g_{ij} , $i, j = 1, 2$ en lugar de \mathbf{w} . Usando la regla de la cadena, el vector gradiente complejo $\nabla_{\mathbf{w}} J$ se puede, a su vez, escribir en términos de las derivadas respecto a g_{ij} decomponiéndolo como

$$\nabla_{\mathbf{w}} J = \frac{\partial J}{\partial g_{11}} \nabla_{\mathbf{w}} g_{11} + \frac{\partial J}{\partial g_{12}} \nabla_{\mathbf{w}} g_{12} + \frac{\partial J}{\partial g_{21}} \nabla_{\mathbf{w}} g_{21} + \frac{\partial J}{\partial g_{22}} \nabla_{\mathbf{w}} g_{22} \quad (\text{D.7})$$

Utilizando una notación matricial, $\nabla_{\mathbf{w}} J$ puede ser escrito de una forma más compacta de la forma

$$\nabla_{\mathbf{w}} J = \mathbf{C} \nabla_{\mathbf{g}} J \quad (\text{D.8})$$

donde \mathbf{C} es una matriz de dimensión $2N \times 4$

$$\mathbf{C} = \begin{bmatrix} \nabla_{\mathbf{w}} g_{11} & \nabla_{\mathbf{w}} g_{12} & \nabla_{\mathbf{w}} g_{21} & \nabla_{\mathbf{w}} g_{22} \end{bmatrix} \quad (\text{D.9})$$

y $\nabla_{\mathbf{g}} = \left[\frac{\partial}{\partial g_{11}}, \frac{\partial}{\partial g_{12}}, \frac{\partial}{\partial g_{21}}, \frac{\partial}{\partial g_{22}} \right]^T$ es el operador gradiente complejo de J con respecto a $\mathbf{g} = [g_{11}, g_{12}, g_{21}, g_{22}]^T$, cuya dimensión es 4×1 . Ya que $g_{ij} = \mathbf{w}_i^H \mathbf{h}_j$, la matriz \mathbf{C} tiene la siguiente forma

$$\mathbf{C} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{h}_1 & \mathbf{h}_2 \end{bmatrix} \quad (\text{D.10})$$

Como consecuencia, la matriz hessiana puede ser expresada de la siguiente forma

$$\begin{aligned} \mathbf{H}_{\mathbf{w}} J &= \begin{bmatrix} \mathbf{C} \nabla_{\mathbf{g}}^H (\nabla_{\mathbf{g}} J) \mathbf{C}^H & \mathbf{C} \nabla_{\mathbf{g}}^T (\nabla_{\mathbf{g}} J) \mathbf{C}^T \\ \mathbf{C}^* (\nabla_{\mathbf{g}}^T (\nabla_{\mathbf{g}} J))^* \mathbf{C}^H & \mathbf{C}^* (\nabla_{\mathbf{g}}^H (\nabla_{\mathbf{g}} J))^* \mathbf{C}^T \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^* \end{bmatrix} \mathbf{H}_{\mathbf{g}} J \begin{bmatrix} \mathbf{C}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^T \end{bmatrix} \end{aligned} \quad (\text{D.11})$$

donde

$$\mathbf{H}_{\mathbf{g}}J = \begin{bmatrix} \nabla_{\mathbf{g}}^H(\nabla_{\mathbf{g}}J) & \nabla_{\mathbf{g}}^T(\nabla_{\mathbf{g}}J) \\ (\nabla_{\mathbf{g}}^T(\nabla_{\mathbf{g}}J))^* & (\nabla_{\mathbf{g}}^H(\nabla_{\mathbf{g}}J))^* \end{bmatrix} \quad (\text{D.12})$$

es la matriz hessiana de J con respecto a \mathbf{g} de dimensión 8×8 . Obsérvese que $\mathbf{H}_{\mathbf{w}}J$ y $\mathbf{H}_{\mathbf{g}}J$ están relacionadas de forma congruente y, debido a la ley de inercia [63], los autovalores de ambas poseen el mismo signo. Por lo tanto, para comprobar si los puntos estacionarios de J son mínimos o no, es suficiente con analizar si $\mathbf{H}_{\mathbf{g}}J$ es definida positiva o no.

Varios son los tests que existen para determinar si una matriz es definida positiva o no. El que nos resultará más útil en este apéndice es el de examinar el signo de los determinantes de todas las submatrices superiores izquierda. Todos ellos serán positivos sí y solo sí la matriz es definida positiva [63].

D.1.1 Puntos Estacionarios

Las derivadas de J con respecto a los g_{ij} están representadas en (3.30). Calculando las derivadas segundas, la matriz $\mathbf{H}_{\mathbf{g}}J$ se puede descomponer en bloques como la submatriz hessiana hermítica

$$\nabla_{\mathbf{g}}^H(\nabla_{\mathbf{g}}J) = \begin{bmatrix} \nabla_1 & \vdots & \nabla_2 \end{bmatrix} \quad (\text{D.13})$$

donde

$$\nabla_1 = \begin{bmatrix} 4|g_{11}|^2q_1 + 4|g_{12}|^2m_1m_2 - & 4g_{11}^*g_{12}m_1m_2 + \\ -2m_1 + \alpha|g_{21}|^2m_1^2 & + \alpha g_{21}^*g_{22}m_1m_2 \\ \\ 4g_{11}g_{12}^*m_1m_2 + & 4|g_{12}|^2q_2 + 4|g_{11}|^2m_1m_2 - \\ + \alpha g_{21}g_{22}^*m_1m_2 & -2m_2 + \alpha|g_{22}|^2m_2^2 \\ \\ \alpha g_{11}g_{21}^*m_1^2 + & 0 \\ + \alpha g_{12}g_{22}^*m_1m_2 & \\ \\ 0 & \alpha g_{12}g_{22}^*m_2^2 + \\ & + \alpha g_{11}g_{21}^*m_1m_2 \end{bmatrix}$$

$$\nabla_2 = \begin{bmatrix} \alpha g_{11}^* g_{21} m_1^2 + & 0 \\ +\alpha g_{12}^* g_{22} m_1 m_2 & \\ \\ 0 & \alpha g_{12}^* g_{22} m_2^2 + \\ & +\alpha g_{11}^* g_{21} m_1 m_2 \\ 4|g_{21}|^2 q_1 + 4|g_{22}|^2 m_1 m_2 - & 4g_{21}^* g_{22} m_1 m_2 + \\ -2m_1 + \alpha |g_{11}|^2 m_1^2 & +\alpha g_{11}^* g_{12} m_1 m_2 \\ \\ 4g_{21} g_{22}^* m_1 m_2 + & 4|g_{22}|^2 q_2 + 4|g_{21}|^2 m_1 m_2 - \\ +\alpha g_{11} g_{12}^* m_1 m_2 & -2m_2 + \alpha |g_{12}|^2 m_2^2 \end{bmatrix}$$

y la submatriz hessiana simétrica

$$\nabla_{\mathbf{g}}^T(\nabla_{\mathbf{g}} J) = \begin{bmatrix} 2g_{11}^{*2} q_1 & 4g_{11}^* g_{12}^* m_1 m_2 & \alpha g_{11}^* g_{21}^* m_1^2 & \alpha g_{21}^* g_{12}^* m_1 m_2 \\ 4g_{11}^* g_{12}^* m_1 m_2 & 2g_{12}^{*2} q_2 & \alpha g_{11}^* g_{22}^* m_1 m_2 & \alpha g_{12}^* g_{22}^* m_2^2 \\ \alpha g_{11}^* g_{21}^* m_1^2 & \alpha g_{11}^* g_{22}^* m_1 m_2 & 2g_{21}^{*2} q_1 & 4g_{21}^* g_{22}^* m_1 m_2 \\ \alpha g_{21}^* g_{12}^* m_1 m_2 & \alpha g_{12}^* g_{22}^* m_2^2 & 4g_{21}^* g_{22}^* m_1 m_2 & 2g_{22}^{*2} q_2 \end{bmatrix} \quad (\text{D.14})$$

Examinaremos a continuación la matriz hessiana correspondiente a los puntos estacionarios de J descritos en la sección 3.7. Recuérdese que los habíamos clasificado en 6 grupos.

Grupo 1 : Ambas salidas cancelan las señales correspondientes a los dos usuarios, es decir

$$g_{11} = g_{12} = g_{21} = g_{22} = 0 \quad (\text{D.15})$$

Los dos tipos de submatriz hessiana en este punto son

$$\begin{aligned} \nabla_{\mathbf{g}}^H(\nabla_{\mathbf{g}} J) &= \begin{bmatrix} -2m_1 & 0 & 0 & 0 \\ 0 & -2m_2 & 0 & 0 \\ 0 & 0 & -2m_1 & 0 \\ 0 & 0 & 0 & -2m_2 \end{bmatrix} \\ \nabla_{\mathbf{g}}^T(\nabla_{\mathbf{g}} J) &= \mathbf{0} \end{aligned} \quad (\text{D.16})$$

Por lo tanto, $\mathbf{H}_{\mathbf{g}} J$ es una matriz diagonal con todos sus elementos negativos, lo cual indica que es definida negativa y que estos puntos son máximos de J .

