

5.

Simulaciones

5.1. Introducción

El objeto de este capítulo es mostrar de forma experimental las cualidades del algoritmo *Thin-ICA* en alguno de los diferentes campos de actuación en los que se puede aplicar [Hyvärinen] [Cichocki01]:

- ❖ *Procesado de Señal de Audio.* Uno de los motivos originales de *ICA* fue el deseo de dotar a un sistema la habilidad que poseen los humanos de concentrarse en una única voz que se encuentra inmersa en una conversación y ante la presencia de ruido, problema conocido como '*cocktail party*'. Así pues, la extracción de señales de audio es un campo de aplicación de especial interés.
- ❖ *Biomedicina.* La *BSS* se presenta como una herramienta de diagnóstico de enfermedades de forma no invasiva. Las señales biomédicas son captadas a través de múltiples electrodos situados sobre la piel del paciente. Dichas señales presentan poca intensidad y están distorsionadas por ruido e interferencia procedente de otros mecanismos fisiológicos. Algunos de las áreas de aplicación son los Electroencefalogramas (EEG), los Magnetoencefalogramas (MEG) y los Electrocardiogramas (ECG).
- ❖ *Procesado de Imágenes.* En el procesado de imágenes, la *BSS* permite la restauración de las mismas a través de la eliminación o minimización de artefactos que degradan su calidad. Un ejemplo de aplicación se encuentra en la necesidad de extraer una imagen objetivo que está alterada o sobreimpresa sobre otras imágenes.
- ❖ *Telecomunicaciones.* El modelo *BSS* puede aplicarse a las comunicaciones con el fin de mejorar el sistema. Un ejemplo de aplicación se encuentra en las comunicaciones inalámbricas con multitrayecto. La aplicación de algoritmos *BSS* permite la estimación del canal sin necesidad de las secuencias de entrenamiento, con la consecuente mejora de la capacidad del sistema.

Para comprobar las prestaciones del algoritmo desarrollado se utilizarán uno o varios de los índices que a continuación se presentan. Estos indicadores servirán para realizar una comparativa con respecto a alguna de las técnicas más empleadas en el panorama de la separación ciega de fuentes.

Este capítulo está organizado de la siguiente forma: en la sección 5.2 se definen los índices de prestaciones (*Cross*, P_i , *GAP*, *MLR* y *GlobInd*) que servirán para medir la capacidad del algoritmo implementado. La sección 5.3 se corresponde con un estudio sobre señales *experimentales*. Este es no es un campo real de aplicación, aunque emplearemos este tipo de señales para probar la eficacia del algoritmo implementado. En este apartado diferenciaremos entre la extracción simple y múltiple y estudiaremos el sistema ante la presencia y ausencia de ruido. Los resultados obtenidos vendrán dados a partir de la forma de onda de las señales a la salida, la función global de transferencia, el ascenso en la función contraste, varios de los índices de prestaciones anteriormente introducidos y una comparativa con el resto de algoritmos de referencia. En la sección 5.4 pasaremos a estudiar el procesado de la señal de audio. Se tratará de extraer la voz de un conferenciante que se encuentra en un entorno donde existen diferentes conversaciones de fondo, música ambiental y ruido. Tanto en este como en apartados sucesivos la extracción considerada será múltiple y los indicadores de calidad serán los mismos que los utilizados en la sección anterior. La sección 5.5 se corresponde con señales biomédicas y trataremos de extraer la actividad cardíaca fetal a través de señales captados por electrodos situados sobre la piel de la madre. En la sección 5.6 pasaremos al procesado de imágenes con la separación de cuatro imágenes que se encuentran superpuestas. El último ejemplo de aplicación lo encontramos en la sección 5.7 con la separación de varias señales para comunicaciones que presentan la peculiaridad de pertenecer al dominio complejo. El último apartado del capítulo se corresponde con la sección 5.8 y en ella se hace un repaso de los resultados alcanzados y se plantean algunas conclusiones.

5.2. Medida de Prestaciones

Como ya hemos comentado, el algoritmo *ThinICA* anteriormente desarrollado lo vamos a poner en práctica sobre señales procedentes de varios de los campos de aplicación. Para tener una idea sobre las prestaciones del mismo se utilizarán algunos de los índices de prestaciones que describiremos a continuación.

5.2.1. Índices de Prestaciones

Una vez se ha concluido el algoritmo y se ha realizado la separación de las fuentes, ¿cómo sabemos si el proceso se ha completado con éxito? Lo ideal sería poder comparar las componentes extraídas con las señales fuente, pero dado que no se tiene información sobre éstas ni del proceso de mezcla, este método no es factible. Así pues, hay que recurrir a unas medidas de prestaciones a partir de señales de prueba conocidas, las cuales se mezclan y que tras su paso por los algoritmos de separación, se extraen las señales estimadas. Con éstas últimas y las originales se realizan las medidas de prestaciones que indican la calidad de la separación. A continuación presentamos los índices de prestaciones más conocidos para corroborar la bondad de los procedimientos utilizados [Puntonet] [Mansour]:

- ❖ **Diafonía (*Crosstalk*).** La diafonía es un estimador que se define para dar información sobre el nivel de diferencia entre dos señales. Cuanto menor es la diafonía mayor será la similitud entre dichas señales. La expresión de la diafonía de la j -ésima fuente estimada viene dada como:



$$\text{Cross}_i = 10 \log \left(\frac{E \left((\hat{S}_j - S_j)^2 \right)}{E(S_j^2)} \right). \quad (5.1)$$

- ❖ **Performance Index o Crosstalk Error.** Este índice de prestaciones se define como:

$$\text{Cross}_e(\mathbf{G}) = P_i(\mathbf{G}) = \sum_{i=1}^N \left(\sum_{j=1}^N \frac{|G_{i,j}|}{\max_l (|G_{i,l}|)} - 1 \right) + \sum_{j=1}^N \left(\sum_{i=1}^N \frac{|G_{i,j}|}{\max_l (|G_{l,j}|)} - 1 \right) \quad (5.2)$$

donde \mathbf{G} es la matriz global de transferencia. Este estimador permanece invariante ante permutaciones aunque es muy sensible a errores numéricos. Los diferentes procedimientos buscan minimizar este índice para así presentar unas mejores prestaciones. La separación perfecta se consigue cuando:

$$P_{\text{index}}(\mathbf{G}) = 0.$$

Normalmente este índice se representa en dB.

- ❖ **GAP.** El *GAP* ofrece una medida de distancia de \mathbf{G} frente a una matriz diagonal. La expresión que define este índice es:

$$\text{GAP}(\mathbf{G}) = \sum_{i=1}^N \left(\sum_{j=1}^N |G_{i,j}| - 1 \right)^2 + \sum_{j=1}^N \left(\sum_{i=1}^N |G_{i,j}| - 1 \right)^2 + \sum_{i=1}^N \left| \sum_{j=1}^N |G_{i,j}|^2 - 1 \right| + \sum_{j=1}^N \left| \sum_{i=1}^N |G_{i,j}|^2 - 1 \right| \quad (5.3)$$

- ❖ **Rechazo.** El *Nivel de Rechazo Medio (Mean Rejection Level, MRL)* o interferencia de una fuente i sobre otra j se define como:

$$\text{MRL}_{i,j} = E(G_{i,j}^2). \quad (5.4)$$

Basándonos en el *MRL* podemos definir el *Nivel de Rechazo Global (GRL)* como:

$$\text{GRL} = \sum_{i \neq j} \text{MRL}_{i,j}, \quad (5.5)$$

aunque este indicador es muy sensible a los valores numéricos y puede dar lugar a equívoco.

- ❖ **Índice Global.** El Índice Global se define como:

$$\rho(\mathbf{G}) = 100 \sum_{j=1}^N \left(\max_i \left(\frac{|G_{i,j}|}{\sum_i |G_{i,j}|} \right) - \frac{1}{N} \right) \quad (5.6)$$

- ❖ **Otros Índices.** Existen otros muchos indicadores que miden las prestaciones de los algoritmos propuestos, entre ellos la *Norma del Error (ErrN)* y la *Tasa de Error de Símbolo (SER)*.

5.3. Señales Experimentales

En una primera instancia, antes de pasar a casos reales de aplicación vamos a comprobar el correcto funcionamiento con 'señales de laboratorio'. Las señales fuente \mathbf{S} que se emplean en los próximos ejemplos tendrán $T=10.000$ muestras y serán del tipo cuadrada, diente de sierra y senoidal. Cumpliendo uno de los requisitos necesarios para la extracción, estas señales son mutuamente independientes, de media cero y varianza unidad.

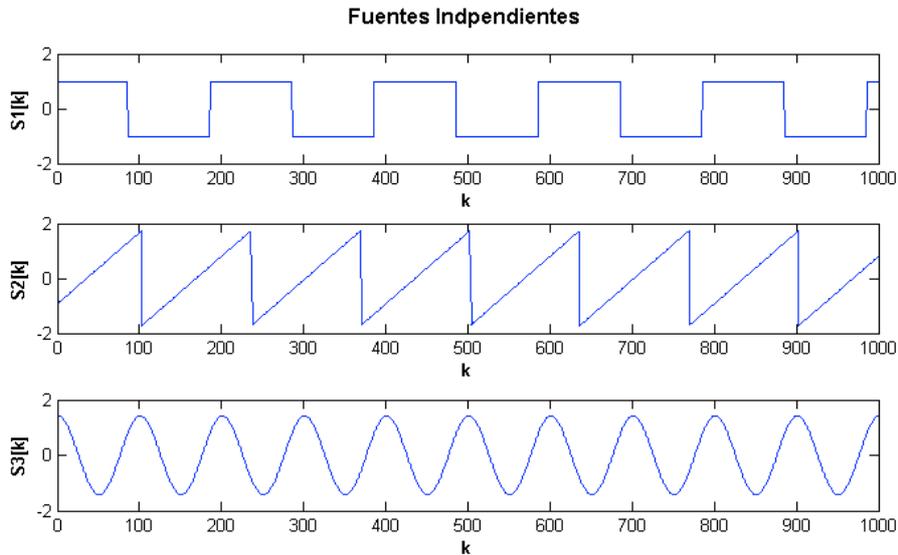


Figura 5.1 Fuentes Experimentales Independientes. En orden descendente: Cuadrada, Diente de sierra y Senoidal.

Para comprobar la capacidad del algoritmo vamos a realizar varios experimentos:

- ❖ *Extracción simple.* Este experimento consistirá en la extracción de una única fuente S_j inmersa en un entorno ruidoso. Se llevarán a cabo una serie de casos de prueba en los cuales el ruido introducido en el sistema irá incrementado su potencia.
- ❖ *Extracción múltiple.* Este experimento será similar al experimento de extracción simple, extrayéndose esta vez todas las fuentes \mathbf{S} de forma simultánea.

Las observaciones \mathbf{X} obtenidas tras el paso por el sistema de mezcla \mathbf{A} estarán contaminadas con un ruido del tipo aditivo blanco gaussiano. En los diferentes casos de prueba, la relación entre la potencia de las fuentes \mathbf{S} y la potencia de la señal interferente \mathbf{N} irá variando, desde una potencia de ruido nula hasta una potencia de ruido similar a la potencia de las fuentes. A esta relación de potencias se la denomina *Relación Señal a Ruido (SNR)*:

$$\text{SNR} = 10 \log \left(\frac{P_{\mathbf{S}}}{P_{\mathbf{N}}} \right) (\text{dB}). \quad (5.7)$$

A medida que la relación señal a ruido vaya disminuyendo, el sistema estará expuesto a un ruido más predominante, con la consecuente degradación en la calidad de la extracción.



El principal medidor de efectividad en la extracción que se empleará será el *Performance Index* P_i introducido en el capítulo anterior, aunque de manera puntual echaremos un vistazo a alguno de los otros indicadores de calidad introducidos.

Además de los índices previamente introducidos, nos apoyaremos para medir la calidad de las señales recuperadas en la relación entre la potencia de la señal extraída y la potencia del error, definida como:

$$\frac{P_Y}{P_{E_j}} = 10 \log \left(\frac{\sum_t (Y)^2}{\sum_t (S_j - Y)^2} \right) \text{ (dB)}, \quad j=1,2,\dots,p. \quad (5.8)$$

Al final de cada caso se incluirá una tabla de control donde se anotarán los resultados obtenidos en la extracción.

5.3.1. Extracción Simple de Señales Experimentales

Este es el experimento más sencillo a realizar y consistirá en la extracción de una única fuente de las anteriormente mencionadas.

Partiremos de un sistema en el que el número de observaciones ($M=8$) será mayor que el número de fuentes ($N=3$). La matriz \mathbf{A} tendrá por tanto dimensiones $[8,3]$ y, con el fin de poder seguir la evolución en la degradación en la extracción, permanecerá constante durante todos los casos que ocupan este experimento.

