



Máster en Ingeniería de Sistemas y Control

Trabajo de Fin de Máster

Estudio de métodos de determinación de número óptimo de componentes independientes en señales multicanal

Alumno: Daniel Mercader Rodríguez

Directores: Raquel Dormido
Natividad Duro
Ángel Mur

Curso 2016-2017
Convocatoria: Junio 2017

Máster en Ingeniería de Sistemas y Control

**Estudio de métodos de determinación
de número óptimo de componentes
independientes en señales multicanal**

Proyecto tipo A: Proyecto específico propuesto por un profesor

Alumno: Daniel Mercader Rodríguez

Directores: Raquel Dormido
Natividad Duro
Ángel Mur



Autorización

Autorizamos a la Universidad Complutense de Madrid (UCM) y a la Universidad Nacional de Educación a Distancia (UNED) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a sus autores, tanto la memoria de este Trabajo Fin de Máster, como el código, la documentación y/o prototipo desarrollado.

Firmado: Daniel Mercader Rodríguez

Firma del Alumno

Agradecimientos

En primer lugar me gustaría agradecer a los directores de este trabajo final de máster, Raquel Dormido, Natividad Duro y Ángel Mur, por darme la oportunidad de desarrollar este trabajo, así como por su paciencia, disponibilidad y guía durante el desarrollo del mismo.

También quiero agradecer a la Universidad Nacional de Educación a Distancia la oportunidad de desarrollar estos estudios que en otras modalidades de estudio presencial me hubiera sido imposible compatibilizar con el trabajo.

De igual forma agradecer este trabajo muy especialmente a Elena, por aguantarme durante tantas horas pegado al ordenador y en definitiva apoyarme incondicionalmente para que haya podido completar estos estudios.

Por último, no puedo olvidar a mi familia, especialmente mis padres, que siempre estuvieron ayudándome y mostrándome el camino que ha hecho posible llegar hasta aquí, suyo es el mérito.

A todos muchas gracias.

Estudio de métodos de determinación de número óptimo de componentes independientes en señales multicanal

Resumen:

En este trabajo fin de máster se presenta un estudio sobre diferentes métodos para la determinación del número óptimo de componentes independientes en señales multicanal.

Cuando se desarrolla un modelo de Análisis de Componentes Independientes (ICA) sobre un conjunto de señales multicanal, éstas se suponen formadas por diversas mezclas de otro conjunto de señales independientes entre sí. El Análisis de Componentes Independientes trata de extraer precisamente estas Componentes Independientes, de las que en numerosas ocasiones no se tiene ningún tipo de información a priori, hasta el punto de que ni siquiera se conoce el número de señales que forman este conjunto de señales independientes.

Se han analizado diversos métodos para la determinación del Número Óptimo de Componentes Independientes (NOCI) a partir del conjunto de señales de mezcla y sin necesitar ningún conocimiento previo sobre las Componentes Independientes. Los principales métodos para determinación de NOCI son el Criterio de Durbin-Watson, el método 'ICA by blocks', el método 'RV ICA by blocks', método basado en Análisis de Componentes Principales (PCA) y recientemente el método basado en Correlación Lineal de Componentes (LCC).

En este trabajo se presentan dichos métodos, se han implementado en *MATLAB* y se han aplicado sobre diferentes conjuntos de señales de prueba sintéticas y reales para demostrar su funcionamiento, comparar los métodos y adquirir experiencia en su utilización. Finalmente se presentan las principales características de cada uno de los métodos y una metodología de utilización de los mismos.

Fruto del trabajo desarrollado en este proyecto fin de máster, he participado como coautor en la siguiente publicación indexada: [\[Mur et al., 2017\]](#).

Palabras clave: análisis de componentes independientes, ICA, número óptimo componentes independientes, análisis componentes principales, PCA, correlación, MATLAB, métodos, distancia correlación, estudio

Índice general

Índice de figuras	7
Índice de tablas	9
1. Introducción	11
1.1. Objetivos	11
1.1.1. 'Cocktail-Party Problem'	11
1.1.2. Separación Ciega de Fuentes	12
1.1.3. Número Óptimo de Componentes Independientes (NOCI)	13
1.2. Organización de la Memoria	14
2. Análisis de Componentes Independientes (ICA)	15
2.1. Análisis de Componentes Independientes (ICA)	15
2.1.1. Definición	15
2.1.2. ICA ilustrado gráficamente	16
2.1.3. Restricciones en ICA	16
2.1.4. Ambigüedades en ICA	18
2.2. Algoritmos ICA	18
2.2.1. FastICA	19
2.2.2. Joint Approximate Diagonalization of Eigenmatrices (JADE)	19
3. Métodos de determinación del Número Óptimo Componentes Independientes	21
3.1. Criterio de Durbin-Watson	21
3.2. Método 'ICA by blocks'	23
3.2.1. 'Random ICA by Blocks'	25
3.3. Método 'RV ICA by blocks'	25
3.4. Método basado en Análisis de Componentes Principales (PCA)	26
3.5. Algoritmo Linear Correlations between Components (LCC)	28