Grupo 2 : Una salida extrae la señal correspondiente a uno de los usuarios mientras que la otra cancela ambas señales, es decir, uno de los g_{ij} es diferente de cero y los otros tres son iguales a cero. Hay por lo tanto cuatro soluciones que se

pueden incluir en este grupo, dependiendo de cuál de los cuatro g_{ij} sea diferente de cero. Una de las cuatro posibles soluciones es

$$|g_{11}|^2 \neq 0, \quad g_{12} = g_{21} = g_{22} = 0 \quad (\text{D.17})$$

Resolviendo el sistema de ecuaciones (3.30), el punto estacionario resulta ser

$$|g_{11}|^2 = \frac{m_1}{q_1}, \quad g_{12} = g_{21} = g_{22} = 0 \quad (\text{D.18})$$

Calculando el primer bloque de la matriz hessiana en este punto, se obtiene

$$\nabla_{\mathbf{g}}^H(\nabla_{\mathbf{g}}J) = \begin{bmatrix} 2m_1 & 0 & 0 & 0 \\ 0 & 2m_2(\frac{2}{k_1} - 1) & 0 & 0 \\ 0 & 0 & 2m_1(\frac{\alpha}{k_1} - 1) & 0 \\ 0 & 0 & 0 & -2m_2 \end{bmatrix} \quad (\text{D.19})$$

donde $k_i = \frac{q_i}{m_i^2}$ es el momento de cuarto orden normalizado de $s_i(t)$. Obsérvese que la matriz es definida negativa ya que el cuarto elemento es negativo y, por lo tanto, la matriz hessiana no puede ser definida positiva. Este punto no es un mínimo de J . Este resultado es idéntico para cualquiera de las soluciones de este grupo.

Grupo 3 : Una de las salidas extrae una combinación lineal de las señales correspondientes a los dos usuarios y la otro cancela ambas. En este grupo existen dos soluciones posibles dependiendo de cual de las salidas extrae la combinación lineal. Una posible solución sería

$$|g_{11}|^2 \neq 0, \quad |g_{12}|^2 \neq 0, \quad g_{21} = g_{22} = 0 \quad (\text{D.20})$$

Resolviendo el sistema de ecuaciones (3.30), el punto estacionario obtenido es

$$m_1|g_{11}|^2 = \frac{k_2 - 2}{k_1 k_2 - 4}, \quad m_2|g_{12}|^2 = \frac{k_1 - 2}{k_1 k_2 - 4}, \quad g_{21} = g_{22} = 0 \quad (\text{D.21})$$

donde $k_i = q_i/m_i^2$ es el momento de cuarto orden normalizado de $s_i(n)$.

El primer bloque de la matriz hessiana es

$$\nabla_{\mathbf{g}}^H(\nabla_{\mathbf{g}}J) = \begin{bmatrix} 2q_1|g_{11}|^2 & 4g_{11}^*g_{12}m_1m_2 & 0 & 0 \\ 4g_{11}g_{12}^*m_1m_2 & 2q_2|g_{12}|^2 & 0 & 0 \\ 0 & 0 & \alpha m_1^2|g_{11}|^2 - 2m_1 & \alpha g_{11}^*g_{12}m_1m_2 \\ 0 & 0 & \alpha g_{11}g_{12}^*m_1m_2 & \alpha m_2^2|g_{12}|^2 - 2m_2 \end{bmatrix} \quad (\text{D.22})$$

Para analizar si el hessiano es definido positivo, se puede estudiar el signo del determinante de la submatriz superior izquierda de $\mathbf{H}_{\mathbf{g}}J$ de dimensión 2×2

$$\Delta_2 = \text{Det} \begin{bmatrix} 2q_1|g_{11}|^2 & 4g_{11}^*g_{12}m_1m_2 \\ 4g_{11}g_{12}^*m_1m_2 & 2q_2|g_{12}|^2 \end{bmatrix} = \frac{2m_1m_2(k_1 - 2)(k_2 - 2)}{k_1k_2 - 4} \quad (\text{D.23})$$

Es evidente comprobar que si $k_1 < 2$ y $k_2 < 2$ (es decir, las kurtosis de las señales transmitidas son negativas) $\Delta_2 < 0$ y el hessiano no es definido positivo. De esta forma se demuestra que este punto no es un mínimo de J . Este análisis es idéntico para cualquiera de las soluciones de este grupo.

Grupo 4 : Ambas salidas extraen la misma señal y cancelan la otra. En este grupo podemos considerar dos posibles soluciones dependiendo de la señal que sea extraída por las dos salidas. Una de las posibles soluciones es de la forma

$$|g_{11}|^2 = |g_{21}|^2 \neq 0, \quad g_{12} = g_{22} = 0 \quad (\text{D.24})$$

Resolviendo el sistema de ecuaciones (3.30) el punto estacionario que se obtiene es

$$|g_{11}|^2 = |g_{21}|^2 = \frac{2m_1}{2q_1 + \alpha m_1^2}, \quad g_{12} = g_{22} = 0 \quad (\text{D.25})$$

Los dos tipos de bloques de la matriz hessiana son

$$\begin{aligned} & \nabla_{\mathbf{g}}^H(\nabla_{\mathbf{g}} J) = \\ = & \begin{bmatrix} (4k_1 + \alpha)m_1^2|g_{11}|^2 - 2m_1 & 0 & \alpha g_{11}^* g_{21} m_1^2 & 0 \\ 0 & 2m_2 \frac{4-\alpha-2k_1}{\alpha+2k_1} & 0 & \alpha g_{11}^* g_{21} m_1 m_2 \\ \alpha g_{11} g_{21}^* m_1^2 & 0 & (4k_1 + \alpha)m_1^2|g_{11}|^2 - 2m_1 & 0 \\ 0 & \alpha g_{11} g_{21}^* m_1 m_2 & 0 & 2m_2 \frac{4-\alpha-2k_1}{\alpha+2k_1} \end{bmatrix} \\ & \nabla_{\mathbf{g}}^T(\nabla_{\mathbf{g}} J) = \begin{bmatrix} 2g_{11}^{*2} q_1 & 0 & \alpha g_{11}^* g_{21}^* m_1^2 & 0 \\ 0 & 0 & 0 & 0 \\ \alpha g_{11}^* g_{21}^* m_1^2 & 0 & 2g_{21}^{*2} q_1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

La naturaleza de $\mathbf{H}_{\mathbf{g}} J$ está determinada por el signo de los determinantes de las primeras cinco submatrices superiores izquierda. Calculando sus determinantes se obtiene

$$\begin{aligned} \Delta_1 &= \frac{2\alpha(-2 + \alpha + 2k_1)m_1 m_2}{(\alpha + 2k_1)^2} \\ \Delta_2 &= \Delta_3 = 0 \\ \Delta_4 &= \frac{\alpha(-4 + \alpha + 4k_1)(-4 + 3\alpha + 4k_1)m_1^3 m_2^2}{(\alpha + 2k_1)^3} \\ \Delta_5 &= \frac{2\alpha(4 - 3\alpha - 4k_1)(-2 + \alpha + 2k_1)(-4 + \alpha + 4k_1)m_1^3 m_2^3}{(\alpha + 2k_1)^4} = \\ &= -\Delta_1 \Delta_4 \frac{\alpha + 2k_1}{2m_1} \quad (\text{D.26}) \end{aligned}$$

Es inmediato comprobar que si Δ_1, Δ_4 y α son positivos Δ_5 es negativo y $\mathbf{H}_g J$ no es definido positivo. Por lo tanto, la hipótesis $\alpha > 0$ es suficiente para asegurar que este punto no es un mínimo de J .