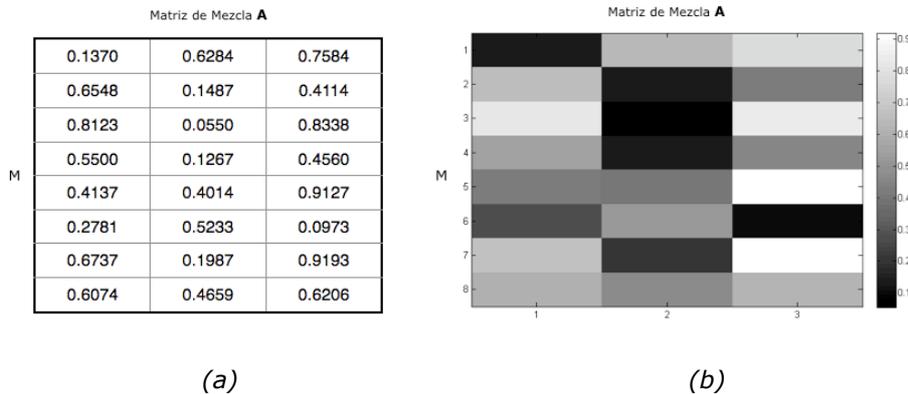


Figura 5.2 Matriz de mezcla \mathbf{A} generada de forma aleatoria y con dimensiones $[8, 3]$. (a) Valores numéricos. (b) Escala de grises.

5.3.1.1. Caso 1: Ausencia de Ruido ($P_N = 0$)

El primer caso de estudio es aquel en el que el sistema se encuentra libre de ruido, siendo por tanto la potencia de ruido nula.

Ante tal situación, la señal extraída se obtiene tras la sexta iteración:

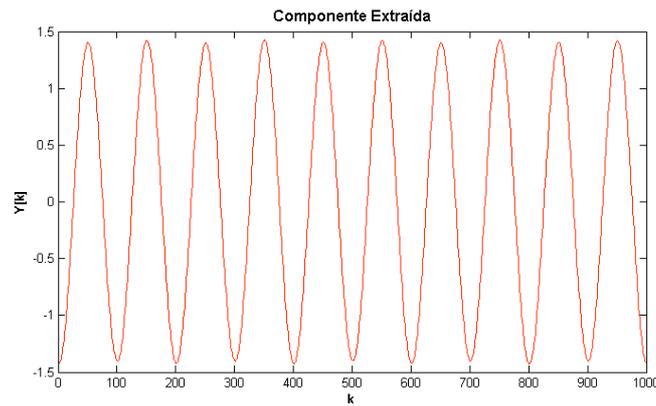


Figura 5.3 Señal extraída Y correspondiente con la tercera fuente original invertida.

En la *figura 5.4* se muestra la evolución en la función de transferencia global del sistema G así como de la relación entre la potencia de la señal extraída y la potencia del error P_Y/P_E a lo largo de las iteraciones. Como cabría esperar tras observar la señal extraída, la función de transferencia alcanzada se encuentra invertida.

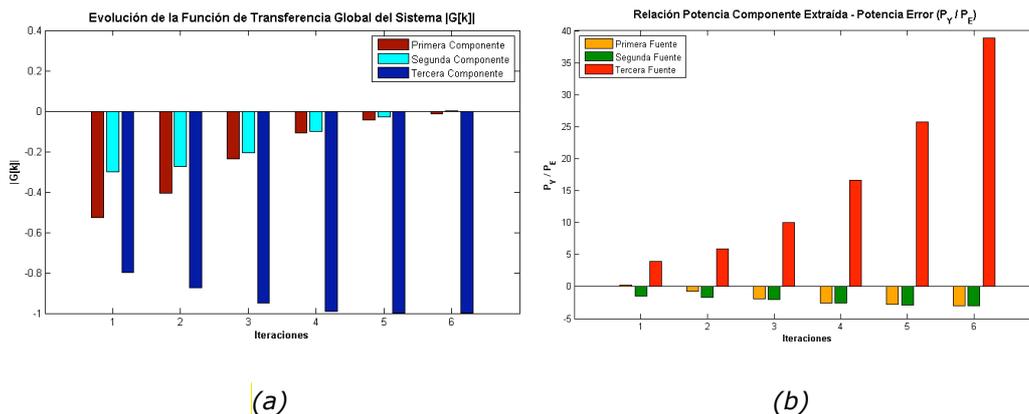


Figura 5.4 Evolución a lo largo de las iteraciones en la extracción de señales experimentales. (a) Función de transferencia global del sistema G . (b) Relación P_Y/P_E .

Por otra parte, en la siguiente figura podemos observar el ascenso de forma monótona hacia un máximo⁹ en la función contraste $\Phi_{\Theta}(U^{(k)})$, dando lugar esta convergencia a la extracción de una de las componentes.

⁹ El máximo hacia el que converge la función contraste es un valor dependiente de la fuente que se está extrayendo, siendo diferente para cada una de ellas.

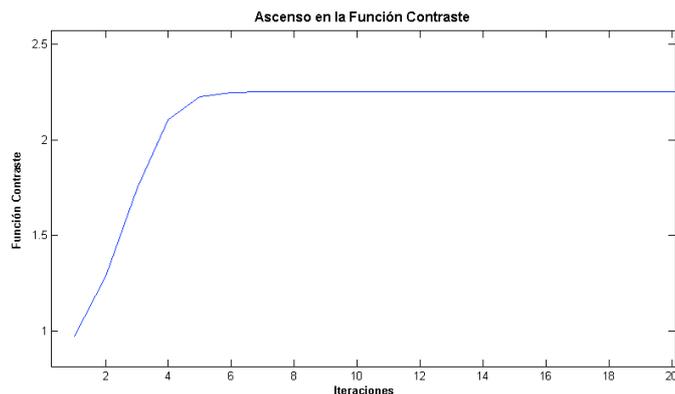


Figura 5.5 Ascenso en la función contraste $\Phi_{\Theta}(U^{(k)})$ en la extracción de señales experimentales. El máximo alcanzado se consigue tras la sexta iteración.

De acuerdo con la extracción, el máximo al que tiende la función contraste se alcanza en torno a la sexta iteración.

En la siguiente figura se muestra la evolución de algunos de los índices de prestaciones introducidos a lo largo de los capítulos.

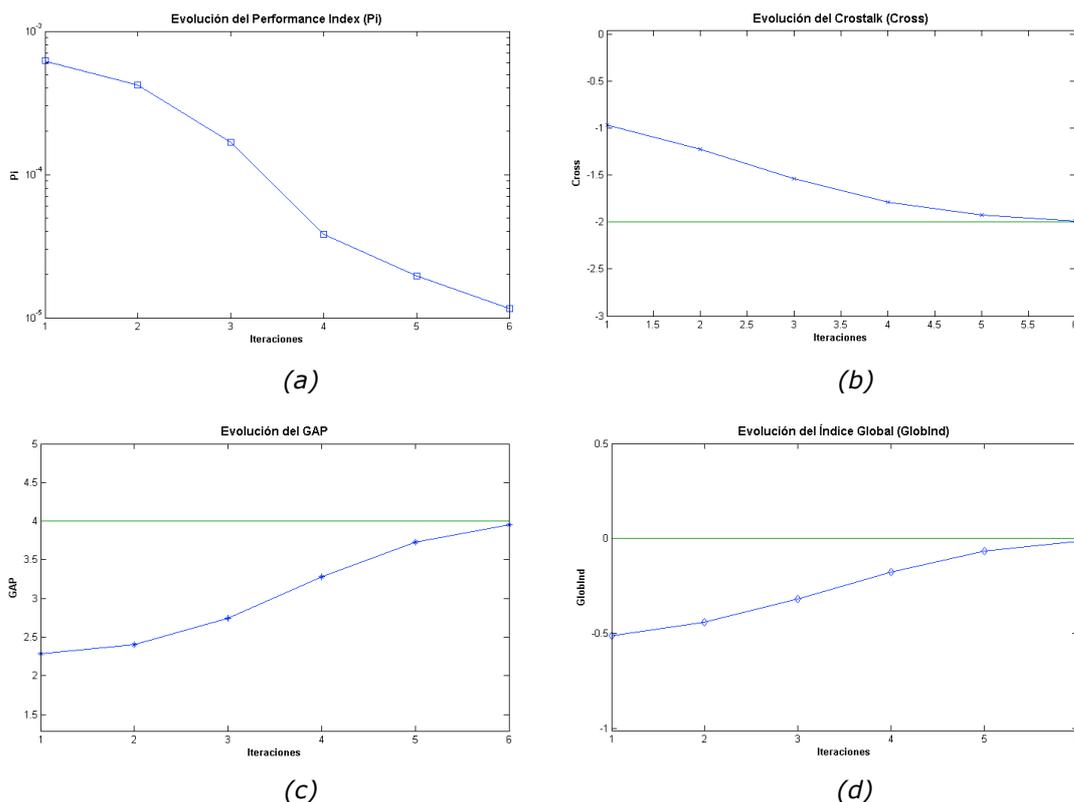


Figura 5.6 Evolución a lo largo de las iteraciones. (a) Performance Index P_i , (b) Crosstalk 'Cross', (c) GAP y (d) Índice Global (GlobInd).

Tras la observación de la forma de la señal extraída, la relación P_Y/P_E obtenida, la convergencia de la función contraste y de varios de los índices de prestaciones existentes podemos concluir que la extracción se ha realizado con éxito.

El ejemplo anterior, dado su carácter individual, no es representativo de la capacidad propia del algoritmo. Para mostrar el verdadero poder del algoritmo vamos a realizar una batería de 50 simulaciones donde la matriz de mezcla del sistema **A** ahora sí variará para poder obtener así respuesta ante escenarios diferentes. Los resultados obtenidos se muestran en las siguientes figuras.

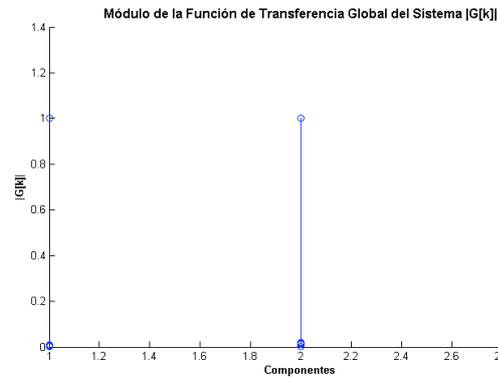
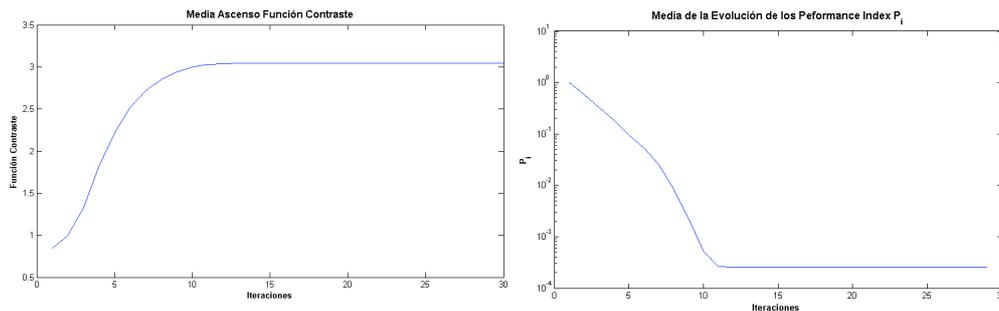


Figura 5.7 Batería de simulaciones. Módulo de la Función de transferencia global del sistema **G** obtenido en la extracción simple de señales de laboratorio.



(a)

(b)

Figura 5.8 Batería de simulaciones. (a) Media del ascenso en la función contraste $\Phi_{\Theta}(U^{(k)})$, (b) Media de la evolución de los Performance Index P_i .

Como puede apreciarse en la *figura 5.8*, el valor medio de las iteraciones en las que el algoritmo hace converger la función contraste a su máximo está en torno a 11. Para este mismo número de iteraciones el *Performance Index* se estabiliza.

La siguiente tabla muestra un resumen de los valores medios del número de iteraciones empleadas y del *Performance Index* del caso de estudio sin ruido:

| Iteraciones | P_i |
|-------------|-----------|
| 11 | 2.599 E-4 |

Tabla 5.1 Media de la extracción simple de señales experimentales.



5.3.1.2. Caso 2: Presencia de Ruido ($P_N \neq 0$)

En este segundo caso el sistema se verá contaminado por un ruido externo gaussiano que irá incrementando su potencia haciendo que la relación SNR disminuya su valor.

El algoritmo de extracción lo que busca es obtener una función de transferencia global del sistema \mathbf{G} lo mas parecida a la matriz unidad, por lo que este ruido externo se verá reflejado en las componentes extraídas:

$$\mathbf{Y} = \mathbf{U}^H \mathbf{Z} = \mathbf{U}^H \mathbf{W} \mathbf{X} = \mathbf{U}^H \mathbf{W} (\mathbf{A} \mathbf{S} + \mathbf{N}) = \mathbf{G} \mathbf{S} + \mathbf{U}^H \mathbf{W} \mathbf{N} \approx \mathbf{S} + \mathbf{N}'. \quad (5.9)$$

Para observar el comportamiento del algoritmo frente al ruido realizaremos una serie de baterías de simulaciones. En cada batería la relación señal a ruido permanecerá constante e iremos incrementado su valor desde 0 decibelios (potencia de ruido similar a la potencia de las fuentes) hasta un valor de 50 decibelios (Potencia de ruido despreciable frente a la potencia de señal). Cada batería estará compuesta por 50 simulaciones y, tras su conclusión, se realizará un promediado de los resultados. Como se indicó al principio del apartado, la matriz de mezclas \mathbf{A} permanecerá constante durante todas y cada una de las simulaciones. Los resultados obtenidos se muestran en las siguientes figuras:

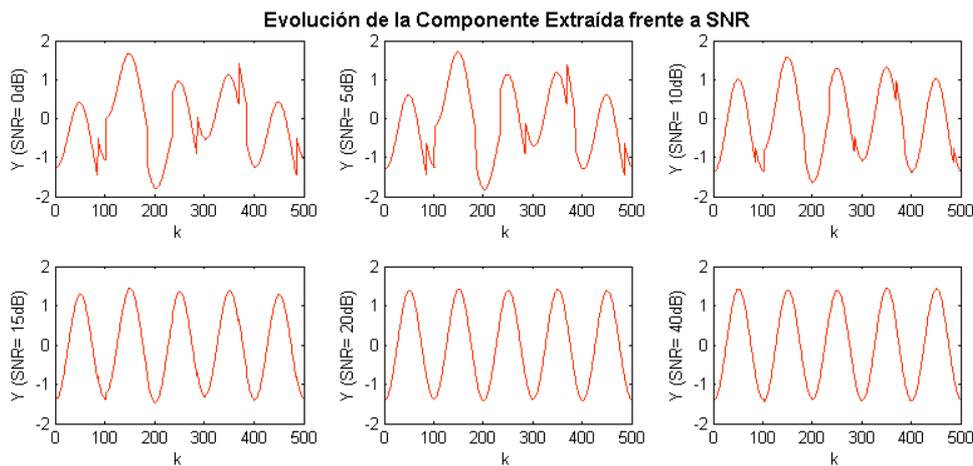


Figura 5.9 Evolución de la señal extraída frente al incremento de la relación señal a ruido. De izquierda a derecha y de arriba abajo: $SNR=0dB$, $SNR=5dB$, $SNR=10dB$, $SNR=15dB$, $SNR=20dB$ y $SNR=40dB$.