4. Resultados del estudio	31
4.1. Señales de prueba	31
4.1.1. Independencia de las señales de prueba	31
4.2. Ensayo sobre el conjunto de señales <i>Set1a</i>	33
4.2.1. Criterio DW sobre el conjunto de señales <i>Set1a</i>	33
4.2.2. Método 'ICA by blocks' sobre el conjunto de señales <i>Set1a</i>	34
4.2.3. Método 'RV ICA by blocks' sobre el conjunto de señales <i>Set1a</i>	34
4.2.4. Método 'PCA' sobre el conjunto de señales <i>Set1a</i>	34
4.2.5. Método 'LCC' sobre el conjunto de señales <i>Set1a</i>	34
4.3. Ensayo sobre el conjunto de señales <i>Set1</i>	39
4.3.1. Criterio DW sobre el conjunto de señales <i>Set1</i>	40
4.3.2. Método 'ICA by blocks' sobre el conjunto de señales <i>Set1</i>	40
4.3.3. Método 'RV ICA by blocks' sobre el conjunto de señales <i>Set1</i>	40
4.3.4. Método 'PCA' sobre el conjunto de señales <i>Set1</i>	44
4.3.5. Método 'LCC' sobre el conjunto de señales <i>Set1</i>	44
4.4. Ensayo sobre el conjunto de señales <i>Set2</i>	44
4.4.1. Criterio DW sobre el conjunto de señales <i>Set2</i>	46
4.4.2. Método 'ICA by blocks' sobre el conjunto de señales <i>Set2</i>	46
4.4.3. Método 'RV ICA by blocks' sobre el conjunto de señales <i>Set2</i>	48
4.4.4. Método 'PCA' sobre el conjunto de señales <i>Set2</i>	48
4.4.5. Método 'LCC' sobre el conjunto de señales <i>Set2</i>	48
4.5. Ensayo sobre el conjunto de señales <i>Set3</i>	48
4.5.1. Criterio DW sobre el conjunto de señales <i>Set3</i>	53
4.5.2. Método 'ICA by blocks' sobre el conjunto de señales <i>Set3</i>	53
4.5.3. Método 'LCC' sobre el conjunto de señales <i>Set3</i>	53
4.6. Análisis de tiempo de ejecución	56
5. Conclusiones	59
5.1. Características de los métodos	59
5.1.1. Características principales del método Criterio de Durbin-Watson (DW)	59
5.1.2. Características principales del método ' <i>ICA by blocks</i> '	59
5.1.3. Características principales del método ' <i>RV ICA by blocks</i> '	60
5.1.4. Características principales del método basado en ' <i>PCA</i> '	61

5.1.5. Características principales del método 'LCC'	61
5.2. Conclusiones	62
5.3. Futuros Trabajos	62
A. Guía rápida de funciones	65
A.1. Guía rápida de funciones	65
A.1.1. DW.m	65
A.1.2. RV_coeff.m	65
A.1.3. matrixdcorr.m	65
A.1.4. DWCriterion.m	66
A.1.5. ICA_by_two_blocks.m	66
A.1.6. RV_ICA_by_two_blocks.m	66
A.1.7. PCA_NOCI.m	67
A.1.8. LCC_NOCI.m	67
A.1.9. Test01.m	67
A.1.10. Test02.m	67
A.1.11. Test03.m	68
A.1.12. Test04.m	68
A.1.13. Test05.m	68

Índice de figuras

2.1. ICA gráficamente.	17
3.1. Conjuntos de señales de prueba	22
3.2. Ejemplo de aplicación del Criterio de DW a 100 señales de mezcla obtenidas a partir de las señales independientes del conjunto <i>Set 1</i>	23
3.3. Ejemplo de aplicación del método 'ICA_by_two_blocks' a 100 señales de mezcla obtenidas a partir de las señales independientes del conjunto <i>Set 2</i>	25
3.4. Ejemplos de aplicación <i>RV ICA by blocks</i>	27
3.5. Ejemplo de aplicación del método basado en PCA a 100 señales de mezcla obtenidas a partir de las señales independientes del conjunto <i>Set2</i>	27
3.6. Representación algoritmo LCC.	30
4.1. Aplicación del criterio DW sobre mezclas de señales <i>Set1a</i>	35
4.2. Aplicación del método ' <i>ICA by blocks</i> ' sobre mezclas de señales <i>Set1a</i>	36
4.3. Aplicación del método ' <i>RV ICA by blocks</i> ' sobre mezclas de señales <i>Set1a</i>	37
4.4. Aplicación del método ' <i>PCA</i> ' sobre mezclas de señales <i>Set1a</i>	38
4.5. Aplicación del método ' <i>LCC</i> ' sobre mezclas de señales <i>Set1a</i>	39
4.6. Aplicación del criterio DW sobre mezclas de señales <i>Set1</i>	41
4.7. Aplicación del método ' <i>ICA by blocks</i> ' sobre mezclas de señales <i>Set1</i>	42
4.8. Aplicación del método ' <i>RV ICA by blocks</i> ' sobre mezclas de señales <i>Set1</i>	43
4.9. Aplicación del método ' <i>PCA</i> ' sobre mezclas de señales <i>Set1</i>	45
4.10. Aplicación del método ' <i>LCC</i> ' sobre mezclas de señales <i>Set1</i>	46
4.11. Aplicación del criterio DW sobre mezclas de señales <i>Set2</i>	47