Grupo 5 : Las salidas obtienen cada una de las señales correspondientes a los usuarios. Dependiendo de qué señal se obtenga en cada salida, obtendremos dos puntos estacionarios en este grupo. Una de estas soluciones es de la forma

$$|g_{11}|^2 \neq 0, \quad |g_{22}|^2 \neq 0, \quad g_{12} = g_{21} = 0 \quad (\text{D.27})$$

Resolviendo el sistema de ecuaciones (3.30) el punto estacionario es que se obtiene es

$$|g_{11}|^2 = \frac{m_1}{q_1}, \quad |g_{22}|^2 = \frac{m_2}{q_2}, \quad g_{12} = g_{21} = 0 \quad (\text{D.28})$$

Desarrollando los dos tipos de bloques de la matriz hessiana para este punto, obtenemos

$$\begin{aligned} \nabla_g^H(\nabla_g J) &= \begin{bmatrix} 2m_1 & 0 & 0 & 0 \\ 0 & m_2\left(\frac{\alpha}{k_2} + \frac{4}{k_1} - 2\right) & 0 & 0 \\ 0 & 0 & m_1\left(\frac{\alpha}{k_1} + \frac{4}{k_2} - 2\right) & 0 \\ 0 & 0 & 0 & 2m_2 \end{bmatrix} \\ \nabla_g^T(\nabla_g J) &= \begin{bmatrix} 2g_{11}^{*2}q_1 & 0 & 0 & 0 \\ 0 & 0 & \alpha g_{11}^* g_{22}^* m_1 m_2 & 0 \\ 0 & \alpha g_{11}^* g_{22}^* m_1 m_2 & 0 & 0 \\ 0 & 0 & 0 & 2g_{22}^{*2}q_2 \end{bmatrix} \end{aligned} \quad (\text{D.29})$$

Para comprobar el carácter definido positivo de esta matriz, analizamos el signo del determinante de todas las matrices superiores izquierda. Calculando estos determinantes se obtiene

$$\begin{aligned} \Delta_1 &= 2m_1 \\ \Delta_2 &= m_2\left(\frac{\alpha}{k_2} + \frac{4}{k_1} - 2\right)\Delta_1 \\ \Delta_3 &= m_1\left(\frac{\alpha}{k_1} + \frac{4}{k_2} - 2\right)\Delta_2 \\ \Delta_4 &= 2m_2\Delta_3 \\ \Delta_5 &= \Delta_6 = \Delta_7 = \Delta_8 = 0 \end{aligned} \quad (\text{D.30})$$

Podemos ver que si $\alpha > 0$, $k_1 < 2$ y $k_2 < 2$, los determinantes Δ_2 y Δ_3 son siempre positivos. Por lo tanto, si $\alpha > 0$ y $k_i < 2$, $i = 1, 2$, es decir, las señales tienen kurtosis negativa, todas las submatrices superiores tienen un determinante no negativo y, por lo tanto, la matriz hessiana es definida positiva. Estos puntos son mínimos de J y corresponden con las soluciones deseadas donde se alcanza una separación perfecta de los usuarios.

Grupo 6 : Cada salida extrae una combinación lineal de ambas señales, es decir, las cuatro variables g_{ij} son diferentes de cero

$$g_{11} \neq 0, \quad g_{12} \neq 0, \quad g_{21} \neq 0, \quad g_{22} \neq 0 \quad (\text{D.31})$$

Para demostrar que este punto no es un mínimo, vamos a analizar la segunda submatriz superior de $\mathbf{H}_{\mathbf{g}}J$

$$\mathbf{U}_2 = \begin{bmatrix} \frac{\partial^2 J}{\partial g_{11} \partial g_{11}^*} & \frac{\partial^2 J}{\partial g_{11} \partial g_{12}^*} \\ \frac{\partial^2 J}{\partial g_{12} \partial g_{11}^*} & \frac{\partial^2 J}{\partial g_{12} \partial g_{12}^*} \end{bmatrix} \quad (\text{D.32})$$

Observando las segundas derivadas (D.13) de la diagonal de esta matriz, estas se pueden expresar en función de las derivadas primeras de la siguiente forma

$$\begin{aligned} g_{11}^* \frac{\partial^2 J}{\partial g_{11} \partial g_{11}^*} &= \frac{\partial J}{\partial g_{11}} + 2|g_{11}|^2 g_{11}^* q_1 - \alpha g_{21}^* g_{12}^* g_{22} m_1 m_2 \\ g_{12}^* \frac{\partial^2 J}{\partial g_{12} \partial g_{12}^*} &= \frac{\partial J}{\partial g_{12}} + 2|g_{12}|^2 g_{12}^* q_2 - \alpha g_{11}^* g_{21} g_{22}^* m_1 m_2 \end{aligned} \quad (\text{D.33})$$

Ahora bien, como en un punto estacionario las derivadas primeras se anulan $\frac{\partial J}{\partial g_{11}} = \frac{\partial J}{\partial g_{12}} = 0$ y las expresiones anteriores se pueden simplificar a

$$\begin{aligned} \frac{\partial^2 J}{\partial g_{11} \partial g_{11}^*} &= 2|g_{11}|^2 q_1 - \alpha \frac{g_{11} g_{12}^* g_{21}^* g_{22}}{|g_{11}|^2} m_1 m_2 \\ \frac{\partial^2 J}{\partial g_{12} \partial g_{12}^*} &= 2|g_{12}|^2 q_2 - \alpha \frac{g_{11}^* g_{12} g_{21} g_{22}^*}{|g_{12}|^2} m_1 m_2 \end{aligned} \quad (\text{D.34})$$

Además, ya que $\frac{\partial^2 J}{\partial g_{11} \partial g_{11}^*}$ y $\frac{\partial^2 J}{\partial g_{12} \partial g_{12}^*}$ deben ser reales, se puede concluir que

$$g_{11} g_{12}^* g_{21}^* g_{22} = g_{11}^* g_{12} g_{21} g_{22}^* = |g_{11}| |g_{12}| |g_{21}| |g_{22}| \quad (\text{D.35})$$

Por último, la submatriz superior izquierda de $\mathbf{H}_{\mathbf{g}}J$ de dimensión 2×2 es de la forma

$$\mathbf{U}_2 = \begin{bmatrix} 2|g_{11}|^2 q_1 - \alpha \frac{|g_{11}| |g_{12}| |g_{21}| |g_{22}|}{|g_{11}|^2} m_1 m_2 & 4g_{11}^* g_{12} m_1 m_2 + \alpha g_{21}^* g_{22} m_1 m_2 \\ 4g_{11} g_{12}^* m_1 m_2 + \alpha g_{21} g_{22}^* m_1 m_2 & 2|g_{12}|^2 q_2 - \alpha \frac{|g_{11}| |g_{12}| |g_{21}| |g_{22}|}{|g_{12}|^2} m_1 m_2 \end{bmatrix} \quad (\text{D.36})$$

su determinante es

$$\begin{aligned} \Delta_2 &= \text{Det}[\mathbf{U}_2] = 4m_1^2 |g_{11}|^2 m_2^2 |g_{12}|^2 (k_1 k_2 - 4) - \\ &\quad - 2\alpha |g_{11}| |g_{12}| |g_{21}| |g_{22}| m_1 m_2 \left(\frac{|g_{11}|^2}{|g_{12}|^2} q_1 + \frac{|g_{12}|^2}{|g_{11}|^2} q_2 + 4 \right) \end{aligned} \quad (\text{D.37})$$

Cuando consideramos $k_1 < 2$ y $k_2 < 2$ entonces $k_1 k_2 < 4$ y Δ_2 es siempre negativo. Por lo tanto, las condiciones $k_1 < 2$ y $k_2 < 2$ son suficientes para asegurar que la matriz hessiana en este punto no es definida positiva y que estos puntos no son mínimos de J .

D.2 Cálculo del MSE

En este apartado del apéndice se realizan todos los desarrollos correspondientes a la sección 3.7.2 donde se calcula el valor del MSE en los puntos que minimizan la función de coste (3.25). Para mayor claridad de los desarrollos de este apéndice consideraremos las siguientes definiciones

$$y_i(n) = \mathbf{g}_i \mathbf{s}(n) + \mathbf{w}_i \mathbf{r}(n) = \sum_{j=1}^N g_{ij} s_j(n) + \sum_{l=1}^N w_{il} r_l(n) \quad (\text{D.38})$$

$$\mathbf{g}_i = [g_{i1}, \dots, g_{iN}]^T \quad (\text{D.39})$$

$$\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_N] \quad (\text{D.40})$$

$$\frac{\partial}{\partial \mathbf{G}} = \left[\frac{\partial}{\partial \mathbf{g}_1}, \dots, \frac{\partial}{\partial \mathbf{g}_N} \right]^T \quad (\text{D.41})$$

$$\frac{\partial}{\partial \mathbf{g}_i} = \left[\frac{\partial}{\partial g_{i1}}, \dots, \frac{\partial}{\partial g_{iN}} \right]^T \quad (\text{D.42})$$

Además, se considerará que los símbolos transmitidos son i.i.d. con media cero y estadísticamente independientes del ruido aditivo.