En la figura anterior se muestra la evolución de la forma de onda de la componente extraída con el incremento de la relación señal a ruido. En las imágenes superiores no se puede decir que la componente se haya extraído de forma correcta dado que la forma de la onda presenta irregularidades con respecto a la de la fuente original. A partir de una SNR de 15 decibelios el algoritmo tiene una extracción de calidad. A esta misma conclusión se puede llegar a través de las *figuras 5.10a* y *5.10b*. En ellas podemos apreciar que la función de transferencia global de sistema \mathbf{G} se aproxima a la matriz unitaria a partir de 15 dB, el mismo valor que presenta la relación P_Y/P_E ante esta situación. Aunque es un valor muy aceptable, la relación señal a ruido a partir de la cual el sistema no presenta mejora con respecto a un escenario libre de ruido está en torno a 20 decibelios, momento en el cual el P_i tiene un valor aproximado de $2.599 \text{ E-}4$.

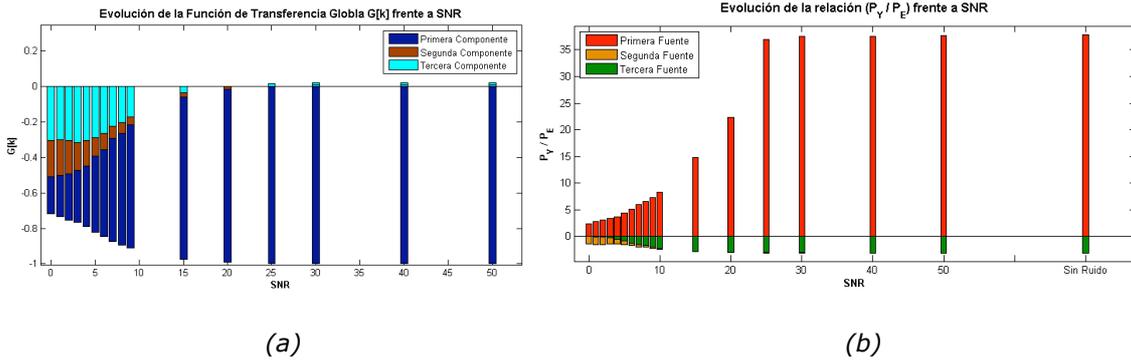


Figura 5.10 Evolución de la función de transferencia global del sistema y de la relación de potencias frente al incremento de la relación señal a ruido en la extracción múltiple de señales experimentales. (a) Evolución de \mathbf{G} . (b) Evolución de P_v/P_e .

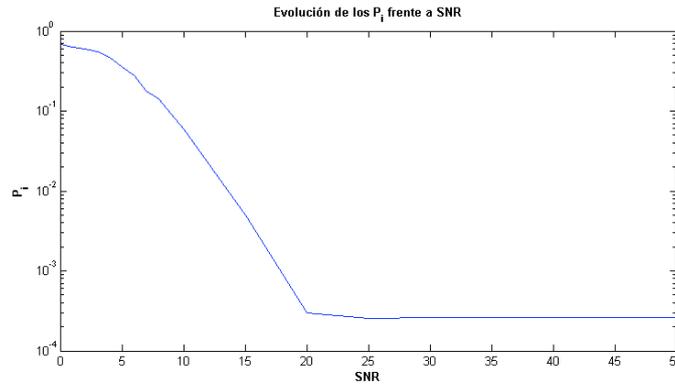


Figura 5.11 Evolución del Performance Index P_i frente al incremento de la relación señal a ruido en la extracción múltiple de señales experimentales.

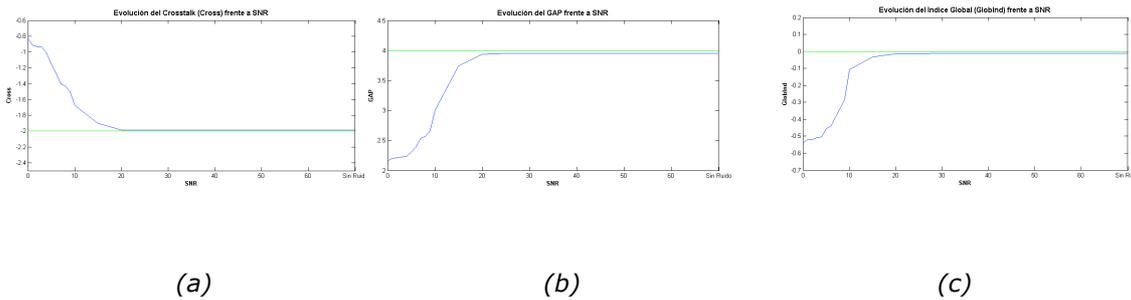


Figura 5.12 Evolución de los diferentes índices de prestaciones frente al incremento de la relación señal a ruido. (a) Crosstalk 'Cross', (b) GAP y (c) Índice Global (GlobInd).

5.3.2. Extracción Múltiple de Señales Experimentales

Este experimento es similar al anterior con la única diferencia de que se extraerán todas las fuentes \mathbf{S} . Para tal fin, la matriz de mezcla empleada será la misma que en el experimento anterior, pudiendo así comparar resultados.

De la misma forma que en la extracción simple de fuentes, dividiremos el apartado en una separación libre de ruido y una separación en un ambiente ruidoso, con potencia de ruido incremental.

5.3.2.1. Caso 1: Ausencia de Ruido ($P_N = 0$)

En las siguientes figuras se observan las componentes extraídas así como la función de transferencia global del sistema y la relación de potencias obtenidas tras la extracción de las tres componentes.

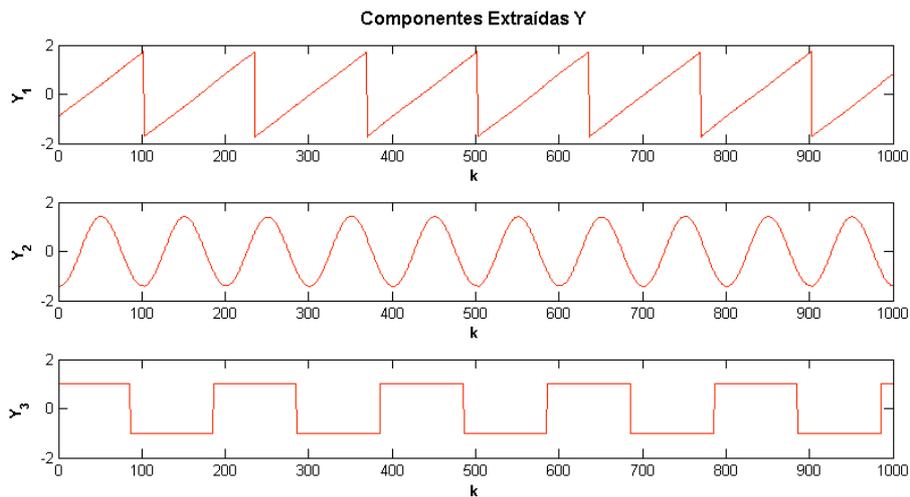
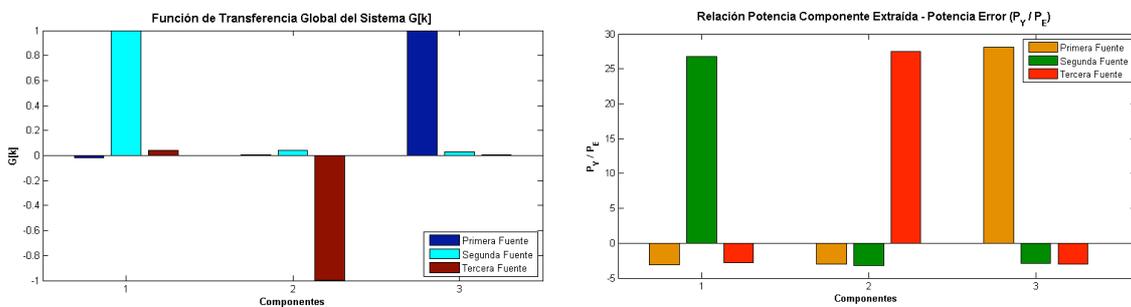


Figura 5.13 Componentes extraídas Y . En orden descendente: Diente de sierra, Senoidal y Cuadrada.



(a)

(b)

Figura 5.14 Extracción múltiple de señales de laboratorio. (a) Función de Transferencia global del sistema obtenida G . (b) Relación de potencias entre la componente extraída y el resto de fuentes p_Y / p_{E_j} .

Para tener un resultado más fiable sobre las prestaciones del algoritmo volveremos a realizar una batería de simulaciones, donde los resultados obtenidos se muestran en las figuras sucesivas.

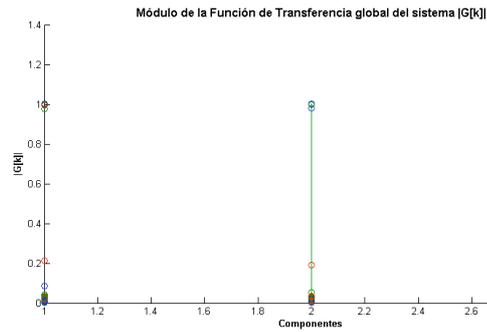


Figura 5.15 Batería de simulaciones. Módulo de la Función de transferencia global del sistema G obtenido en la extracción múltiple de señales de laboratorio.

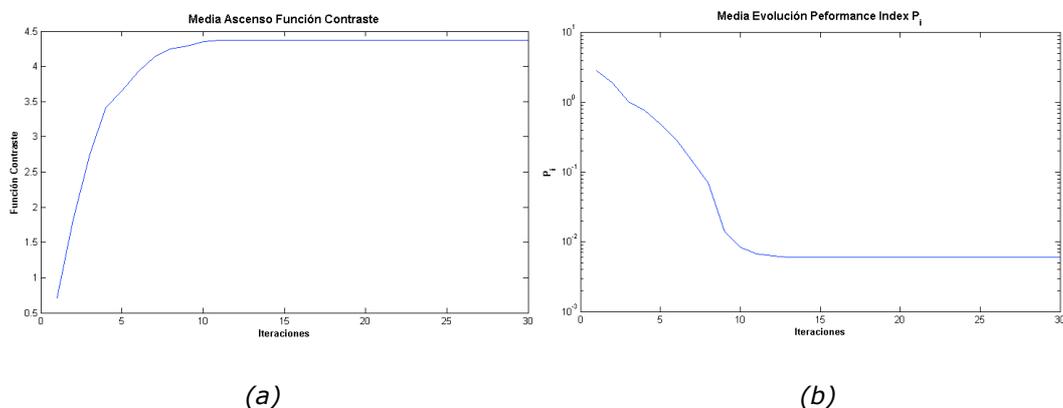


Figura 5.16 Batería de simulaciones. (a) Media del ascenso en la función contraste $\Phi_{\Theta}(U^{(k)})$, (b) Media de la evolución de los Performance Index P_i .

En la figura 5.16 puede observarse que el número medio de iteraciones empleadas para la extracción de todas las fuentes ($it=11$) es coincidente con el número de iteraciones para la extracción simple. En cambio, el valor medio en el que se estabiliza el índice de separación es menor que en el caso de extracción individual de fuentes.

A continuación se presenta la tabla de control con los valores medios del número de iteraciones empleadas y del Performance Index del caso de extracción múltiple y ausencia de ruido:

| Iteraciones | P_i |
|-------------|-----------|
| 11 | 5.737 E-3 |

Tabla 5.2 Media de la extracción múltiple de señales experimentales.



A simple vista, el nuevo algoritmo de extracción implementado tiene buena calidad, pero, ¿cómo de buena? Con el fin de poder situar su capacidad dentro del espectro de mecanismos de extracción de fuentes existentes vamos a realizar una batería de 100 simulaciones del mismo experimento con algunos de los métodos de referencia mencionados en capítulos anteriores. Los resultados se muestran en la siguiente figura:

| | JADE | AMUSE | SOBI | THIN-ICA |
|--------------------|-----------|-----------|-----------|-----------|
| Rapidez (segundos) | 8.29 | 6.28 | 7.72 | ~ 50 |
| P_i | 6.448 E-3 | 6.286 E-3 | 6.901 E-3 | 5.737 E-3 |

Tabla 5.3 Tabla Comparativa. Rapidez y calidad en la separación de componentes de los algoritmos de extracción JADE, AMUSE, SOBI y Thin-ICA.