4.12. Aplicación del método ' <i>ICA by blocks</i> ' sobre mezclas de señales <i>Set2</i>	49
4.13. Aplicación del método ' <i>RV ICA by blocks</i> ' sobre mezclas de señales <i>Set2</i>	50
4.14. Aplicación del método ' <i>PCA</i> ' sobre mezclas de señales <i>Set2</i>	51
4.15. Aplicación del método ' <i>LCC</i> ' sobre mezclas de señales <i>Set2</i>	52
4.16. Aplicación del criterio DW sobre mezclas de señales <i>Set3</i>	54
4.17. Aplicación del método ' <i>ICA by blocks</i> ' sobre mezclas de señales <i>Set3</i>	55
4.18. Aplicación del método ' <i>LCC</i> ' sobre mezclas de señales <i>Set3</i>	56

Índice de tablas

4.1. Características de las señales de prueba.	32
4.2. Distancia correlación de las señales <i>Set1</i>	32
4.3. Distancia correlación de las señales <i>Set2</i>	32
4.4. Distancia correlación de las señales <i>Set3</i>	33
4.5. Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de <i>Set1a</i>	33
4.6. Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de <i>Set1</i>	39
4.7. Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de <i>Set2</i>	44
4.8. Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de <i>Set3</i>	53
4.9. Tiempos de ejecución con el algoritmo DW en segundos. .	57
4.10. Tiempos de ejecución con el algoritmo 'ICA by blocks' en segundos.	57
4.11. Tiempos de ejecución con el algoritmo 'RV ICA by blocks' en segundos.	57
4.12. Tiempos de ejecución con el algoritmo 'PCA' en segundos.	58
4.13. Tiempos de ejecución con el algoritmo 'LCC' en segundos.	58

Capítulo 1

Introducción

1.1. Objetivos

1.1.1. 'Cocktail-Party Problem'

Imaginemos que estamos en una habitación donde dos personas están hablando simultáneamente y tenemos dos micrófonos situados en localizaciones distintas. Estos micrófonos nos proporcionan dos señales grabadas en el dominio del tiempo que denotaremos como $x_1(t)$ y $x_2(t)$. Cada una de estas señales grabadas está compuesta por una suma ponderada de las señales emitidas por cada una de las personas que están hablando, denotaremos estas señales emitidas como $s_1(t)$ y $s_2(t)$. Esto se puede representar mediante una ecuación lineal:

$$\begin{aligned}x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) \\x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t)\end{aligned}\tag{1.1}$$

Los parámetros a_{ij} son constantes que dependerán de las distancias de cada uno de los micrófonos a las personas que están hablando.

Sería extremadamente útil poder estimar las señales originales de las dos personas que hablan $s_1(t)$ y $s_2(t)$ a partir únicamente de las dos señales grabadas $x_1(t)$ y $x_2(t)$. Mientras que en $x_1(t)$ y $x_2(t)$ tenemos ambas voces mezcladas, en $s_1(t)$ y $s_2(t)$ tendremos de forma individualizada la locución de cada una de las dos personas que están hablando. Este problema se conoce como el *cocktail-party problem* [Haykin and Chen, 2005].

Si se conocieran a priori los parámetros a_{ij} , el problema sería extremadamente sencillo, pero la dificultad radica precisamente en el desconocimiento de dichos parámetros.

De forma más generalizada el problema se hace mucho más complejo si además desconocemos el número de personas que están hablando en la habitación.

Una aproximación para resolver el problema sería utilizar información de las propiedades estadísticas de las señales $s_1(t)$ y $s_2(t)$ para estimar el valor de los parámetros a_{ij} . Es suficiente considerar que $s_1(t)$ y $s_2(t)$ son estadísticamente independientes en cada instante de tiempo t .

1.1.2. Separación Ciega de Fuentes

La metodología para la **Separación Ciega de Fuentes** proporciona una solución al *cocktail-party problem* presentado en 1.1.1. Esto es, permitirá recuperar señales independientes desconocidas a partir de una combinación lineal de las mismas. Además se puede observar en gran medida el trasfondo del Análisis de Componentes Independientes (ICA).