D.2.1 Desarrollo de la Ecuación (3.53)

Este subapartado pretende expresar la ecuación (3.52) en forma matricial. Para ello, es necesario desarrollar las esperanzas considerando (D.38) y $r_l = 0 \quad \forall l$, es decir

$$\begin{aligned} \frac{1}{m_s} \sum_{i=1}^N E\{y_i^2(n)\} &= \frac{1}{m_s} \sum_{i=1}^N E\left[\sum_{m=1}^N g_{im} s_m(n) \sum_{l=1}^N g_{il} s_l(n) \right] = \sum_{i=1}^N \sum_{m=1}^N g_{im}^2 = \\ &= \sum_{i=1}^N \mathbf{g}_i^T \mathbf{g}_i = \text{Tr}\{\mathbf{G}^T \mathbf{G}\} \\ \frac{1}{m_s^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N (E\{y_i(n) y_j(n)\})^2 &= \frac{1}{m_s^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N (E\left[\sum_{m=1}^N g_{im} s_m(n) \sum_{l=1}^N g_{jl} s_l(n) \right])^2 \\ &= \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \left(\sum_{m=1}^N g_{im} g_{jm} \right)^2 = \sum_{i=1}^N \sum_{j=1}^N (\mathbf{g}_i^T \mathbf{g}_j)^2 - \sum_{l=1}^N (\mathbf{g}_l^T \mathbf{g}_l)^2 = \\ &= \text{Tr}\{(\mathbf{G}^T \mathbf{G})^2\} \\ \frac{1}{m_s^2} \sum_{i=1}^N E\{y_i^4(n)\} &= \frac{1}{m_s^2} \sum_{i=1}^N E \left[\sum_{m=1}^N \sum_{p=1}^N \sum_{l=1}^N \sum_{j=1}^N g_{im} g_{ip} g_{il} g_{ij} s_m(n) s_p(n) s_l(n) s_j(n) \right] \end{aligned}$$

$$\begin{aligned}
&= \begin{cases} m = p = l = j & \Rightarrow k \sum_{i=1}^N \sum_{m=1}^N g_{im}^4 \\ \left\{ \begin{array}{ll} m = p & l = j \\ m = l & p = j \\ m = j & p = l \end{array} \right\} & \Rightarrow 3 \sum_{i=0}^N \sum_{m=1}^N \sum_{\substack{l=1 \\ m \neq l}}^N g_{im}^2 g_{il}^2 \end{cases} \\
&= \sum_{i=1}^N \left((k-3) \sum_{m=1}^N g_{im}^4 + 3 \sum_{m=1}^N \sum_{l=1}^N g_{im}^2 g_{il}^2 \right) = \\
&= \sum_{i=1}^N \left((k-3) \sum_{m=1}^N g_{im}^4 + 3(\mathbf{g}_i^T \mathbf{g}_i)^2 \right) = \\
&= \sum_{i=1}^N \left((k-3) \sum_{m=1}^N g_{im}^4 \right) + 3 \text{Tr}\{(\text{diag}(\mathbf{G}^T \mathbf{G}))^2\} \quad (\text{D.43})
\end{aligned}$$

donde $\text{diag}(\mathbf{G}^T \mathbf{G})$ representa una matriz diagonal cuyos elementos corresponden con los de la diagonal de $\mathbf{G}^T \mathbf{G}$. Sustituyendo estas expresiones en (3.52) obtenemos la ecuación (3.53)

$$\begin{aligned}
J(\mathbf{G}) &= \frac{J}{m_s^2} = (k-3) \sum_{i=1}^N \sum_{j=1}^N g_{ij}^4 + k^2 N + \left(3 - \frac{\alpha}{2}\right) \text{Tr}\{(\text{diag}(\mathbf{G}^T \mathbf{G}))^2\} - \\
&\quad - 2k \text{Tr}\{\mathbf{G}^T \mathbf{G}\} + \frac{\alpha}{2} \text{Tr}\{(\mathbf{G}^T \mathbf{G})^2\} \quad (\text{D.44})
\end{aligned}$$

D.2.2 Cálculo del Gradiente de (3.53)

En la sección 3.7.2 se hace uso del gradiente de (3.53) con respecto a la matriz \mathbf{G} para el cálculo de los puntos estacionarios.

$$\frac{\partial J(\mathbf{G})}{\partial \mathbf{G}} = (k-3) \frac{\partial T_1}{\partial \mathbf{G}} + \left(3 - \frac{\alpha}{2}\right) \frac{\partial T_2}{\partial \mathbf{G}} - 2k \frac{\partial T_3}{\partial \mathbf{G}} + \frac{\alpha}{2} \frac{\partial T_4}{\partial \mathbf{G}} \quad (\text{D.45})$$

donde cada uno de los términos T_i y sus parciales son mostrados a continuación

$$\bullet T_1 = \sum_{i=1}^N \sum_{j=1}^N g_{ij}^4$$

$$\frac{\partial T_1}{\partial g_{ml}} = 4g_{ml}^3 \Rightarrow \frac{\partial T_1}{\partial \mathbf{G}} = 4\mathbf{G} \otimes \mathbf{G} \otimes \mathbf{G} \quad (\text{D.46})$$

donde $\mathbf{A} \otimes \mathbf{B}$ es una operación que obtiene una matriz cuyos elementos se corresponden con el producto individual de los elementos de la matriz \mathbf{A} con los de la matriz \mathbf{B} .

$$\begin{aligned}
\bullet T_2 &= \text{Tr}\{(\text{diag}(\mathbf{G}^T \mathbf{G}))^2\} = \sum_{i=1}^N (\mathbf{g}_i^T \mathbf{g}_i)^2 \\
\frac{\partial T_2}{\partial \mathbf{g}_l} &= 4(\mathbf{g}_l^T \mathbf{g}_l) \mathbf{g}_l \Rightarrow \frac{\partial T_2}{\partial \mathbf{G}} = 4 \text{diag}(\mathbf{G}^T \mathbf{G}) \mathbf{G}
\end{aligned} \tag{D.47}$$

$$\begin{aligned}
\bullet T_3 &= \text{Tr}\{\mathbf{G}^T \mathbf{G}\} = \sum_{i=1}^N \mathbf{g}_i^T \mathbf{g}_i \\
\frac{\partial T_3}{\partial \mathbf{g}_l} &= 2\mathbf{g}_l \Rightarrow \frac{\partial T_3}{\partial \mathbf{G}} = 2\mathbf{G}
\end{aligned} \tag{D.48}$$

$$\begin{aligned}
\bullet T_4 &= \text{Tr}\{(\mathbf{G}^T \mathbf{G})^2\} = \sum_{i=1}^N \sum_{j=1}^N (\mathbf{g}_i^T \mathbf{g}_j)^2 \\
\frac{\partial T_4}{\partial \mathbf{g}_l} &= 4(\mathbf{g}_l^T \mathbf{g}_l) \mathbf{g}_l + 4 \sum_{\substack{j=1 \\ j \neq l}}^N (\mathbf{g}_l^T \mathbf{g}_j) \mathbf{g}_j + 4 \sum_{\substack{i=1 \\ i \neq l}}^N (\mathbf{g}_i^T \mathbf{g}_l) \mathbf{g}_i \\
&= 4 \sum_{j=1}^N (\mathbf{g}_j^T \mathbf{g}_l) \mathbf{g}_j = 4 \sum_{j=1}^N \mathbf{g}_j \mathbf{g}_j^T \mathbf{g}_l \\
&\Downarrow \\
\frac{\partial T_4}{\partial \mathbf{G}} &= 4 \underbrace{\sum_{j=1}^N \mathbf{g}_j \mathbf{g}_j^T}_{\mathbf{G} \mathbf{G}^T} \underbrace{[\mathbf{g}_1, \dots, \mathbf{g}_N]}_{\mathbf{G}} = 4 \mathbf{G} \mathbf{G}^T \mathbf{G}
\end{aligned} \tag{D.49}$$

Sustituyendo todos estos términos en (D.45) se obtiene el gradiente de J

$$\frac{\partial J(\mathbf{G})}{\partial \mathbf{G}} = 4(k-3) \mathbf{G} \otimes \mathbf{G} \otimes \mathbf{G} + 4(3 - \frac{\alpha}{2}) \text{diag}(\mathbf{G}^T \mathbf{G}) \mathbf{G} - 4k \mathbf{G} + 4 \frac{\alpha}{2} (\mathbf{G} \mathbf{G}^T) \mathbf{G} \tag{D.50}$$

D.2.3 Desarrollo de la Ecuación (3.55)

Para obtener la expresión de la función de coste J normalizada para un canal con ruido es necesario sustituir $y_i = y_{is}(n) + y_{ir}(n)$ en (3.52) y descomponer las esperanzas estadísticas resultantes de la siguiente forma

$$\begin{aligned}
E[y_i^4(n)] &= E[y_{is}^4(n)] + E[y_{ir}^4(n)] + 6E[y_{is}^2(n)]E[y_{ir}^2(n)] \\
E[y_i^2(n)] &= E[y_{is}^2(n)] + E[y_{ir}^2(n)] \\
(E[y_i(n)y_j(n)])^2 &= (E[(y_{is}(n) + y_{ir}(n))(y_{js}(n) + y_{jr}(n))])^2 = \\
&= (E[y_{is}(n)y_{js}(n)])^2 + (E[y_{ir}(n)y_{jr}(n)])^2 + \\
&\quad 2E[y_{is}(n)y_{js}(n)]E[y_{ir}(n)y_{jr}(n)]
\end{aligned} \tag{D.51}$$