Según la tabla anterior, el algoritmo implementado obtiene una mejor separación de las componentes pagando por ello un mayor tiempo de ejecución.

5.3.2.2. Caso 2: Presencia de Ruido ($P_N \neq 0$)

En este caso la extracción múltiple se realizará en un ambiente ruidoso y el procedimiento a seguir será el mismo que en la extracción simple: realizaremos baterías de simulaciones para diferentes valores de relación señal a ruido, promediaremos los resultados y los incluiremos en una gráfica para ver su evolución. Al igual que hicimos en el caso de extracción en ausencia de ruido, compararemos las prestaciones del algoritmo frente a los mecanismos de extracción de referencia y sacaremos conclusiones.

En las siguientes figuras se mostrará la evolución de la forma de onda de las componentes extraídas, su correspondiente función de transferencia \mathbf{G} y la relación de potencias P_Y/P_E conforme mejora la relación señal a ruido. En ellas podemos observar que aunque la calidad de la extracción es elevada en términos de la P_Y/P_E , la forma de onda mantiene el efecto del ruido.

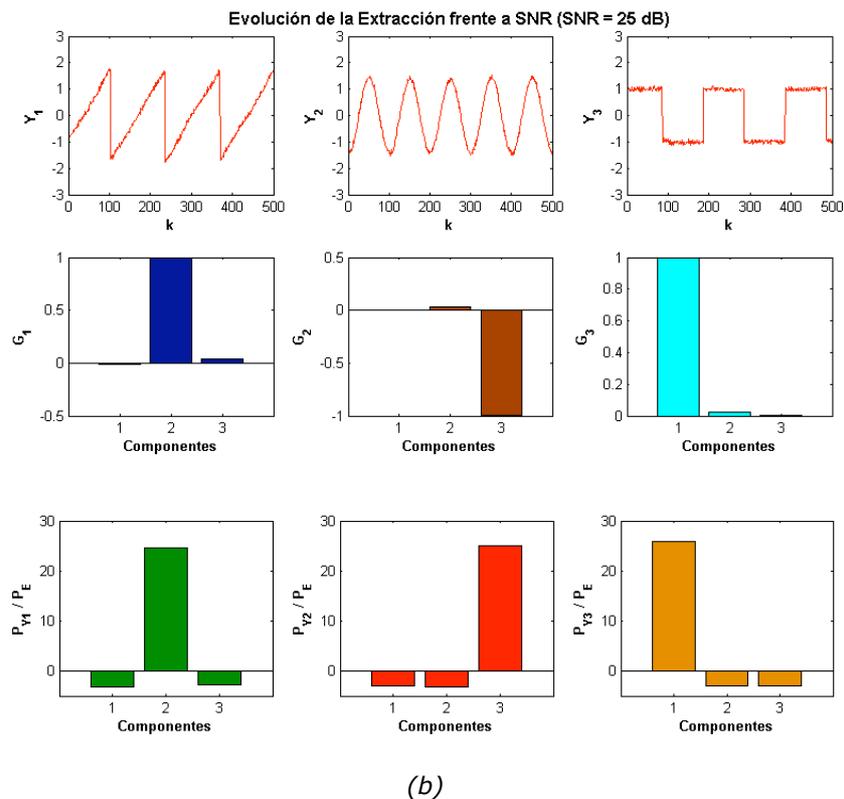
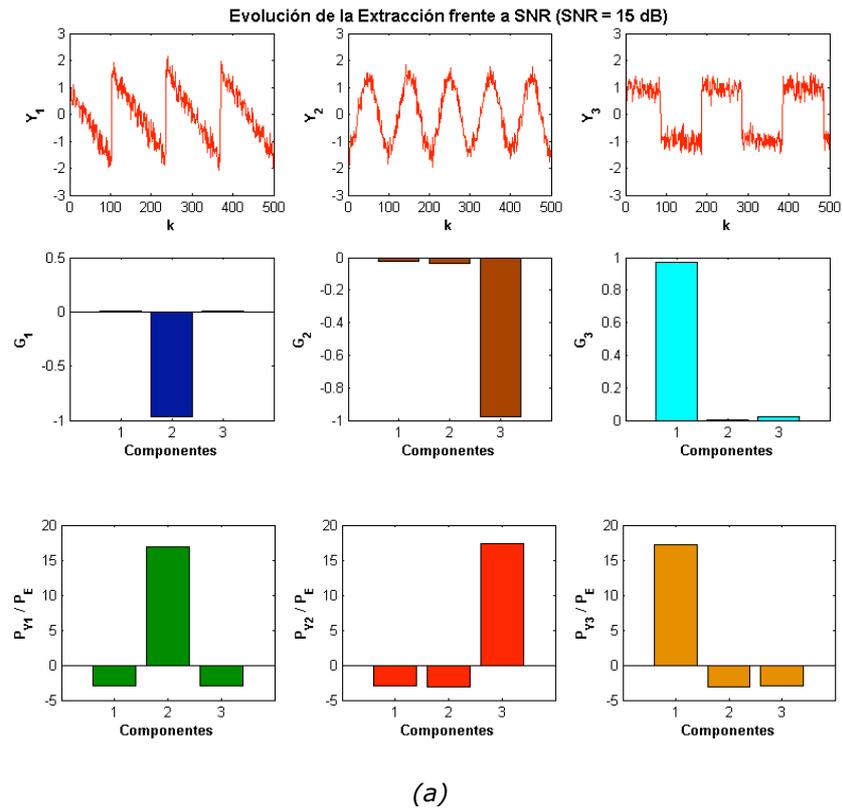


Figura 5.17 Evolución de la señal extraída, la función de transferencia global del sistema y la relación de potencia P_Y/P_N frente al incremento de la relación señal a ruido. (a) SNR=15dB y (b) SNR=25 dB.



En la *figura 5.18* se muestra la evolución del indicador de separación de componentes P_i con la mejora de la SNR . Esta imagen nos muestra que la extracción llega a su tope de calidad cuando la SNR alcanza aproximadamente los 25 decibelios, hecho que se ve corroborado en la imagen de la *figura 5.17b*, donde la relación de potencias P_Y/P_E tiene un valor aproximado al de la situación libre de ruido.

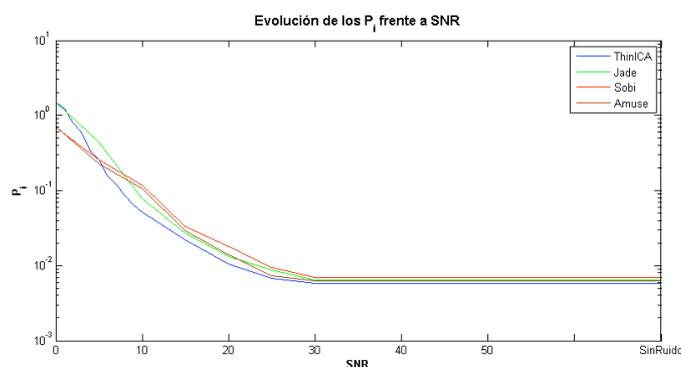


Figura 5.18 Evolución del Performance Index P_i frente al incremento de la relación señal a ruido para los métodos *ThinICA*, *Jade*, *Sobi* y *Amuse*.

En esta misma figura también se muestra la evolución del P_i frente a la relación señal a ruido de los algoritmos de extracción de referencia.

Como se puede apreciar, todos siguen un comportamiento similar: para valores muy bajos de SNR el índice de separación P_i es dependiente de la potencia de ruido, mientras que para valores altos de SNR es constante. Este comportamiento es debido a que todos estos algoritmos se basan en la diagonalización conjunta de matrices de cumulantes de cuarto orden o de matrices de correlación y , aunque el comportamiento del sistema no debería verse alterado por la presencia de un ruido aditivo blanco y gaussiano, a niveles de potencia de las fuentes del orden del nivel de potencia ruido su efecto sí es apreciable.

5.4. Procesado Señal Audio

Uno de los muchos campos de aplicación de la separación ciega de fuentes se encuentra en el procesado de señales de audio. Una de las principales motivaciones en el análisis de componentes independientes fue el problema del '*cocktail-party*'. La idea general de este problema consiste en separar una única fuente recopilada por un micrófono que recoge todos los sonidos ambiente. Así por ejemplo, se pretende rescatar la voz procedente de una persona considerando el resto de fuentes recogidas como ruido, encontrándonos entonces ante un problema de cancelación de ruido. Existen diversas técnicas para la cancelación de ruido pero únicamente son válidas cuando las fuentes ruidosas presentan características espectrales diferentes a las de la voz del conferenciante.

Dado que en la vida real la posición de los micrófonos respecto a las fuentes es un tanto arbitraria, el proceso de mezcla es desconocido, estando pues ante un problema de separación ciega de fuentes.

El experimento que realizaremos para comprobar el comportamiento del algoritmo *ThinICA* ante las señales de audio consistirá en la separación de la voz de una persona que mantiene una conversación telefónica en un entorno donde otros

individuos mantienen conversaciones ajenas a nuestro interés. Además, dicho entorno presenta música ambiental y ruido de fondo.

5.4.1. Extracción Múltiple de Señales de Audio

Las fuentes \mathbf{S} que formarán parte de este experimento se corresponden con señales de audio de tres segundos de duración, obtenidas con una tasa de muestreo de 8 KHz y 8 bits por muestra.

La primera de las fuentes se corresponde con la conversación que mantiene un individuo particular, la segunda de las fuentes se corresponde con el murmullo procedente del entorno en el que se encuentra inmerso el individuo, la tercera de las fuentes procede de la música ambiente presente en el entorno mientras que la cuarta fuente se corresponde con ruido. En la *figura 5.19* se muestran la forma de onda de cada una de las fuentes originales.

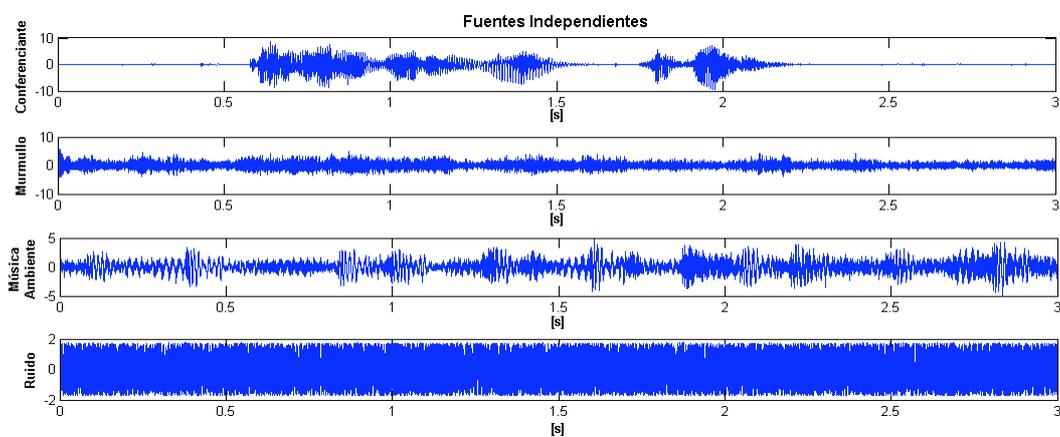


Figura 5.19 Fuentes de audio originales \mathbf{S} . En orden descendente: conferenciante, murmullo, música ambiente y ruido.

Las fuentes originales son procesos independientes como se puede observar en la siguiente figura. En ella se representa la dispersión de las muestras de cada una de las fuentes S_i con respecto a cualquier otra fuente S_j ($j \neq i$).

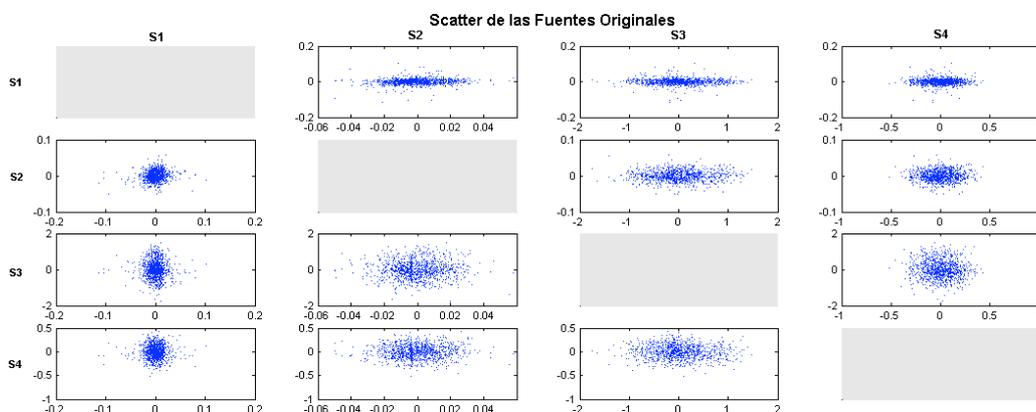


Figura 5.20 Dispersión de las muestras de las fuentes originales \mathbf{S} incorreladas.