Análogamente a lo expuesto en el 1.1.1, consideremos varias fuentes que emiten un cierto número de señales y además existe una serie de sensores que van a permitir obtener una grabación de dichas señales mezcladas con unos ciertos coeficientes de ponderación. Para simplificar y a efectos ilustrativos, consideremos 3 fuentes de señal y 3 receptores o sensores. Llamemos $x_1(t)$, $x_2(t)$ y $x_3(t)$ a las señales obtenidas por los receptores y $s_1(t)$, $s_2(t)$ y $s_3(t)$ a las señales emitidas por las fuentes, siendo la relación entre ambas señales:

$$\mathbf{x} = \mathbf{A} \times \mathbf{s} \quad (1.2)$$

En 1.2, \mathbf{A} es una matriz de mezcla 3×3 con los coeficientes a_{ij} constantes que determinan la relación entre las fuentes y las señales grabadas, $\mathbf{x} = [x_1(t), x_2(t), x_3(t)]^T$ es un vector con las señales grabadas y $\mathbf{s} = [s_1(t), s_2(t), s_3(t)]^T$ es un vector de las señales emitidas por las fuentes.

De forma general los coeficientes de la matriz \mathbf{A} son desconocidos porque se desconocen las propiedades físicas del medio de transmisión de las señales emitidas por las fuentes.

El problema de la **separación ciega de fuentes** consiste en obtener las señales originales \mathbf{s} a partir del único conocimiento de las señales mezcladas.

Se asumirá que la matriz de mezcla \mathbf{A} en 1.2 es invertible, de forma que se puede decir que existe una matriz \mathbf{W} que permite obtener las señales originales \mathbf{s} a partir de las observaciones:

$$\mathbf{s} = \mathbf{W} \times \mathbf{x} \quad (1.3)$$

Si conociéramos los coeficientes de la matriz \mathbf{A} se podría obtener \mathbf{W} como la inversa de \mathbf{A} . El problema radica precisamente en el desconocimiento de \mathbf{A} . Todo lo que se tiene son las señales mezcladas \mathbf{x} y se desea encontrar una matriz \mathbf{W} que permita estimar las señales fuente originales \mathbf{s} mediante 1.3 con la mayor precisión posible. Una posible solución puede

encontrarse suponiendo únicamente que existe **independencia estadística** entre las señales fuente originales. De hecho, si las señales originales fuente son *no-Gaussianas*, esto es suficiente para determinar los coeficientes de la matriz \mathbf{W} , de modo que las señales

$$\mathbf{y} = \mathbf{W} \times \mathbf{x} \quad (1.4)$$

son estadísticamente independientes. Si las señales $\mathbf{y} = [y_1(t), y_2(t), y_3(t)]^T$ son independientes entre sí, entonces serán iguales a las señales originales $\mathbf{s} = [s_1(t), s_2(t), s_3(t)]^T$.

ICA puede utilizarse para resolver el problema de *separación ciega de fuentes*; de hecho, fue inicialmente desarrollado para resolver este problema y otros similares. Sin embargo, en la actualidad ICA despierta un interés mayor debido a la posibilidad de aplicarlo a un amplio abanico de problemas.

Consideremos, por ejemplo, la grabación de las señales eléctricas obtenidas mediante Electroencefalografía (EEG). Los datos de EEG consisten en la grabación de los potenciales eléctricos obtenidos mediante electrodos distribuidos en diferentes localizaciones en el peri-cráneo. Se puede presumir que las señales grabadas serán la mezcla de algún tipo de componentes subyacentes de la actividad cerebral. Así considerado este problema tienen grandes similitudes con el *cocktail-party problem* y en teoría, puede utilizarse ICA para obtener las componentes originales de la actividad cerebral. Sin embargo, la principal utilización de ICA con las señales EEG es el filtrado de los artefactos [Vigário, 1997].

Otra aplicación diferente de ICA es la extracción de características en imágenes, señales de audio u otros tipos de datos [Hyvriinen, 1999].

1.1.3. Número Óptimo de Componentes Independientes (NOCI)

ICA, como se verá en el siguiente capítulo, se basa en la construcción de Componentes Independientes (CIs). El procedimiento para determinar el NOCI, k , es un serio problema cuando se desarrolla un modelo ICA [Mur et al., 2017], especialmente cuando no se tiene una idea previa sobre el número de componentes independientes que puede haber en las señales bajo estudio. El NOCI a usar en un modelo ICA es un parámetro importante, y cuando no se conoce *a priori* su valor se debe invertir tiempo y esfuerzo en su determinación [Bouweresse and Rutledge, 2016].

En este trabajo se presentan los principales métodos para la determinación de NOCI. Se aplican dichos métodos a diferentes conjuntos de señales de prueba con el objetivo de conocer su funcionamiento, conocer las ventajas e inconvenientes de cada uno de ellos, y realizar una comparativa. En definitiva se trata de un estudio de los métodos de determinación de número óptimo de componentes independientes en señales multicanal.