Sustituyendo estas expresiones en (3.52), la función de coste resultante es

$$\begin{aligned} \tilde{J}(\mathbf{G}) &= J(\mathbf{G}) + \frac{3}{m_s^2} \sum_{i=1}^N E^2[y_{ir}^2(n)] + \frac{6}{m_s^2} \sum_{i=1}^N E^2[y_{is}^2(n)] E^2[y_{ir}^2(n)] - \frac{2}{m_s^2} \sum_{i=1}^N E[y_{ir}^2(n)] + \\ &\quad \frac{\alpha}{2m_s^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N E^2[y_{ir}(n)y_{jr}(n)] + \frac{\alpha}{m_s^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N E[y_{is}(n)y_{js}(n)] E[y_{ir}(n)y_{jr}(n)] \end{aligned} \quad (\text{D.52})$$

donde $J(\mathbf{G})$ es la función de coste en una situación libre de ruido (3.53). Las esperanzas estadísticas de esta expresión se pueden desarrollar de la siguiente forma

$$\begin{aligned} \sum_{i=1}^N E[y_{ir}^2(n)] &= \sum_{i=1}^N \left(\sum_{m=1}^N w_{im} r_m(n) \sum_{l=1}^N w_{il} r_l(n) \right) = \sum_{i=1}^N \sum_{m=1}^N w_{im}^2 m_r = \\ &= m_r \sum_{i=1}^N \mathbf{w}_i^T \mathbf{w}_i \\ \sum_{i=1}^N \sum_{j=1}^N (E[y_{ir}(n)y_{jr}(n)]) &= \sum_{i=1}^N \sum_{j=1}^N \left(\sum_{m=1}^N w_{im} r_m(n) \sum_{l=1}^N w_{jl} r_l(n) \right) = \\ &= m_r \sum_{i=1}^N \sum_{j=1}^N \left(\sum_{m=1}^N w_{im} w_{jm} \right) = m_r \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N (\mathbf{w}_i^T \mathbf{w}_j) \\ \sum_{i=1}^N E[y_{ir}^4(n)] &= \sum_{i=1}^N 3E^2[y_{ir}^2(n)] = 3m_r^2 \sum_{i=1}^N ((\mathbf{w}_i^T \mathbf{w}_i)^2) \end{aligned} \quad (\text{D.53})$$

donde $m_r = E[r_i^2(n)]$ y hemos utilizado el hecho de que el ruido es un proceso gaussiano y por lo tanto $E[y_{ir}^4(n)] = 3E^2[y_{ir}^2(n)]$. Sustituyendo las expresiones (D.53) en (D.52), resulta la expresión

$$\begin{aligned} \tilde{J}(\mathbf{G}) &= J(\mathbf{G}) + 3\gamma^2 \sum_{i=1}^N ((\mathbf{w}_i^T \mathbf{w}_i)^2) + 6\gamma \sum_{i=1}^N \mathbf{g}_i^T \mathbf{g}_i \mathbf{w}_i^T \mathbf{w}_i - 2\gamma \sum_{i=1}^N \mathbf{w}_i^T \mathbf{w}_i + \\ &\quad + \frac{\alpha}{2} \gamma^2 \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N (\mathbf{w}_i^T \mathbf{w}_j)^2 + \alpha\gamma \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \mathbf{g}_i^T \mathbf{g}_j \mathbf{w}_i^T \mathbf{w}_j \end{aligned} \quad (\text{D.54})$$

donde considerando una SNR alta, podemos despreciar los términos multiplicados por γ^2 y además

$$\sum_{i=1}^N \mathbf{g}_i^T \mathbf{g}_i \mathbf{w}_i^T \mathbf{w}_i = \text{Tr}\{\text{diag}(\mathbf{G}^T \mathbf{G}) \text{diag}(\mathbf{W}^T \mathbf{W})\}$$

$$\begin{aligned}
\sum_{i=1}^N \mathbf{w}_i^T \mathbf{w}_i &= \text{Tr}\{\mathbf{W}^T \mathbf{W}\} \\
\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \mathbf{g}_i^T \mathbf{g}_j \mathbf{w}_i^T \mathbf{w}_j &= \text{Tr}\{\mathbf{G}^T \mathbf{G} \mathbf{W}^T \mathbf{W}\}
\end{aligned} \tag{D.55}$$

obteniendo la siguiente expresión para la función de coste

$$\begin{aligned}
\frac{\tilde{J}}{m_s^2} &\simeq \frac{J}{m_s^2} + \gamma(6 - \alpha) \text{Tr}\{\text{diag}(\mathbf{G}^T \mathbf{G}) \text{diag}(\mathbf{W}^T \mathbf{W})\} - \\
&\quad - 2\gamma k \text{Tr}\{\mathbf{W}^T \mathbf{W}\} + \gamma \alpha \text{Tr}\{\mathbf{G}^T \mathbf{G} \mathbf{W}^T \mathbf{W}\}
\end{aligned} \tag{D.56}$$

Considerando de nuevo que $\mathbf{W}^T = \mathbf{G}^T \mathbf{H}^{-1}$, obtenemos (3.55)

$$\begin{aligned}
\tilde{J}(\mathbf{G}) &= J(\mathbf{G}) + \gamma(6 - \alpha) \text{Tr}\{\text{diag}(\mathbf{G}^T \mathbf{G}) \text{diag}(\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G})\} - \\
&\quad - 2\gamma k \text{Tr}\{\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}\} + \gamma \alpha \text{Tr}\{\mathbf{G}^T \mathbf{G} \mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}\}
\end{aligned} \tag{D.57}$$

D.2.4 Cálculo del Gradiente de (3.55)

Para calcular el gradiente de $\tilde{J}(\mathbf{G})$ (3.55) respecto a \mathbf{G} consideraremos los gradientes correspondientes a los distintos sumandos de esta ecuación

$$\begin{aligned}
\bullet T_1 &= \text{Tr}\{\text{diag}(\mathbf{G}^T \mathbf{G}) \text{diag}(\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G})\} = \sum_{i=1}^N \mathbf{g}_i^T \mathbf{g}_i \mathbf{g}_i^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_i \\
\frac{\partial T_1}{\partial \mathbf{g}_l} &= 2\mathbf{g}_l \mathbf{g}_l^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l + 2(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l \mathbf{g}_l^T \mathbf{g}_l \\
&\quad \Downarrow \\
\frac{\partial T_1}{\partial \mathbf{G}} &= 2[\text{diag}(\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}) + \text{diag}(\mathbf{G}^T \mathbf{G}) (\mathbf{H}^T \mathbf{H})^{-1}] \mathbf{G}
\end{aligned} \tag{D.58}$$

$$\begin{aligned}
\bullet T_2 &= \text{Tr}\{\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}\} = \sum_{i=1}^N \mathbf{g}_i^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_i \\
\frac{\partial T_2}{\partial \mathbf{g}_l} &= 2(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l \Rightarrow \frac{\partial T_2}{\partial \mathbf{G}} = 2(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}
\end{aligned} \tag{D.59}$$

$$\bullet T_3 = \text{Tr}\{\mathbf{G}^T \mathbf{G} \mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}\} = \sum_{i=1}^N \sum_{j=1}^N \mathbf{g}_i^T \mathbf{g}_j \mathbf{g}_i^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_j$$

$$\frac{\partial T_3}{\partial \mathbf{g}_l} = 2\mathbf{g}_l \mathbf{g}_l^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l + 2\mathbf{g}_l^T \mathbf{g}_l (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l +$$

$$\begin{aligned}
& + \sum_{\substack{j=1 \\ j \neq l}}^N (\mathbf{g}_j \mathbf{g}_l^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_j + \mathbf{g}_l^T \mathbf{g}_j (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l) + \\
& + \sum_{\substack{j=1 \\ j \neq l}}^N \mathbf{g}_j \mathbf{g}_j^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l + \sum_{\substack{j=1 \\ j \neq l}}^N \mathbf{g}_j^T \mathbf{g}_l (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_j = \\
& = 2 \sum_{j=1}^N (\mathbf{g}_j \mathbf{g}_j^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_l + 2 \sum_{j=1}^N \underbrace{\mathbf{g}_j^T \mathbf{g}_l (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_j}_{(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{g}_j \mathbf{g}_j^T \mathbf{g}_l}) \\
& \Downarrow \\
\frac{\partial T_3}{\partial \mathbf{G}} & = 2 \left(\left(\sum_{j=1}^N (\mathbf{g}_j \mathbf{g}_j^T) (\mathbf{H}^T \mathbf{H})^{-1} + (\mathbf{H}^T \mathbf{H})^{-1} \left(\sum_{j=1}^N \mathbf{g}_j \mathbf{g}_j^T \right) \right) [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N] = \right. \\
& = 2(\mathbf{G} \mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} + (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G} \mathbf{G}^T) \mathbf{G} \quad (D.60)
\end{aligned}$$

Por lo tanto, el gradiente de $\tilde{J}(\mathbf{G})$ respecto a \mathbf{G} corresponde con la expresión (3.56)

$$\begin{aligned}
\frac{\partial \tilde{J}(\mathbf{G})}{\partial \mathbf{G}} & = \frac{\partial J(\mathbf{G})}{\partial \mathbf{G}} + 2\gamma \{ (6 - \alpha) [\text{diag}(\mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G}) + \text{diag}(\mathbf{G}^T \mathbf{G}) (\mathbf{H}^T \mathbf{H})^{-1}] \mathbf{G} - \\
& - 2k (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G} + \alpha [\mathbf{G} \mathbf{G}^T (\mathbf{H}^T \mathbf{H})^{-1} + (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{G} \mathbf{G}^T] \mathbf{G} \} \quad (D.61)
\end{aligned}$$

donde $\frac{\partial J(\mathbf{G})}{\partial \mathbf{G}}$ corresponde con la expresión del gradiente de la función de coste en una situación sin ruido (3.54).