Estas fuentes se mezclan de forma aleatoria mediante una matriz **A** de dimensiones [4,4] simulando la arbitrariedad de la situación de los micrófonos respecto las fuentes. En la siguiente figura se muestran las observaciones a las que se le aplicará el algoritmo de extracción:

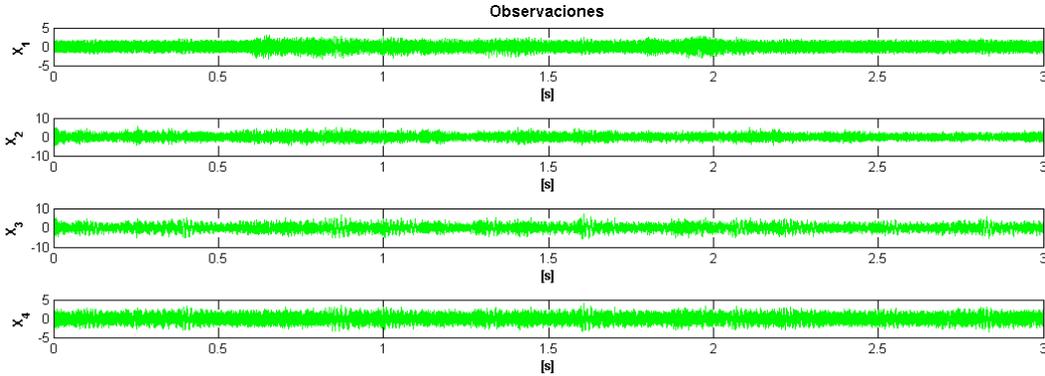


Figura 5.21 Observaciones **X**. Fuentes originales mezcladas a partir de la matriz **A**.

El proceso de mezcla provoca que las fuentes pasen a estar correladas. Así pues, pasamos de una matriz de correlación de **S** cuasi-identidad a una matriz donde los términos no pertenecientes a la diagonal principal tienen un peso significativo:

$$\text{Corr}(\mathbf{S}) = \begin{bmatrix} 1.0000 & -0.0032 & 0.0003 & 0.0100 \\ -0.0032 & 1.0000 & -0.0101 & -0.0022 \\ 0.0003 & -0.0101 & 1.0000 & 0.0026 \\ 0.0100 & -0.0022 & 0.0026 & 1.0000 \end{bmatrix} \rightarrow \text{Corr}(\mathbf{X}) = \begin{bmatrix} 1.0000 & 0.4102 & 0.4153 & 0.5221 \\ 0.4102 & 1.0000 & 0.5530 & 0.5114 \\ 0.4153 & 0.5530 & 1.0000 & 0.5646 \\ 0.5221 & 0.5114 & 0.5646 & 1.0000 \end{bmatrix}$$

(a) (b)

Figura 5.22 Matrices Correlación. (a) Correlación de las fuentes, $\text{Corr}(\mathbf{S})$; (b) Correlación de las observaciones, $\text{Corr}(\mathbf{X})$.

Este hecho se ve reflejado en la dispersión de las muestras de las observaciones **X** mostradas en la siguiente figura:

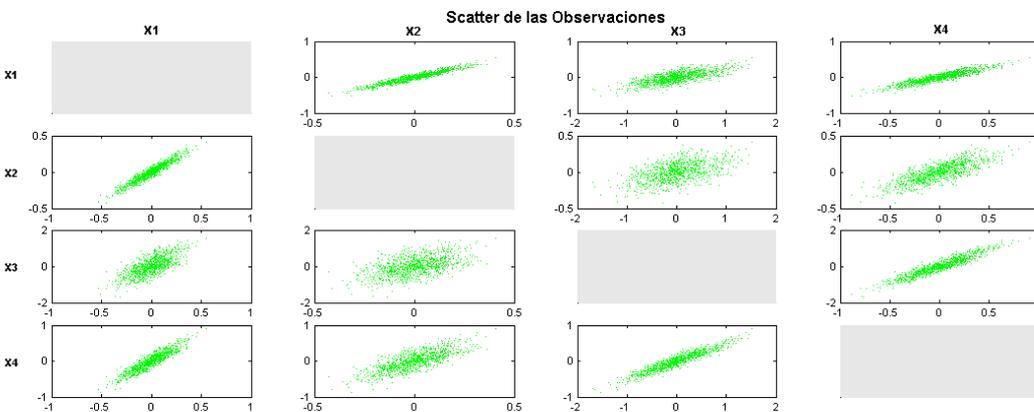


Figura 5.23 Dispersión de las muestras de las observaciones **X** correladas.

Tras la aplicación del algoritmo de extracción *ThinICA* se comprueba que las salidas obtenidas se corresponden con cada una de las fuentes originales, aunque en un orden distinto, como se muestra en la *figura 5.24*.

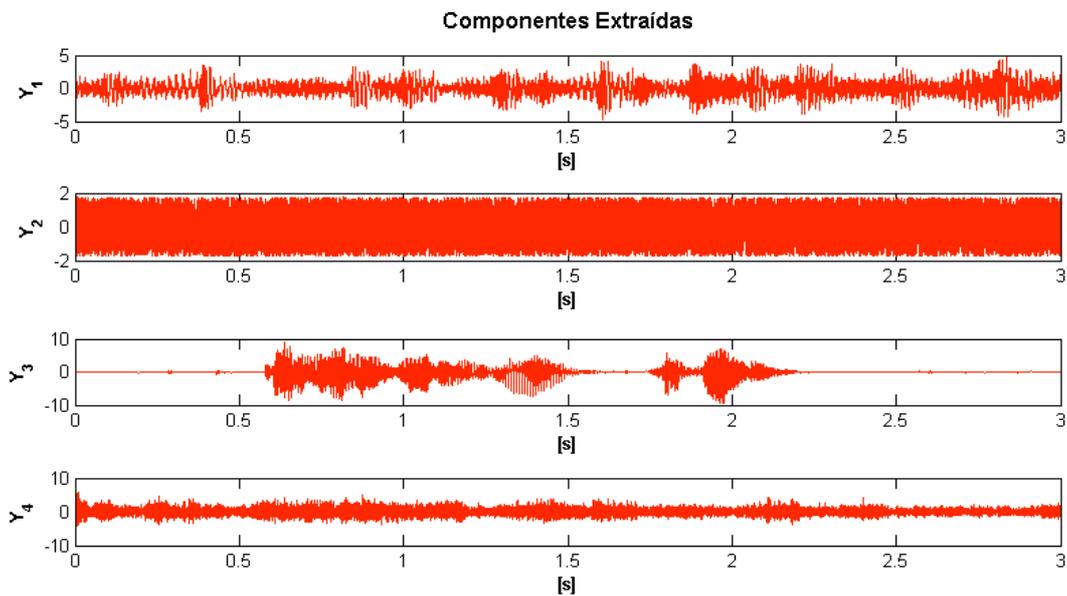


Figura 5.24 Salidas obtenidas Y . En orden descendente: música ambiente, ruido, conferenciante y murmullo.

La función de transferencia global del sistema y la relación de potencias P_Y/P_E obtenidas en la extracción demuestran la calidad de la extracción realizada por el algoritmo:

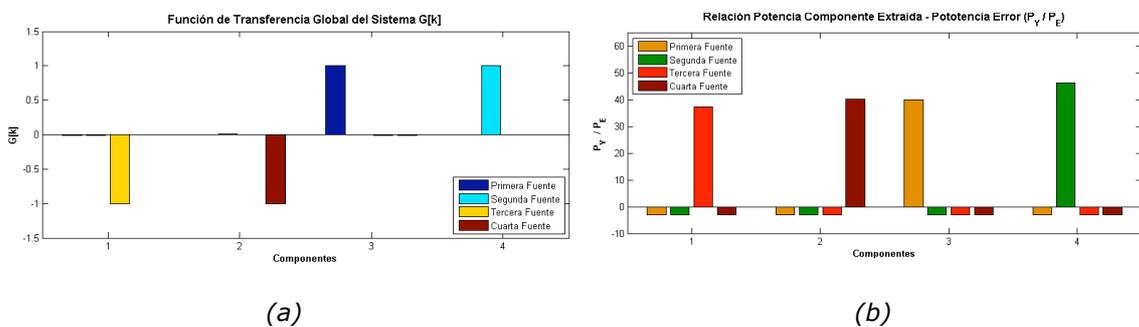


Figura 5.25 Extracción múltiple de señales de audio. (a) Función de Transferencia Global del Sistema G . (b) Relación de potencias P_Y/P_{E_j} de las fuentes extraídas.

Tras una batería de 50 simulaciones obtenemos la media del ascenso en la función contraste del algoritmo así como la media de la evolución del índice de separación de las componentes P_i . En la siguiente figura se muestra la evolución de ambas medias a lo largo de las iteraciones.

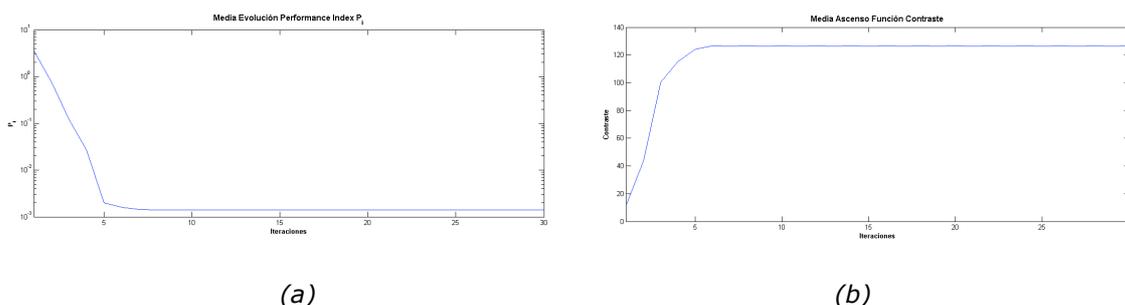


Figura 5.26 Media de la evolución a lo largo de las iteraciones tras 50 simulaciones. (a) Índice de separación P_i , (b) Ascenso en a función contraste.

Observando ambas gráficas podemos decir que se alcanza el límite del índice de separación mucho antes que en el resto de experimentos, en la séptima iteración. La rapidez en la separación está unida a la calidad, ya que la respuesta a este tipo de fuentes es la que mejor resultado ha obtenido:

| Iteraciones | P_i |
|-------------|-----------|
| 7 | 1.304 E-3 |

Tabla 5.4 Media de la extracción múltiple de señales de audio.

La gráfica que se presenta a continuación muestra la media del performance index P_i con la evolución de la Relación Señal a Ruido. Para cada nivel de SNR se han realizado una batería de 50 simulaciones, obteniéndose la media:

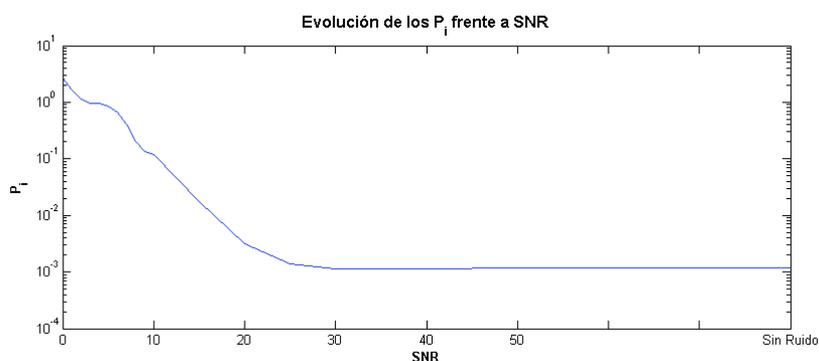


Figura 5.27 Evolución del Índice de separación P_i frente al incremento de la SNR en la separación de señales de audio.

Al igual que ocurre en todos y cada uno de los experimentos realizados, para valores muy bajos de SNR la separación es defectuosa o de muy mala calidad. A partir de una relación señal a ruido en torno a 15 decibelios se puede considerar la separación satisfactoria. Con un nivel de 25 decibelios, el algoritmo realiza la separación con su límite de calidad.

5.5. Biomedicina

Uno de los campos más importantes de aplicación de la separación ciega de señales es la biomedicina. El aumento de la tecnología hace posible la adquisición de una gran cantidad de datos que, procesados, podrían ofrecer una información muy valiosa para el diagnóstico de una posible enfermedad.

En este apartado vamos a realizar un experimento sobre un caso real de aplicación. Dicho experimento consistirá en la extracción de una señal correspondiente al latido de un feto obtenida a través de un electrocardiograma realizado sobre una mujer embarazada.

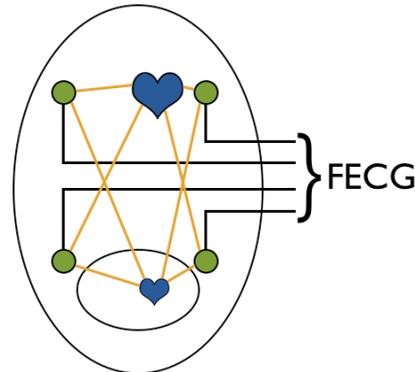


Figura 5.28 Ejemplo de aplicación de la BSS en biomedicina. Obtención del Fetal Heart Rate (FHR) mediante Electrocardiograma fetal (FECG).

5.5.1. Extracción de la Actividad Fetal

El Electrocardiograma (*ECG*) es una herramienta efectiva, barata y no invasiva que permite el diagnóstico de posibles enfermedades cardíacas. De la misma manera que es posible realizar un *ECG* a una persona adulta, se puede monitorizar el estado de salud de un feto mediante un Electrocardiograma Fetal (*FECG*). A través de la *BSS* es posible observar la forma de onda del latido del feto, pudiendo así realizar un diagnóstico más acertado sobre posibles afecciones.