1.2. Organización de la Memoria

La memoria se organiza en 4 capítulos. El primer capítulo es una introducción al trabajo que se pretende desarrollar. El segundo capítulo presenta una introducción a ICA. En el tercer capítulo se presentan los diferentes métodos de determinación de Número Óptimo de Componentes Independientes (NOCI) utilizados durante el estudio. Para los distintos métodos se ha desarrollado código en *MATLAB* con el objetivo de poder aplicarlos a diferentes señales de prueba. En el cuarto capítulo se hace una presentación del estudio realizado y particularmente de los resultados obtenidos. En los anexos se presenta el código desarrollado para cada uno de los métodos.

Capítulo 2

Análisis de Componentes Independientes (ICA)

2.1. Análisis de Componentes Independientes (ICA)

2.1.1. Definición

Se expone aquí someramente la definición de ICA que se puede encontrar en la introducción de [Hyvärinen et al., 2004].

Dado un conjunto de observaciones de variables aleatorias $\mathbf{x} = [x_1(t), x_2(t), \dots, x_n(t)]^T$, siendo t el tiempo o el índice de las muestras, se asumirá que están generadas por una combinación lineal de componentes independientes, el modelo ICA se define como:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \mathbf{A} \begin{pmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_n(t) \end{pmatrix} \quad (2.1)$$

donde \mathbf{A} es una matriz de mezcla desconocida y $\mathbf{s} = [s_1(t), s_2(t), \dots, s_n(t)]^T$ es un vector con las componentes independientes, también desconocido. La ecuación 2.1 describe por lo tanto cómo son generados los datos observados mediante la mezcla de las fuentes s_i .

ICA consistirá en estimar tanto la matriz \mathbf{A} como las fuentes \mathbf{s} a partir de las observaciones \mathbf{x} . Se puede suponer que el número de observaciones coincide con el número de fuentes independientes originales, si bien esta simplificación no es completamente necesaria para resolver el problema y en determinadas ocasiones puede no ser aceptable.

De forma alternativa se podría definir ICA como el problema de obtención de la transformación lineal dada por la \mathbf{W} tal que las variables

aleatorias estimadas $\mathbf{y} = [y_1(t), y_2(t), \dots, y_n(t)]^T$ mediante:

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \mathbf{W} \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \quad (2.2)$$

sean tan independientes como sea posible. Una vez obtenida la matriz \mathbf{W} , la matriz \mathbf{A} se obtiene como su inversa, y viceversa.

El problema estará completamente definido, es decir, el modelo planteado puede ser estimado sí y sólo sí las componentes $\mathbf{s} = [s_1(t), s_2(t), \dots, s_n(t)]^T$ son *no-Gaussianas*. Esta condición es esencial y sirve para explicar la diferencia fundamental que existe entre ICA y PCA, dónde la *no-Gaussianidad* no se tiene en cuenta.

2.1.2. ICA ilustrado gráficamente

En [Hyvärinen and Oja, 2000] se puede encontrar una muestra del funcionamiento de ICA de una forma gráfica y sencilla. Se consideran dos CIs generadas con una distribución uniforme y valores en el rango $[-1, 1]$ denominadas $\mathbf{s} = [s_1(t), s_2(t)]^T$. Ambas componentes se ilustran en la figura 2.1a. Su función de densidad de probabilidad se ilustra en la figura 2.1c(izquierda), simplemente representando las muestras generadas aleatoriamente de una componente con respecto a la otra.

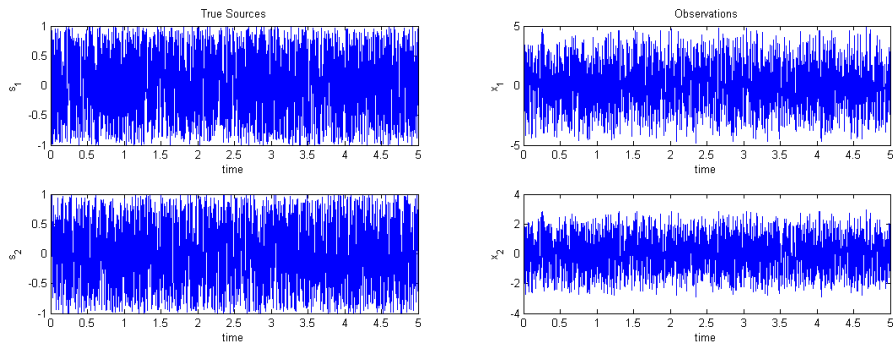
Si se considera una matriz de mezcla $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$ se pueden obtener dos variables mezcladas utilizando 1.2. Estas variables mezcladas se muestran en la figura 2.1b. Estas variables mezcladas no son independientes, puesto que ambas contienen información de \mathbf{s} , esto se corrobora mediante su función de densidad de probabilidad conjunta que se muestra en la figura 2.1c(derecha).