D.2.5 Obtención de la Expresión (3.57)

Suponemos que la solución al sistema de ecuaciones (3.56) es de la forma $\tilde{\mathbf{G}}_0 = \mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM} = \mathbf{I} + \gamma \Delta \mathbf{G}_{CM}$ y que de nuevo, tenemos una alta SNR ($\gamma^2 \simeq 0$). Por lo tanto, sustituyendo $\tilde{\mathbf{G}}_0$ en los sumandos de (3.54), estos se pueden expresar como sigue

$$\begin{aligned}
\mathbf{G} \otimes \mathbf{G} \otimes \mathbf{G} \Big|_{\tilde{\mathbf{G}}_0} & = (\mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM}) \otimes (\mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM}) \otimes (\mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM}) \simeq \\
& \simeq \mathbf{I} + 3\gamma \Delta \mathbf{G}_{CM} \\
\text{diag}(\mathbf{G}^T \mathbf{G}) \mathbf{G} \Big|_{\tilde{\mathbf{G}}_0} & = \text{diag}((\mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM})^T (\mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM})) (\mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM}) = \\
& = [\text{diag}(\mathbf{G}_0) + \gamma (\text{diag}(\Delta \mathbf{G}_{CM}^T) + \text{diag}(\Delta \mathbf{G}_{CM})) + \\
& + \gamma^2 \text{diag}(\Delta \mathbf{G}_{CM}^T \Delta \mathbf{G}_{CM})] (\mathbf{G}_0 + \gamma \Delta \mathbf{G}_{CM}) \simeq \\
& \simeq \text{diag}(\mathbf{G}_0) + \gamma [\text{diag}(\Delta \mathbf{G}_{CM}) + \text{diag}(\Delta \mathbf{G}_{CM}^T) + \\
& + \text{diag}(\mathbf{G}_0) \Delta \mathbf{G}_{CM}] = \\
& = \mathbf{I} + \gamma [\Delta \mathbf{G}_{CM} + 2 \text{diag}(\Delta \mathbf{G}_{CM})]
\end{aligned}$$

$$\begin{aligned}
(\mathbf{G}\mathbf{G})\mathbf{G}|_{\tilde{\mathbf{G}}_0} &= (\mathbf{G}_0 + \gamma\Delta\mathbf{G}_{CM})(\mathbf{G}_0 + \gamma\Delta\mathbf{G}_{CM})^T(\mathbf{G}_0 + \gamma\Delta\mathbf{G}_{CM}) \simeq \\
&\simeq [\mathbf{G}_0\mathbf{G}_0^T + \gamma(\mathbf{G}_0\Delta\mathbf{G}_{CM}^T + \Delta\mathbf{G}_{CM}\mathbf{G}_0^T)](\mathbf{G}_0 + \gamma\Delta\mathbf{G}_{CM}) \simeq \\
&\simeq \mathbf{G}_0\mathbf{G}_0^T\mathbf{G}_0 + \gamma[\mathbf{G}_0\mathbf{G}_0^T\Delta\mathbf{G}_{CM} + \mathbf{G}_0\Delta\mathbf{G}_{CM}^T\mathbf{G}_0 + \\
&\quad + \Delta\mathbf{G}_{CM}\mathbf{G}_0^T\mathbf{G}_0] = \\
&= \mathbf{I} + \gamma[2\Delta\mathbf{G}_{CM} + \Delta\mathbf{G}_{CM}^T] \tag{D.62}
\end{aligned}$$

además el resto de los términos en (3.56) pueden ser simplificados de forma muy sencilla al considerar que $\gamma^2 \simeq \gamma^3 \simeq \gamma^4 \simeq 0$. De esta forma, el sistema de ecuaciones (3.56) se reduce a

$$\begin{aligned}
\frac{\partial \tilde{J}(\mathbf{G})}{\partial \mathbf{G}} &= \gamma \{12(k-3)\Delta\mathbf{G}_{CM} + 2(6-\alpha)(\Delta\mathbf{G}_{CM} + 2\text{diag}(\Delta\mathbf{G}_{CM})) - 4k\Delta\mathbf{G}_{CM} + \\
&\quad + 2\alpha(2\Delta\mathbf{G}_{CM} + \Delta\mathbf{G}_{CM}^T) + 2(6-\alpha)(\text{diag}((\mathbf{H}^T\mathbf{H})^{-1}) + (\mathbf{H}^T\mathbf{H})^{-1}) - \\
&\quad - 4k(\mathbf{H}^T\mathbf{H})^{-1} + 2\alpha(2(\mathbf{H}^T\mathbf{H})^{-1})\} = 0 \\
&\quad \Downarrow \\
&[6(k-3) + (6-\alpha) - 2k + 2\alpha]\Delta\mathbf{G}_{CM} + 2(6-\alpha)\text{diag}(\Delta\mathbf{G}_{CM}) + \alpha\Delta\mathbf{G}_{CM}^T + \\
&\quad + [(6-\alpha) - 2k + 2\alpha](\mathbf{H}^T\mathbf{H})^{-1} + (6-\alpha)\text{diag}((\mathbf{H}^T\mathbf{H})^{-1}) = 0 \\
&\quad \Downarrow \\
&\Delta\mathbf{G}_{CM}(4k - 12 + \alpha) + 2(6-\alpha)\text{diag}(\Delta\mathbf{G}_{CM}) + \alpha\Delta\mathbf{G}_{CM}^T + \\
&\quad + (6-2k+\alpha)(\mathbf{H}^T\mathbf{H})^{-1} + (6-\alpha)\text{diag}((\mathbf{H}^T\mathbf{H})^{-1}) = 0 \\
&\quad \Downarrow \\
&\Delta\mathbf{G}_{CM}(4k - 12 + 2\alpha) + 2(6-\alpha)\text{diag}(\Delta\mathbf{G}_{CM}) + \\
&\quad + (6-2k+\alpha)(\mathbf{H}^T\mathbf{H})^{-1} + (6-\alpha)\text{diag}((\mathbf{H}^T\mathbf{H})^{-1}) = 0 \tag{D.63}
\end{aligned}$$

donde se ha supuesto que la matriz $\Delta\mathbf{G}_{CM}$ es simétrica, es decir $\Delta\mathbf{G}_{CM} = \Delta\mathbf{G}_{CM}^T$.

Para despejar $\Delta\mathbf{G}_{CM}$ de este sistema de ecuaciones es necesario considerar independientemente los elementos de su diagonal de los demás. Denominando Δg_{ii} y b_{ii} los elementos de la diagonal de las matrices $\Delta\mathbf{G}_{CM}$ y $(\mathbf{H}^T\mathbf{H})^{-1}$, respectivamente, obtenemos

$$\begin{aligned}
(4(k-3) + 2\alpha)\Delta g_{ii} + 2(6-\alpha)\Delta g_{ii} + (2(3-k) + \alpha)b_{ii} &= 0 \\
&\quad \Downarrow \\
\Delta g_{ii} &= \frac{k-6}{2k}b_{ii} \tag{D.64}
\end{aligned}$$

Por otra parte, si Δg_{ij} y b_{ij} son los elementos de fuera de la diagonal de las matrices $\Delta\mathbf{G}_{CM}$ y $(\mathbf{H}^T\mathbf{H})^{-1}$, respectivamente, obtenemos

$$\begin{aligned}
(4(k-3) + 2\alpha)\Delta g_{ij} + (2(3-k) + \alpha)b_{ij} &= 0 \\
&\quad \Downarrow \\
\Delta g_{ij} &= \frac{2(k-3)-\alpha}{2(2(k-3)+\alpha)}b_{ij} = \frac{1}{2}\left(1 - \frac{2\alpha}{2(k-3)+\alpha}\right)b_{ij} \tag{D.65}
\end{aligned}$$