Para mantener la no invasividad del mecanismo se situarán electrodos sobre la madre, captando así multitud de señales bioeléctricas, entre ellas la actividad cardíaca de la madre (*MECG*) y el feto. Las principales características que deben obtenerse del *ECG* para el diagnóstico de cardiopatologías son el ritmo cardíaco (*Fetal Heart Rate, FHR*) y las amplitudes y duraciones de las ondas que componen la *FECG*.

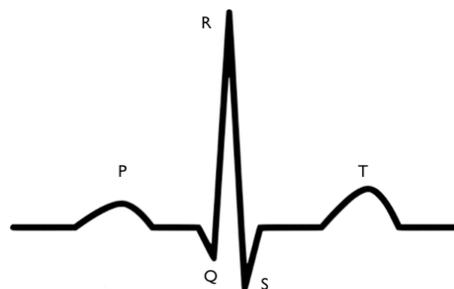


Figura 5.29 Conjunto PQRST correspondiente a una señal ECG.

Debido al tamaño del corazón del feto, la señal que genera es de muy baja intensidad, que sumado con el alto ruido procedente del *MECG* y los electrodos, hacen que la *SNR* del sistema sea muy baja. Además, dado que el *MECG* y *FECG* recogidos por los electrodos están solapados en frecuencia y tiempo no es posible un filtrado para separar ambas señales. El principal objetivo será por tanto el extraer la onda *R* de una forma estable y fiable, dado que es la que presenta una mayor amplitud.

Para tener la mejor *SNR* posible, los electrodos se dispondrán en el abdomen de la madre, donde la distancia a la fuente de interés es mínima. Por otro lado, se situarán electrodos en el pecho de la madre que, aunque no permitirán captar información alguna sobre el feto, sí permitirán visualizar de forma limpia la *MECG*. La obtención de esta *MECG* puede ayudar a su extracción *FECG* en los electrodos situados en el abdomen.

En este experimento se emplearán 8 electrodos sobre la madre, cinco de ellos situados sobre la zona abdominal y los tres restantes en la zona torácica. Las medidas empleadas tendrán una duración de cinco segundos y han sido muestreadas a 500 Hz [Callaerts] [Lathauwer02]. En la siguiente figura se muestran las observaciones captadas por estos electrodos.

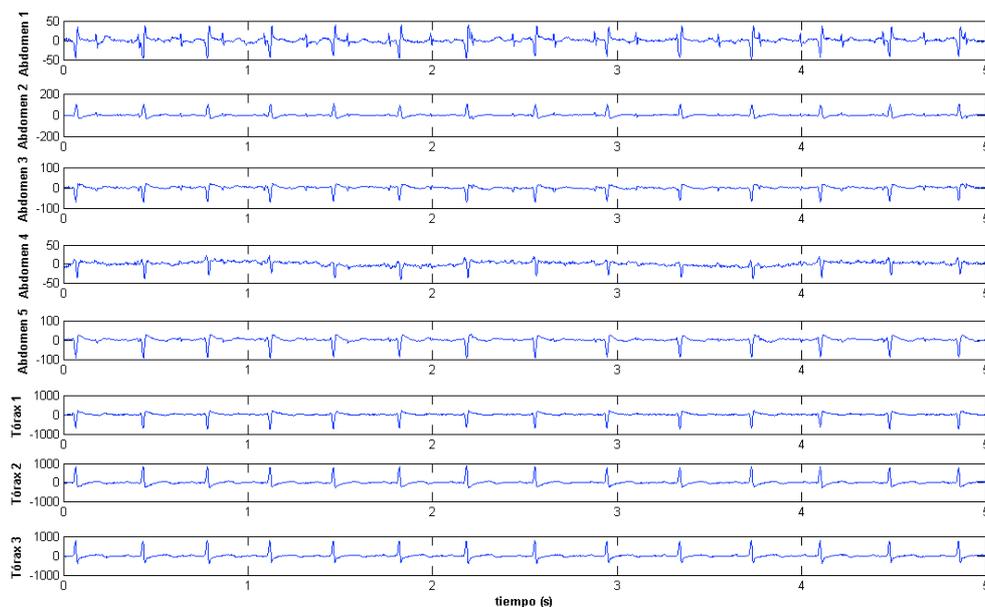


Figura 5.30 Observaciones obtenidas por los electrodos situados sobre la piel de la madre. Las cinco primeras se corresponden a electrodos situados sobre el abdomen mientras que las restantes se corresponden a observaciones obtenidas por electrodos en el tórax.

Como era de esperar, las primeras muestras procedentes del abdomen tienen una amplitud muy inferior a las obtenidas por los electrodos situados en el tórax, donde la *MECG* es dominante. Por otro lado puede observarse el efecto de la respiración materna en la muestra correspondiente al 'Abdomen 4'.

La siguiente figura muestra un detalle de la *FECG* y la *MECG* que se puede observar en la observación correspondiente al 'Abdomen1'. En ella podemos observar que la actividad cardíaca materna es más intensa mientras que la fetal tiene una mayor frecuencia.

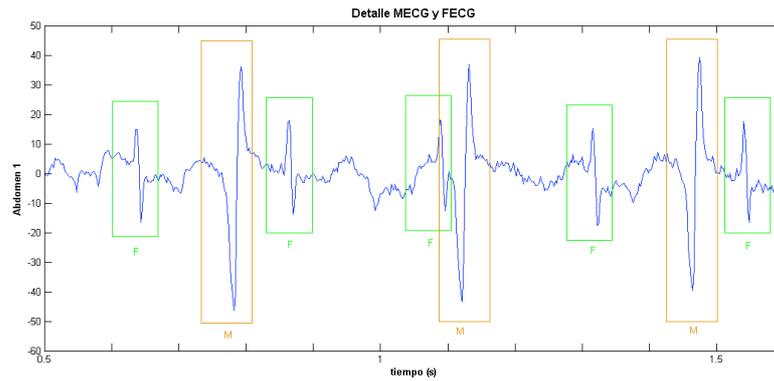


Figura 5.31 Detalle de la MEGG y FECG obtenida por el electrodo 'Abdomen1'. Recuadrado en verde se observa la componente correspondiente al feto (F). Recuadrado en rojo se observa la componente correspondiente a la madre (M).

Una vez aplicamos el algoritmo de extracción *ThinICA* se consigue separar ambas componentes, pero, mientras que la *PQRST* materna se obtiene con claridad, en la componente fetal solo pueden distinguirse la forma de onda *R* y *S*. En las siguientes figuras se muestra el resultado de la extracción.

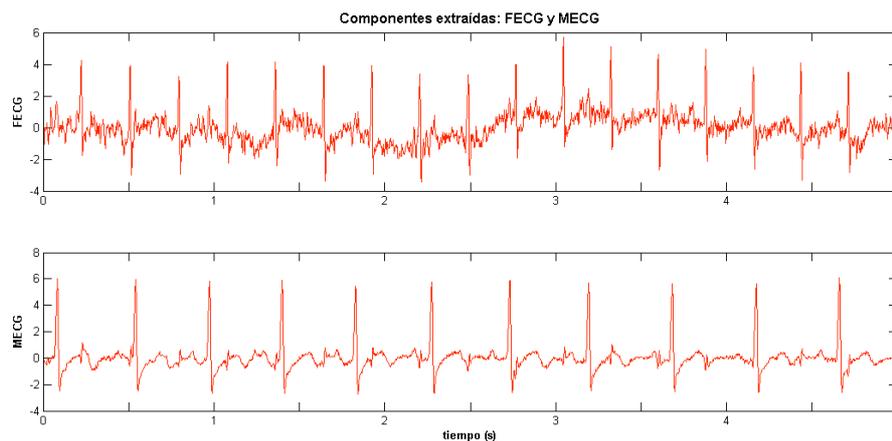


Figura 5.32 MEGG y FECG extraídas. En la parte superior se muestra la FECG donde se observa un ritmo cardíaco superior a la MEGG mostrada en la parte inferior.

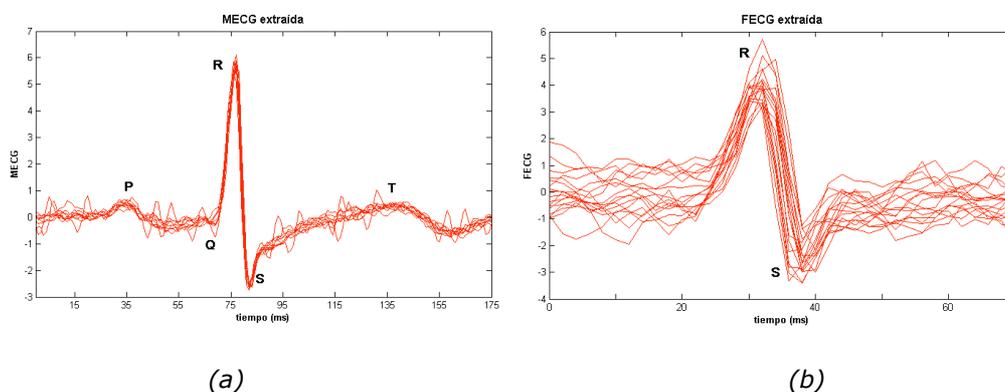


Figura 5.33 Complejo *PQRST* obtenido. En la MEGG (a) se percibe el conjunto *PQRST* de forma completa mientras que en la FECG (b) solo es apreciable las componentes *R* y *S*.

Como hemos observado, *BSS* basado en *ThinICA* es un mecanismo robusto y efectivo para la extracción *FECG* no invasiva. Aunque la calidad alcanzada en la extracción no es óptima, los resultados alcanzados pueden mejorarse haciendo uso de mecanismos de cancelación de ruido, amplificación del complejo *PQRST* o del filtrado de la *MECG*.

Debido a que a partir de los electrodos únicamente podemos obtener las observaciones \mathbf{X} pero desconocemos las fuentes \mathbf{S} , no es posible observar la evolución de los diferentes índices de prestaciones tal y como hemos realizado en apartados anteriores.

5.6. Procesado Imágenes

Otra de las posibilidades de aplicación de la separación ciega de fuentes se encuentra en el campo del procesado de imágenes. El origen de las imágenes puede ser de muy diversa naturaleza (médicas, naturales, sintéticas,...) y su procesado puede tener diferentes finalidades como, por ejemplo, la extracción de ciertas características, la compresión o reducción de ruido.



Figura 5.34 Ejemplos de imágenes de diferente naturaleza. (a) Imagen Médica, (b) Imagen Natural y (c) Imagen sintética correspondiente a un texto.

El experimento que vamos a realizar para comprobar la respuesta del algoritmo *ThinICA* desarrollado consistirá en la separación de imágenes de diferente naturaleza.

5.6.1. Imágenes Origen

Las imágenes fuente \mathbf{S} que formarán parte del experimento han sido extraídas del banco de imágenes de MATLAB. Dichas imágenes tendrán un tamaño de 128 x 128 píxeles en escala de grises, donde la intensidad de cada uno de ellos vendrá dada por 8 bits, pudiendo así conseguirse 256 niveles. La primera fuente se corresponde con una imagen sintética de ruido, la segunda con un retrato y la tercera y cuarta con imágenes naturales. En la siguiente figura se muestran dichas imágenes fuente, donde cada una de las columnas que la forma representan las imágenes originales, sus histogramas y su representación normalizada en una dimensión.