La figura 2.1c(derecha) muestra una forma intuitiva de obtener la matriz de mezcla \mathbf{A} y por tanto de resolver ICA en un caso sencillo como este, con distribuciones uniformes. Como se puede comprobar los ejes del paralelogramo generado están en la dirección de las columnas de \mathbf{A} .

2.1.3. Restricciones en ICA

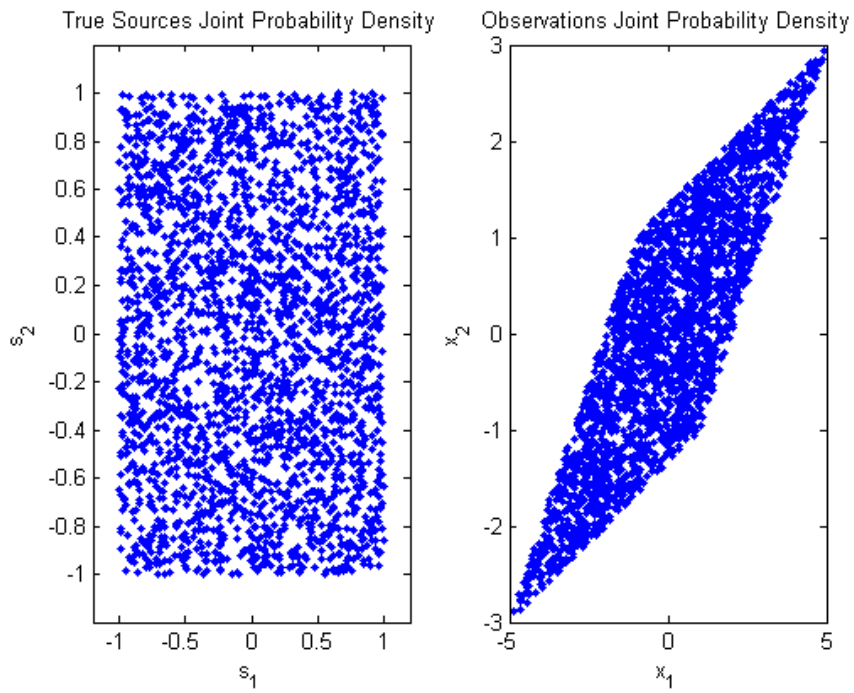
Para que el modelo ICA planteado en 2.1.1 tenga solución es necesario establecer una serie de restricciones/suposiciones de acuerdo a [Hyvärinen et al., 2004]. Estas restricciones se enumeran a continuación:

1. *El número de observaciones ha de ser mayor o igual al número de componentes independientes a estimar.* En el caso de que sea mayor,



(a) Variables Independientes.

(b) Variables Mezcladas.



(c) Funciones de densidad de probabilidad conjunta.

Figura 2.1: ICA gráficamente.

es posible aplicar PCA para reducir su dimensión al número de CIs a estimar.

2. *Las CIs se supondrán estadísticamente independientes.* En este principio descansa la base de ICA.
3. *Las CIs deben presentar distribuciones no gaussianas.* ICA se basa en estadísticos de orden superior, que toman el valor cero para distribuciones gaussianas y estos estadísticos son básicos en ICA, por lo que se puede deducir que ICA es imposible para este tipo de fuentes.
4. *Por simplicidad, se considera que la matriz de mezcla \mathbf{A} es cuadrada e invertible.* Que \mathbf{A} sea cuadrada se desprende de la primera restricción. Que la matriz sea invertible permite obtener \mathbf{W} de forma sencilla mediante \mathbf{A}^{-1} .

2.1.4. Ambigüedades en ICA

En el modelo ICA introducido en el apartado 2.1.1 se presentan una serie de ambigüedades que es necesario tener en cuenta. Dichas ambigüedades se exponen con detalle en [Hyvärinen and Oja, 2000]:

1. *No se puede determinar el orden de las CIs.* Esto es debido a que tanto \mathbf{s} como \mathbf{A} son desconocidas a priori, de forma que si se plantea el modelo ICA como $\mathbf{x} = \sum_{i=1}^n \mathbf{a}_i s_i$ siendo \mathbf{a}_i cada una de las columnas de la matriz \mathbf{A} y s_i las fuentes, entonces en todo momento sería posible cambiar el orden de los términos del sumatorio y colocar cualquiera de las componentes independientes en primer lugar.
2. *No es posible determinar las energías de las CIs.* Es decir, estimaremos las CIs con la incertidumbre de un factor de escala, que en ocasiones puede implicar incluso un cambio de signo. La forma más habitual de proceder es suponer que las CIs a estimar van a tener varianza unidad.

2.2. Algoritmos ICA

Los algoritmos de separación ciega de fuentes se pueden clasificar, según el método utilizado para la separación, en *algoritmos ICA estadísticos* y *algoritmos ICA geométricos*.