Considerando que los elementos fuera de la diagonal de $(\mathbf{H}^T \mathbf{H})^{-1}$ forman una matriz $(\mathbf{H}^T \mathbf{H})^{-1} - \text{diag}((\mathbf{H}^T \mathbf{H})^{-1})$, obtenemos la expresión de $\Delta \mathbf{G}_{CM}$ que corresponde con la de la ecuación (3.57)

$$\begin{aligned}
 \Delta \mathbf{G}_{CM} &= \frac{1}{2} \left(1 - \frac{2\alpha}{2(k-3) + \alpha} \right) ((\mathbf{H}^T \mathbf{H})^{-1} - \text{diag}((\mathbf{H}^T \mathbf{H})^{-1})) + \\
 &\quad + \frac{k-6}{2k} \text{diag}((\mathbf{H}^T \mathbf{H})^{-1}) = \tag{D.66} \\
 &= \frac{1}{2} \left(1 - \frac{2\alpha}{2(k-3) + \alpha} \right) (\mathbf{H}^T \mathbf{H})^{-1} + \frac{(\alpha-6)(k-3)}{k(2(k-3) + \alpha)} \text{diag}((\mathbf{H}^T \mathbf{H})^{-1})
 \end{aligned}$$

Apéndice E

Notación y Abreviaturas Empleadas

E.1 Notación

Representación	Significado	Ejemplos
Letra minúscula normal	Variable escalar o entera	n, t, i, j
Letra minúscula negrilla	Vector columna	\mathbf{x}, \mathbf{s}
Letra mayúscula normal	Constante entera	N, M, P
Letra mayúscula negrilla	Matriz	$\mathbf{H}, \mathbf{I}, \mathbf{G}$
Subíndices de minúsculas	Elementos de un vector	w_i
Superíndice $*$	Conjugado de un complejo	w_i^*, \mathbf{w}^*
Superíndice T	Transpuesto	$\mathbf{w}^T, \mathbf{G}^T$
Superíndice H	Transpuesto conjugado	$\mathbf{w}^H, \mathbf{G}^H$
Superíndice -1	Matriz inversa	\mathbf{H}^{-1}
Superíndice \dagger	Matriz pseudoinversa	\mathbf{H}^\dagger
$\text{Tr}\{\cdot\}$	Traza de una matriz	$\text{Tr}\{\mathbf{G}^T \mathbf{G}\}$
$\text{diag}(\cdot)$	Matriz diagonal	$\text{diag}(\mathbf{G}^T \mathbf{G})$
$\cdot_{(i,-)}$	Fila i -ésima de una matriz	$\mathbf{D}_{(l+1,-)}$
$\cdot_{(i,j)}$	Elemento (i, j) de una matriz	$\mathbf{D}_{(L,l+1)}$
$\cdot_{(n)}$	Dependencia temporal	$s(n), \mathbf{s}(n)$
$E[\cdot]$	Operador esperanza	$E[\mathbf{r}(n)\mathbf{r}^H(n)]$
$ \cdot $	Módulo de un complejo	$ y(n) , \mathbf{w} $
$\lfloor \cdot \rfloor$	Redondeo al entero por debajo	$\lfloor \frac{n}{M} \rfloor$
$*$	Convolución	$s_i(n) * c_i(n)$
\otimes	Producto entre los elementos de matrices	$\mathbf{G} \otimes \mathbf{G}$

E.2 Abreviaturas

AR	Auto-Regressive
ARMA	Auto-Regressive Moving Average
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
CM	Constant Modulus
CMA	Constant Modulus Algorithm
DD	Decision Directed
DS	Direct Sequence
FDMA	Frequency Division Multiple Access
FH	Frequency Hoping
FIR	Finite Impulse Response
FSE	Fractionally Spaced Equalizer
GSC	Generalized Sidelobe Canceller
HOS	Higher Order Statistics
ICA	Independent Component Analysis
iid	Independent, Identically Distributed
IIR	Infinite Impulse Response
ISI	Inter-Symbol Interference
LCMV	Linearly Constrained Minimum Variance
LDLM	Length Dependent Local Minima
LMS	Least Mean Squares
MA	Moving Average
MAI	Multiple Access Interference
MISO	Multiple Input Single Output
MIMO	Multiple Input Multiple Output
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MUSIC	MUltiple SIgnal Classification
PAM	Pulse Amplitude Modulation
QAM	Quadrature Amplitude Modulation
QPSK	Quaternary Phase Shift Keying
SDMA	Space Division Multiple Access
SINR	Signal to Interference plus Noise Ratio
SIR	Signal to Interference Ratio
SNR	Signal to Noise Ratio
TDMA	Time Division Multiple Access

Bibliografía

- [1] M. Aaron and D. Tufts. Intersymbol Interference and Error Probability. *IEEE Transactions on Information Theory*, IT-12, January 1966.
- [2] B. G. Agee. Convergent Behavior of Modulus-Restoring Adaptive Arrays in Gaussian Interference Environments. In *proceedings of 23rd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 1988.
- [3] R. Amari. Gradient Learning in Structured Parameter Spaces: Adaptive Blind Separation of Signal Sources. In *proceedings of WCNN96*, pages 951–956, San Diego, USA, 1996.
- [4] M. Austin. Decision-Feedback Equalization for Digital Communication Over Dispersive Channels. Technical Report 437, MIT Lincoln Laboratory, Lexington, Mass, 1967.
- [5] A. Bell and T. Sejnowski. An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, vol. 7(no. 6), November 1995.
- [6] P. A. Bello. Characterization of Randomly Time-Variant Linear Channels. *IEEE Transactions on Communications*, CS-11:pp. 360–393, December 1963.
- [7] A. Benveniste and G. Goursat. Blind Equalizers. *IEEE Transactions on Communications*, COM-32:pp. 871–883, 1984.
- [8] D. H. Brandwood. A Complex Gradient Operator and Its Applications in Adaptive Array Theory. In *proceedings Inst. Elect. Eng.*, volume 130, parts F & H, pages 11–16, February 1983.
- [9] X. Cao and R. Liu. General Approach to Blind Source Separation. *IEEE Transactions on Signal Processing*, vol. 44(no. 3):pp. 562–571, March 1996.
- [10] L. Castedo, C. J. Escudero, and A. Dapena. A Blind Signal Separation Method for Multiuser Communications. *IEEE Transactions on Signal Processing*, vol. 45(no. 5):pp. 1343–1348, May 1997.

- [11] L. Castedo and O. Macchi. Maximizing the Information Transfer for Adaptive Unsupervised Source Separation. In *proceedings of IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Paris, France, April 1997.
- [12] J. R. Cerquides. *Contribución a la Conformación Ciega de Haz en Comunicaciones*. PhD thesis, Universidad Politécnica de Catalunya, Spain, November 1996.
- [13] Y. Chen and C. Nikias. Adaptive Blind Equalization. Technical report, University of Southern California. Department of Electrical Engineering, Los Angeles, CA, USA, May 1991.
- [14] Y. Chen, C. Nikias, and J. Proakis. CRIMNO: CRIterion with Memory Non-linearity for Blind Equalization. In *proceedings of 26th Asilomar Conference on Signals, Systems and Computers*, pages 694–698, Pacific Grove, CA, USA, November 1991.
- [15] P. Comon. Independent Component Analysis, A New Concept? *IEEE Transactions on Signal Processing*, vol. 36:pp. 287–314, April 1994.
- [16] R. Jr. Compton. *Adaptive Antennas: Concepts and Performance*. Prentice-Hall, New Jersey, 1988.
- [17] Z. Ding, C. R. Johnson Jr., and R. A. Kennedy. On the (Non)Existence of Undesirable Equilibria of Godard Equalizers. *IEEE Transactions on Signal Processing*, vol. 40(no. 10):pp. 2425–2432, October 1992.
- [18] Z. Ding, R. A. Kennedy, Anderson B. D. O., and C. R. Johnson Jr. Ill-Convergence of Godard Blind Equalizers in Data Communication Systems. *IEEE Transactions on Communications*, vol. 39(no. 9):pp. 1313–1327, September 1991.
- [19] C. J. Escudero and L. Castedo. CMA-Based Joint Adaptation of Antenna Arrays for Multiuser Communications. In *proceedings of IEEE ATHOS Workshop on Higher Order Statistics*, Begur, Girona, Spain, June 1995.
- [20] C. J. Escudero, L. Castedo, and U. Mitra. A Modified CMA Equalizer for the Removal of Length Dependent Local Minima in AR Channels. *submitted to IEEE Transactions on Signal Processing*, July 1997.
- [21] C. J. Escudero, L. Castedo, and U. Mitra. Removal of Length Dependent Local Minima in CMA Equalizers. In *proceedings of European Workshop on Emerging Techniques for Communication Terminals*, Toulouse, France, July 1997.