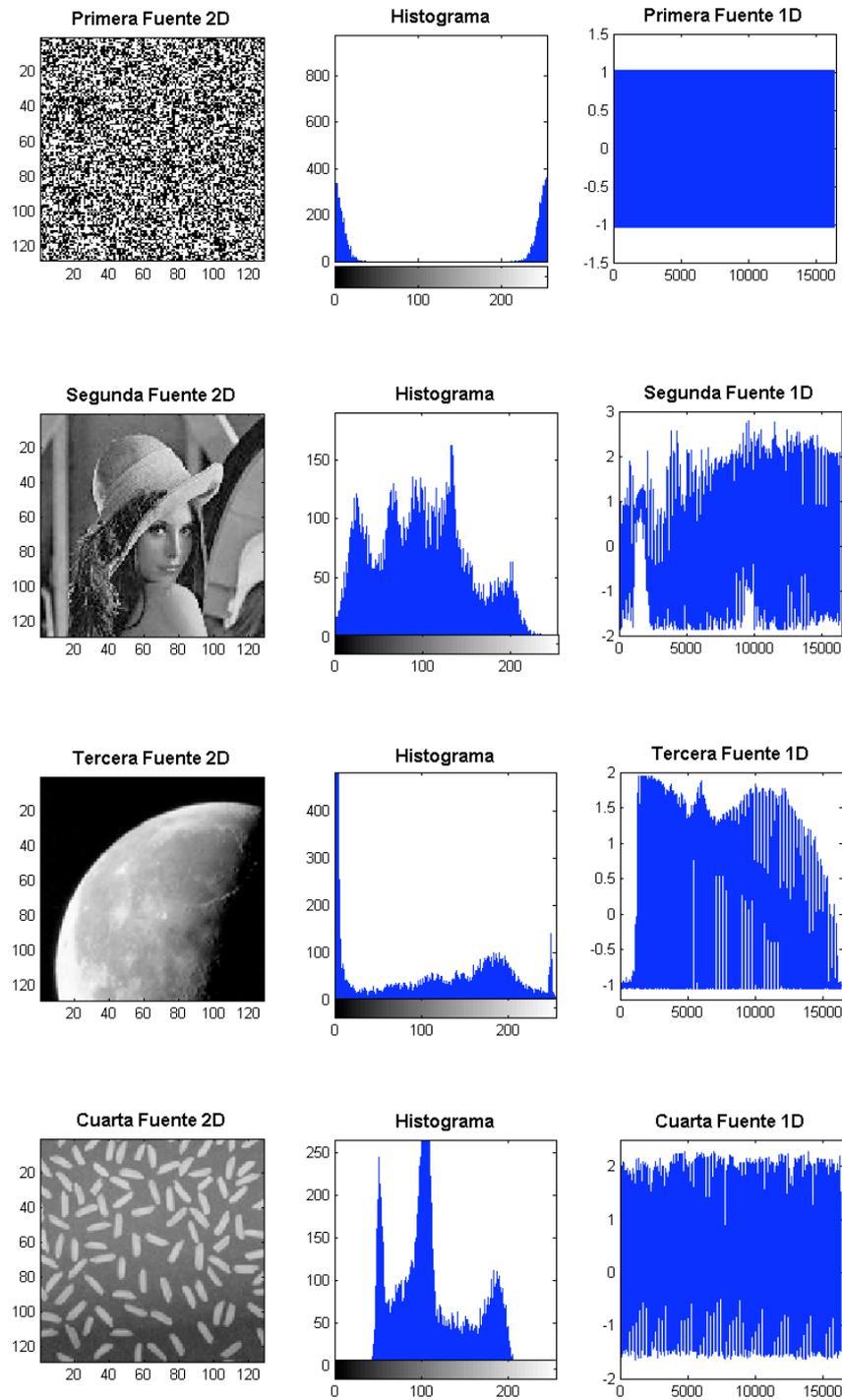


Figura 5.35 Conjunto de imágenes fuente **S**. La primera columna está formada por las fuentes en dos dimensiones, la segunda columna se corresponde con sus histogramas y la tercera columna con las fuentes normalizadas en una dimensión.

5.6.2. Extracción Múltiple de Imágenes

Estas imágenes se mezclan a través de la matriz **A** de dimensiones [7,4], obteniéndose como resultado las observaciones **X** que se muestran en la *figura*

5.36. Debido a dicha mezcla, los histogramas de las observaciones presentan una mayor gaussianeidad que los histogramas originales.

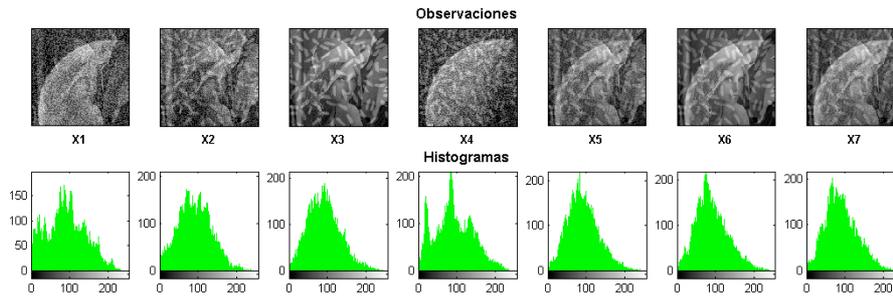


Figura 5.36 Conjunto de observaciones X . La primera fila está formada por las imágenes de las observaciones mientras que la segunda fila se corresponde con sus histogramas.

A continuación se muestran las componentes extraídas por el algoritmo así como sus histogramas:

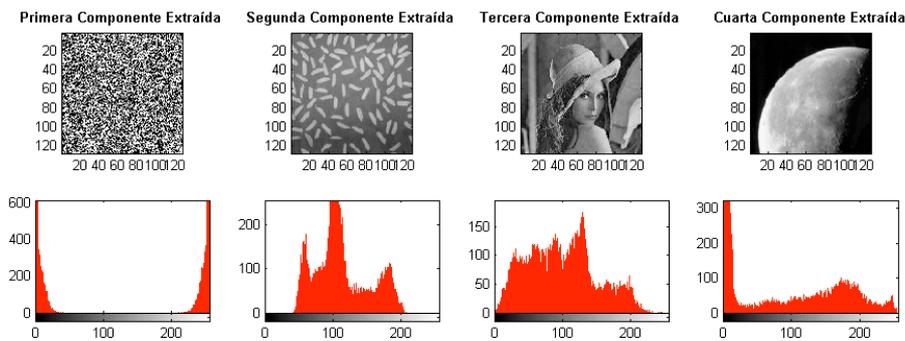


Figura 5.37 Conjunto de componentes extraídas Y . La primera fila está formada por las imágenes obtenidas mientras que la segunda fila se corresponde con sus histogramas.

Como se puede deducir al observar las imágenes obtenidas el algoritmo ha realizado una separación eficaz de las observaciones. La siguiente figura muestra la evolución en la separación de la tercera componente:

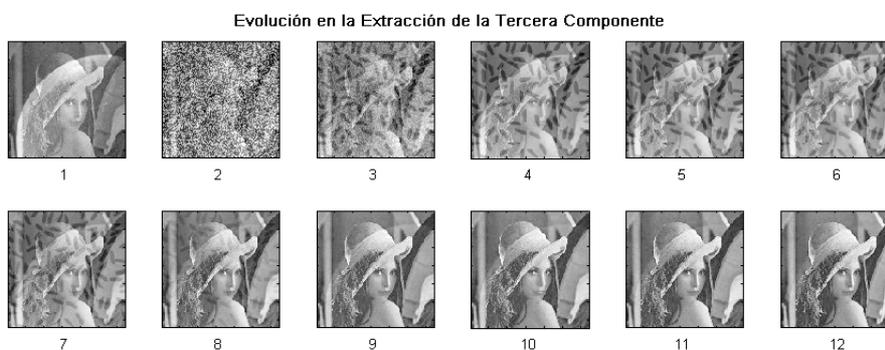


Figura 5.38 Evolución en la extracción de la tercera componente a lo largo de las iteraciones. A partir de la novena iteración el resultado es satisfactorio.

Ratificando este resultado tenemos las gráficas de la función de transferencia global G junto con la relación P_Y/P_E mostradas en la siguiente figura:

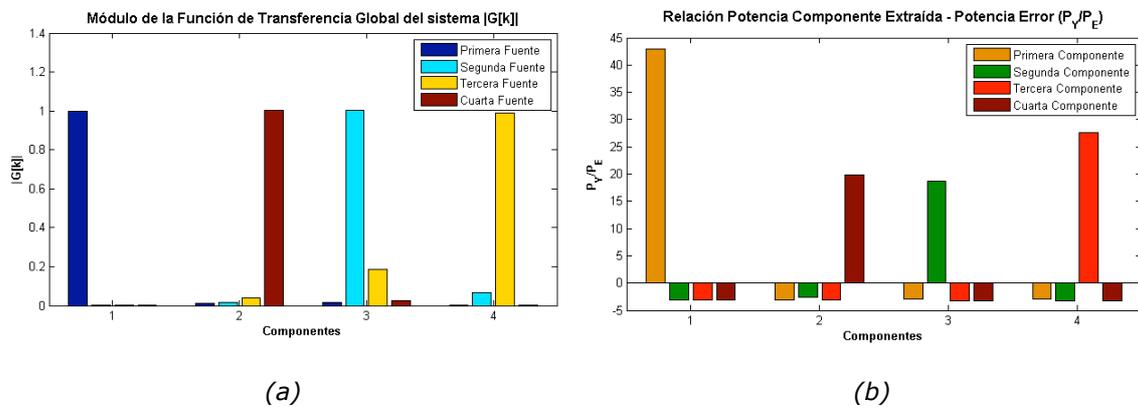


Figura 5.39 Extracción múltiple de imágenes. (a) Módulo de la Función de Transferencia Global del Sistema. (b) Relación de las potencias P_Y/P_{E_j} de las fuentes extraídas.

La siguiente figura muestra la media de la evolución del índice de separación P_i y del ascenso de la función contraste a lo largo de las iteraciones, obtenidas tras una batería de 50 simulaciones. En estas gráficas podemos observar que el ascenso se hace de una forma más rápida que en el resto de casos de estudio. Mientras que en el resto de aplicaciones se necesita una media de más de diez iteraciones para llegar al mejor nivel de separación, en la separación de imágenes el número es menor. Este resultado puede venir propiciado por el hecho de que la separación de las componentes se realiza con una calidad menor que en el resto de aplicaciones.

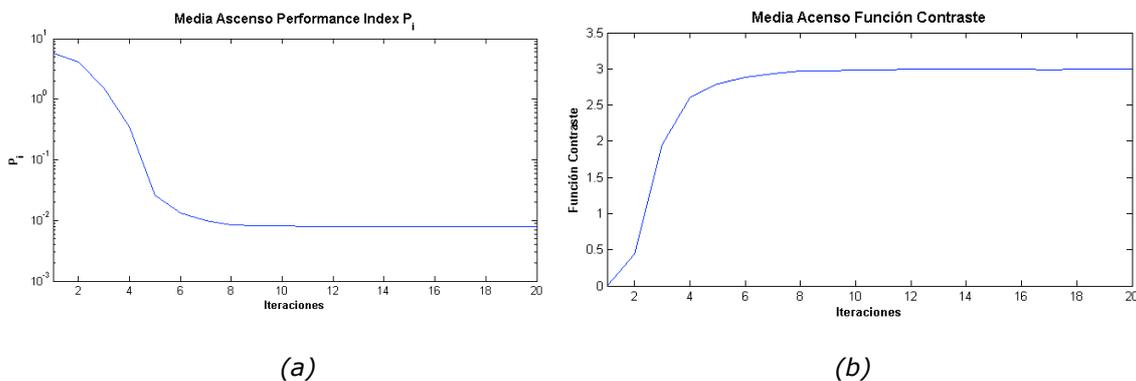


Figura 5.40 Media de la evolución a lo largo de las iteraciones tras 50 simulaciones. (a) Índice de separación P_i , (b) Ascenso en a función contraste.

La siguiente tabla muestra un resumen de los valores medios del número de iteraciones empleadas y del *Performance Index* del experimento sobre señales para comunicaciones:



| Iteraciones | P_i |
|-------------|-----------|
| 9 | 8.137 E-3 |

Tabla 5.5 Media de la extracción múltiple de fuentes tipo imagen.

A continuación se presenta una gráfica con la evolución del índice de separación P_i con respecto la mejora en la Relación Señal a Ruido. Para valores muy bajos de SNR el índice de separación resulta elevados y los resultados no son satisfactorios. Para un P_i superior a 15 decibelios la separación se considera aceptable, llegando a su límite de calidad en torno a los 25 decibelios.

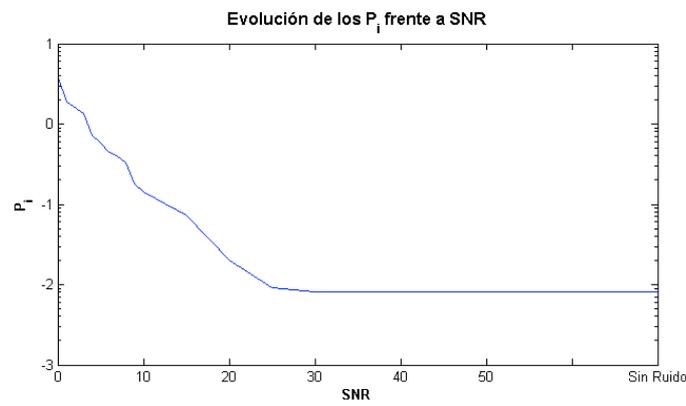


Figura 5.41 Evolución del Índice de separación P_i frente al incremento de la SNR en la separación de imágenes.

5.7. Telecomunicaciones

Son muchas las posibilidades de aplicación de los algoritmos BSS y entre ellas se encuentra las telecomunicaciones. Éste es un campo de trabajo muy amplio y no es objetivo del proyecto el tratar todos los casos posibles de aplicación. En este apartado nos limitaremos únicamente a tratar de separar una serie de fuentes que se encuentran en el mismo medio y que presentan diferentes modulaciones. La particularidad del experimento que a continuación se presenta radica en la naturaleza compleja de las fuentes a extraer.

5.7.1. Extracción Múltiple de Modulaciones

El siguiente experimento consiste en la extracción de varias fuentes de forma simultánea. Las señales originales en este caso se corresponden con diferentes modulaciones digitales empleadas en comunicaciones: $BPSK$, $\pi/4$ - $QPSK$, 16 - QAM y 4 - PCM .

Estas fuentes se mezclan a partir de una matriz de mezclas \mathbf{A} aleatoria uniforme, el número de sensores es igual al de fuentes, coincidiendo también con el número de fuentes a extraer ($M \equiv N \equiv p$).



En la *figura 5.42* se muestra el resultado la extracción, donde puede observarse que el orden de las componentes obtenidas ha variado con respecto al original y se ha producido un pequeño giro en alguna de ellas.