- [22] C. J. Escudero, L. Castedo, and U. Mitra. Analysis of a Modified Constant Modulus Equalizer for Noisy AR Channels. In *proceedings of IEEE International Conference on Signal Processing and Communications*, Las Palmas de Gran Canaria, Spain, February 1998.
- [23] C. J. Escudero, C. Mejuto, and L. Castedo. Performance Analysis of a CMA-Based Adaptive Multiuser Detector. In *proceedings of IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Paris, France, April 1997.
- [24] I. Fijalkow, A. Touzni, and J. R. Treichler. Fractionally Spaced Equalization Using CMA: Robustness to Channel Noise and Lack of Disparity. *IEEE Transactions on Signal Processing*, vol. 45:pp. 56–66, January 1997.
- [25] G. J. Foschini. Equalization Without Altering or Detecting Data. *AT&T Tech. J.*, pages 1885–1911, October 1985.
- [26] W. Gardner. Exploitation of Spectral Redundancy in Cyclostationary Signals. *IEEE Signal Processing Magazine*, vol. 8(no.2):pp. 14–36, April 1991.
- [27] R. Gitlin, J. Hayes, and S. Weinstein. *Data Communications Principles*. Plenum Press, 1992.
- [28] R. Gitlin and S. Weinstein. Fractionally-Spaced Equalization: An Improved Digital Transversal Equalizer. *B.S.T.J.*, vol. 60:pp. 275–296, February 1981.
- [29] D. Godard. Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems. *IEEE Transactions on Communications*, COM-28:pp. 1867–1875, November 1980.
- [30] R. Gooch and J. Lundell. The CM Array: An Adaptive Beamformer for Constant Modulus Signals. In *proceedings of ICASSP86*, pages 2523–2526, Tokyo, Japan, April 1986.
- [31] L. Griffiths and C. Jim. An Alternative Approach to Linearly Constrained Adaptive Beamforming. *IEEE Transactions on Antennas and Propagation*, vol. AP-30:pp. 27–34, January 1982.
- [32] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Macmillan, New York, 1994.
- [33] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, third edition, 1996.
- [34] M. Honig, U. Madhow, and S. Verdú. Blind Adaptive Multiuser Detection. *IEEE Transactions on Information Theory*, vol. 41:pp. 944–960, July 1995.

- [35] N. Jablon. Joint Blind Equalization, Carrier Recovery, and Timing Recovery for High-Order QAM Constelations. *IEEE Transactions on Signal Processing*, vol. 40:pp. 1383–1398, 1992.
- [36] D. Johnson and D. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Prentice-Hall, New Jersey, 1993.
- [37] C. Jutten and J. Héroult. Blind Separation of Sources. Part I: An Adaptive Algorithm Based on Neuromimetic Architecture. *Signal Processing*, vol. 24:pp. 1–10, July 1991.
- [38] W. Lee. *Mobile Communications Design Fundamentals*. Wiley Series in Telecommunications, 1993.
- [39] W. Lee. *Mobile Cellular Telecommunications*. MacGraw Hill, 1995.
- [40] Y. Li and Z. Ding. Convergence Analysis of Finite Length Blind Adaptive Equalizers. *IEEE Transactions on Signal Processing*, vol. 43(no. 9):pp. 2120–2129, September 1995.
- [41] Y. Li, K. J. Liu, and Z. Ding. Length- and Cost-Dependent Local Minima of Unconstrained Blind Channel Equalizers. *IEEE Transactions on Signal Processing*, vol. 44(no. 11):pp. 2726–2735, November 1996.
- [42] R. Lucky, J. Salz, and E. Weldon. *Principles of Data Communication*. MacGraw Hill, 1968.
- [43] J. Lundell and B. Widrow. Application of the Constant Modulus Adaptive Beamformer to Constant and Non-Constant Modulus Algorithms. In *proceedings of 22nd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 1987.
- [44] R. Lupas and S. Verdú. Linear Multiuser Detectors for Synchronous Code Division Multiple Access Channels. *IEEE Transactions on Information Theory*, vol. IT-35(no. 1):pp. 123–136, January 1989.
- [45] R. Lupas and S. Verdú. Near-Far Resistance of Multiuser Detectors in Asynchronous Channels. *IEEE Transactions on Communications*, vol. COM-38(no. 4):pp. 496–508, April 1990.
- [46] A. Mathur, A. V. Keerthi, and J. Shink. Cochanel Signal Recovery Using the MUSIC Algorithm and the Constant Modulus Array. *IEEE Signal Processing Letters*, vol. 2, October 1995.
- [47] E. Moreau and O. Macchi. A One Stage Self-Adaptive Algorithm for Source Separation. In *proceedings of ICASSP94*, volume vol. III, pages 49–52, Adelaide, Australia, April 1994.

- [48] C. Nikias and J. Mendel. Signal Processing with Higher-Order Spectra. *IEEE Signal Processing Magazine*, pages pp. 10–37, July 1993.
- [49] C. Papadias. *Methods for Blind Equalization and Identification of Linear Channels*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, Paris, France, 1995.
- [50] C. Papadias and J. Arogyaswami. Blind Separation of Independent Co-Channel Signals. In *In proceedings of DSP*, Santorini, Greece, July 1997.
- [51] C. Papadias and D. Slock. New Adaptive Blind Equalization Algorithms for Constant Modulus Constellations. In *proceedings of ICASSP94*, volume vol. III, pages 321–324, Adelaide, Australia, April 1994.
- [52] A. Pérez Neira. *Técnicas de Arrays para Acceso Multiusuario*. PhD thesis, Universidad Politécnica de Catalunya, Spain, May 1995.
- [53] A. Pérez Neira and M. A. Lagunas. Multiuser Array Beamforming Based on a Neural Network Mapping. In *proceedings of ICASSP94*, volume vol. IV, pages 9–12, Adelaide, Australia, April 1994.
- [54] G. Picchi and G. Prati. Blind Equalization and Carrier Recovery Using a "Stop-and-Go" Decision Directed Algorithm. *IEEE Transactions on Communications*, COM-35:pp. 877–887, 1987.
- [55] S. Pillai. *Array Signal Processing*. Springer-Verlag, New York, 1989.
- [56] J. Proakis. *Digital Communications*. McGraw-Hill, third edition, 1995.
- [57] D. Sarwate and M. Pursley. Crosscorrelation Properties of Pseudorandom and Related Sequences. *Proceedings of the IEEE*, vol. 68(5):pp. 593–620, May 1980.
- [58] Y. Sato. A Method for Self-Recovering Equalization. *IEEE Transactions on Communications*, COM-23:pp. 679–682, June 1975.
- [59] R. Schmidt. Multiple Emiteer Location and Signal Parameter Estimation. In *proceedings of RADC Spectral Estimation Workshop*, pages 243–258, 1979.
- [60] J. Shynk, C. Chan, and R. Gooch. Comparative Performance Study of Several Blind Equalization Algorithms. In *SPIE'91*, San Diego, CA, 1991.
- [61] J. Shynk and R. Gooch. Convergence Properties of the Multistage CMA Adaptive Beamformer. In *proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 1993.
- [62] Signal Processing Society and the National Science Foundation. Signal processing information base. URL: <http://spib.rice.edu/>.

- [63] G. Strang. *Linear Algebra and its Applications*. Harcourt Brace Jovanovich, San Diego, CA, 1988.
- [64] L. Tong, F. Xu, and T. Kailath. Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach. *IEEE Transactions on Information Theory*, vol. 40(no. 2):pp. 340–349, March 1994.
- [65] J. R. Treichler and B. G. Agee. A New Approach to Multipath Correction of Constant Modulus Signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-31:pp. 459–472, April 1983.
- [66] J. R. Treichler, I. Fijalkow, and C. Johnson. Fractionally Spaced Equalizers. How Long Should they Really Be? *IEEE Signal Processing Magazine*, pages 65–83, May 1996.
- [67] B. Widrow and S. Sterans. *Adaptive Signal Processing*. Prentice Hall, Englewood Cliffs, New Jersey, USA, 1985.
- [68] Z. Xie, R. Short, and C. Rushforth. A Family of Suboptimum Detectors for Coherent Multiuser Communications. *IEEE Journal on Selected Areas in Communications*, pages 683–690, May 1990.
- [69] S. Zazo. *Estudio de Nuevas Técnicas Adaptativas para la Igualación Ciega de Canales Basadas en Algoritmos Tipo Bussgang*. PhD thesis, Universidad Politécnica de Madrid. ETSIT. Departamento de Señales, Sistemas y Radiocomunicaciones, Spain, April 1995.
- [70] S. Zeng, H. Zeng, and L. Tong. Blind Equalization using CMA: Performance Analysis and New Algorithm. In *proceedings of IEEE International Conference Communications*, pages 847–851, Dallas, Texas, USA, June 1996.