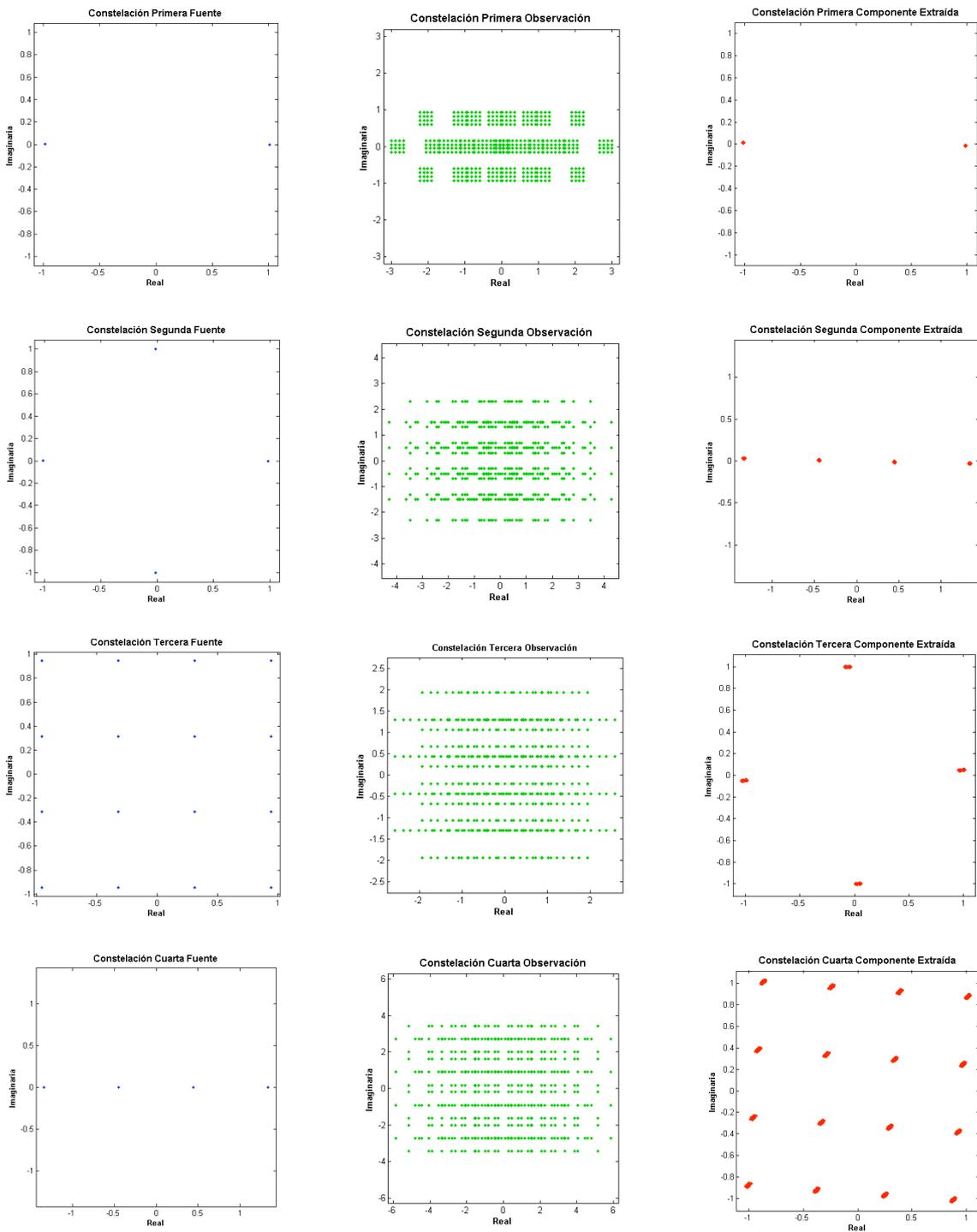


Figura 5.42 Separación de señales para comunicaciones. La primera columna se corresponde con las fuentes originales **S**: BPSK, $\pi/4$ -QPSK, 16-QAM y 4-PCM. La segunda columna se corresponde con las señales observadas **X**. La tercera columna se corresponde con las señales extraídas **Y**: BPSK, 4-PCM, $\pi/4$ -QPSK y 16-QAM.



La función de transferencia global del sistema **G** junto con la relación P_Y/P_E obtenida en el experimento se muestra en la *figura 5.43*.

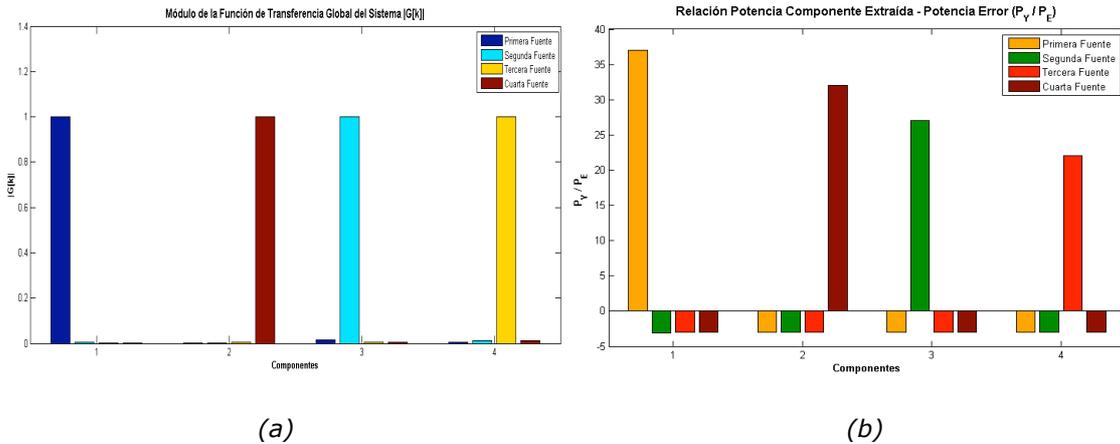


Figura 5.43 Extracción múltiple de fuentes para comunicaciones. (a) Módulo de la Función de Transferencia Global del Sistema. (b) Relación de potencia P_Y/P_E de las fuentes extraídas.

En la siguiente figura se muestra la media de la evolución del índice de separación P_i y del ascenso de la función contraste a lo largo de las iteraciones, obtenidas tras una batería de 50 simulaciones.

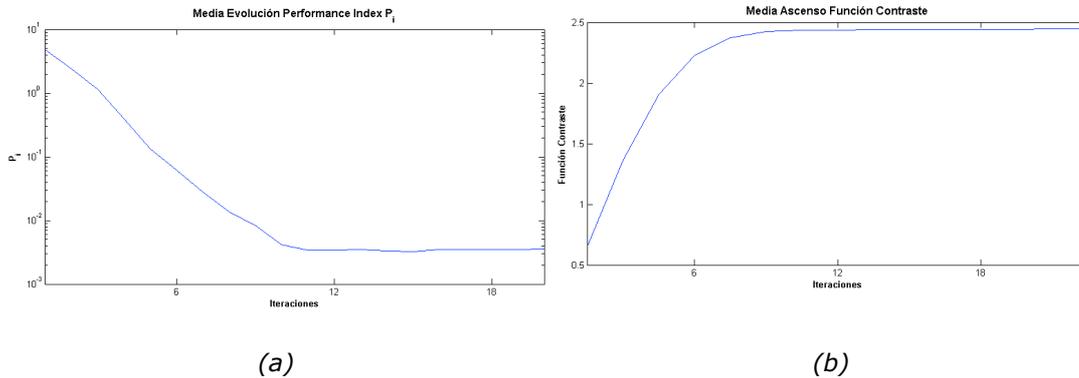


Figura 5.44 Media de la evolución a lo largo de las iteraciones tras 50 simulaciones. (a) Índice de separación P_i , (b) Ascenso en a función contraste.

La siguiente tabla muestra un resumen de los valores medios del número de iteraciones empleadas y del *Performance Index* del experimento sobre señales para comunicaciones:

| Iteraciones | P_i |
|-------------|-----------|
| 11 | 3.459 E-3 |

Tabla 5.6 Media de la extracción múltiple de señales para comunicaciones.

A continuación se presenta una gráfica con el nivel medio del *Performance Index* obtenido en baterías de 50 simulaciones. En éstas se ha ido incrementado el valor de la Relación Señal a Ruido, con la correspondiente mejora en las prestaciones, hasta llegar a la ausencia del mismo.

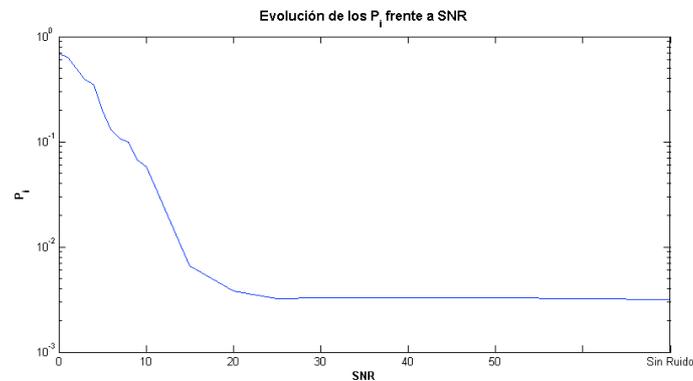


Figura 5.45 Evolución del Índice de separación P_i frente al incremento de la SNR en la separación de señales para comunicaciones.

Como hemos podido comprobar en las imágenes anteriores, el algoritmo de extracción *ThinICA* presenta resultados satisfactorios ante fuentes de naturaleza compleja. Si comparamos con estos mismos resultados con las señales de laboratorio propuestas en el apartado 5.1 observamos que las prestaciones que ofrece el algoritmo no se ven mermadas ante el cambio de naturaleza de las fuentes.

5.8. Conclusiones

Este capítulo ha sido la culminación de todo el estudio teórico realizado a lo largo del proyecto y donde se han puesto en práctica gran parte de los conceptos incluidos previamente.

Previo al desarrollo de las simulaciones se realizó una revisión sobre los diferentes índices de prestaciones existentes para la medida de las cualidades de los algoritmos. Entre ellos se definió el *Performance Index* P_i que posteriormente se empleó para la comparación con los algoritmos de referencia.

Antes de comenzar con los casos reales de aplicación se realizaron diferentes experimentos sobre señales *experimentales* que nos ayudaron a comprobar el correcto funcionamiento del algoritmo implementado. Durante dichos experimentos pudimos comprobar el avance en la separación de las fuentes a lo largo de las iteraciones a través de los distintos índices de prestaciones empleados. Gracias a una batería de simulaciones pudimos extraer la media del ascenso en la función contraste y la media de la evolución del índice de prestaciones P_i . Los resultados obtenidos dejaron entrever que la extracción simple se realiza de una forma más veloz y de mayor calidad que la extracción múltiple. Por otro lado, realizamos una nueva batería de simulaciones donde esta vez variamos la relación señal a ruido de las fuentes y pudimos comprobar que con una $SNR=13$ dB los resultados son muy satisfactorios. Para terminar con las señales *de laboratorio* comparamos el modelo *ThinICA* implementado con los algoritmos de referencia y obtuvimos una mejor calidad de la extracción a costa de un excesivo tiempo de cómputo.



Tras el análisis de señales *experimentales* pasamos a extraer señales reales de audio. El experimento realizado consistió en separar la voz procedente de una persona que se encuentra en un ambiente donde reina un murmullo, existe música ambiental y ruido. En este experimento comprobamos la dispersión de las fuentes antes y después del proceso de mezcla, donde pudimos ver cómo la matriz de correlaciones dejó de ser la matriz identidad. Tras una batería de simulaciones llegamos a la conclusión de que la calidad en la separación de este tipo de señales es muy alta y se alcanza mucho antes que en el resto de experimentos.

En el siguiente apartado llegó el turno de las señales biomédicas. El experimento realizado consistió en la extracción de la actividad cardíaca fetal de manera no invasiva a través de electrodos situados sobre la madre. La señal procedente del corazón del feto (*FECG*) es de muy baja intensidad y se encuentra enmascarada por la de la madre (*MECG*) por lo que únicamente pudimos extraer con calidad las componentes R y S del complejo PQRST. En este experimento, al ser señales reales no pudimos comprobar el avance del algoritmo a lo largo de las iteraciones ante la falta de una señal de referencia original.

El cuarto experimento nos llevó al procesado de imágenes. Dicho experimento consistió en la separación de cuatro imágenes de diversa naturaleza (retrato, naturales y artificial) que se encontraron superpuestas. Dado el carácter gráfico del experimento pudimos apreciar el proceso de mezcla y separación de las fuentes a través de las imágenes intermedias obtenidas. En este experimento volvimos a realizar una batería de simulaciones donde extrajimos la media del ascenso en la función contraste, la media de la evolución del índice P_i y la respuesta del algoritmo ante la variación de la *SNR*. Como conclusión podemos decir que la separación de las imágenes fuente es rápida, con una calidad satisfactoria y que alcanza su límite a partir de una relación señal a ruido de 25 dB.

El último experimento realizado se llevó a cabo a partir de señales para comunicaciones. Este experimento consistió en la separación de señales procedentes de modulaciones digitales: *BPSK*, $\pi/4$ -*QPSK*, *16-QAM* y *4-PCM*. La particularidad del mismo radica en la naturaleza compleja de algunas de las fuentes originales. Observamos cómo el proceso de mezcla hace que las distintas constelaciones queden ocultas y cómo tras el paso por el algoritmo de separación vuelven a estar de nuevo visibles, eso sí, desordenadas y con un pequeño giro. El algoritmo implementado vuelve a tener un comportamiento acertado consiguiendo, de media, separar las fuentes tras once iteraciones y su calidad no se ve mermada ante el cambio de naturaleza de las fuentes.