Los *algoritmos ICA estadísticos* utilizan propiedades estadísticas como cumulantes, kurtosis, etc., en la búsqueda de la independencia estadística de las componentes de salida.

Los *algoritmos ICA geométricos* utilizan la distribución geométrica de las señales provenientes de las fuentes para la estimación de la matriz de mezclas.

A continuación se presentan dos de los algoritmos ICA más utilizados en la actualidad.

2.2.1. FastICA

Es un algoritmo desarrollado por Aapo Hyvärinen y Erkki Oja [Hyvärinen and Oja, 2000] del tipo ICA *estadístico* y recibe su nombre por su rapidez con respecto a otros algoritmos. El algoritmo es una adaptación de una algoritmo de punto fijo de una regla de aprendizaje de una red neuronal, encontrando las componentes independientes una a una. Utiliza técnicas estadísticas y analíticas, llegando a utilizar estadísticos de 4º orden. Es uno de los mejores algoritmos ICA.

2.2.2. Joint Approximate Diagonalization of Eigenmatrices (JADE)

Las siglas JADE son el acrónimo de *Joint Approximate Diagonalization of Eigenmatrices* y se puede traducir como *Unión de Diagonalizaciones Aproximadas de Automatrices*. Fue desarrollado por J.F. Cardoso [Cardoso and Souloumiac, 1993]. Al igual que FastICA, se trata de un algoritmo de tipo *estadístico* que utiliza técnicas analíticas y estadísticas para obtener las componentes independientes, utilizando también estadísticos de hasta 4º orden. Para el estudio realizado en este trabajo se ha utilizado Joint Approximate Diagonalization of Eigenmatrices (JADE) como algoritmo ICA de referencia, ya que proporciona resultados muy estables y es el algoritmo utilizado en algunos trabajos previos como [Mur et al., 2017].

Capítulo 3

Métodos de determinación del Número Óptimo Componentes Independientes

En este capítulo se presentan brevemente los métodos de determinación de Número Óptimo de Componentes Independientes (NOCI) que se han estudiado en este trabajo y para lo cual han sido implementados en *MATLAB* [MATLAB, 2014]. En la mayoría de casos se trata de algoritmos relativamente sencillos de implementar. El algoritmo ICA utilizado en todos los casos ha sido JADE. Los detalles del *Criterio DW*, *ICA by Blocks* y *RV ICA by Blocks* se han obtenido de [Hyvärinen et al., 2004] y [Bouweresse and Rutledge, 2016]. La información sobre el método basado en PCA se ha obtenido de [Mur et al., 2017], donde también se ha presentado recientemente el algoritmo Linear Correlations between Components (LCC), método no supervisado para obtener NOCI de forma automática. Las figuras que se muestran en este capítulo han sido generadas en *MATLAB* mediante la aplicación de los métodos presentados a mezclas obtenidas a partir de dos conjuntos de señales de prueba. El conjunto *Set 1* es de producción propia, mientras que el conjunto *Set 2* es el conjunto de señales presentado y utilizado en [Mur et al., 2017] con la única diferencia de la componente aleatoria de cada una de las señales. El conjunto de señales *Set 1* se muestra en la figura 3.1a, y el conjunto de señales *Set 2* se muestra en la figura 3.1b. En el capítulo 4 se presentan con mas detalle estos conjuntos de señales junto con otros utilizados.

3.1. Criterio de Durbin-Watson

El DW es un estadístico utilizado para detectar la presencia de *autocorrelación* en los residuos de un análisis de regresión.

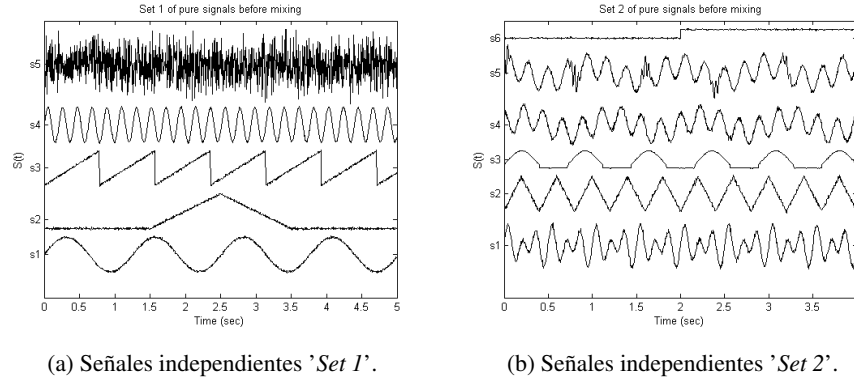


Figura 3.1: Conjuntos de señales de prueba

En [Rutledge and Barros, 2002] se propone el DW como un estimador de la Relación Señal-Ruido (SNR) en el caso de *señales estructuradas*. Su valor es cercano a 0 cuando no hay mucho ruido en las señales; incrementando a medida que el ruido aumenta, y alcanza un valor cercano a 2 cuando el contenido de la señal es ruido es su práctica totalidad. El valor del criterio de DW aplicado a una señal e de longitud n se obtiene mediante:

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (3.1)$$

DW puede utilizarse como un criterio de validación de ICA utilizando el siguiente procedimiento recomendado en [Bouweresse et al., 2012]:

- Se computan $maxCIs$ modelos ICA de 1 a $maxCIs$ CIs.
- Para cada modelo:
 - Se estima una aproximación de la matriz original \mathbf{X} con 2.1. La matriz estimada en el modelo con k -CIs se denotará como \mathbf{X}_k .
 - Las matrices residuales \mathbf{R}_k se computan mediante $\mathbf{R}_k = \mathbf{X} - \mathbf{X}_k$.
 - Para cada fila de \mathbf{R}_k , correspondiente a la señal residual se calcula el DW. Si el DW se aproxima a 0, entonces todavía hay cierta estructura en la señal residual. Si el DW empieza a crecer, significa que la señal residual pierde estructura y hay más presencia de ruido.

Una vez se han obtenido $maxCIs$ matrices de residuos y n veces el valor del criterio de DW para cada una de esas matrices, se puede construir una matriz de dimensiones $n \times maxCIs$ que representada gráficamente ayuda

a seleccionar el valor de NOCI. En la figura 3.2 se muestra un ejemplo de representación gráfica de esta matriz. En ella se puede ver que el valor de DW se mantiene bastante bajo para todas las señales mezcladas cuando el número de CIs está entre 1 y 3, presenta valores mínimos para 4 CIs y a partir de 5 comienza a crecer. En este caso se podría deducir que el NOCI recomendado por este método es 5, ya que es el primer valor alto en el que la matriz de residuos comienza a contener mucho mas ruido que información, este resultado coincide con el valor esperado para NOCI.

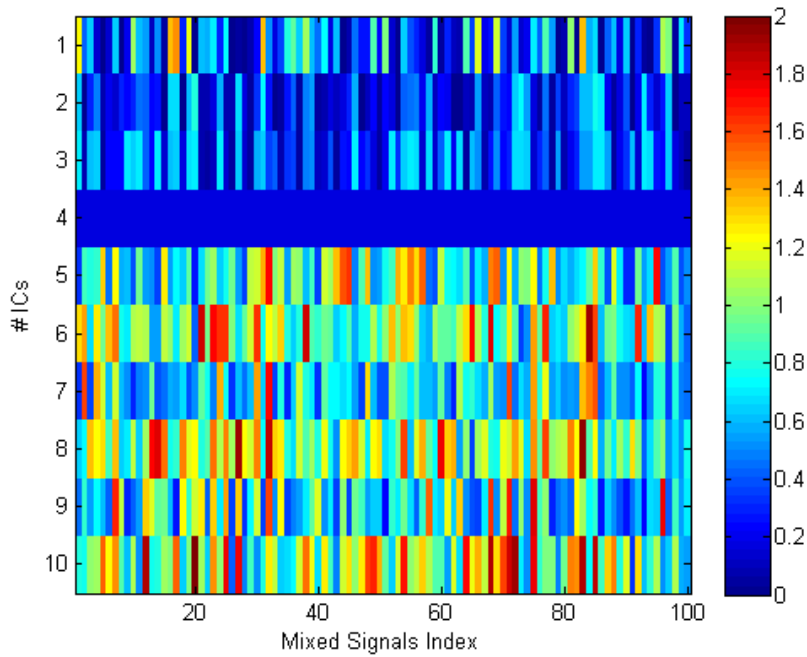


Figura 3.2: Ejemplo de aplicación del Criterio de DW a 100 señales de mezcla obtenidas a partir de las señales independientes del conjunto *Set 1*.

3.2. Método 'ICA by blocks'

En [Bouweresse et al., 2012] se presenta este método para la determinación del NOCI. El método consiste en dividir el conjunto de señales mezcladas en dos o más subgrupos (bloques) representativos de la señales de mezcla original. Es decir, la matriz de señales mezcladas \mathbf{X} se divide en dos o mas submatrices (bloques) que contienen líneas de \mathbf{X} suficientemente representativas del conjunto de mezclas original.

La forma en que se construyen los bloques es muy importante para el éxito final del método y por tanto debe ponerse especial cuidado en esta parte del proceso. Los bloques construidos deben tener aproximadamente

el mismo tamaño y además deben tener un tamaño suficiente como para permitir aplicar ICA y obtener suficientes CIs. Para cada uno de los bloques se computan $maxCIs$ modelos ICA de 1 a $maxCIs$ CIs y se calculan las correlaciones entre cada una de las CIs extraídas de un subconjunto con cada una de las CIs extraídas de cada uno de los demás subconjuntos. La idea en la que se fundamenta este método es que si una Componente Independiente (CI) es suficientemente representativa, está será extraída igualmente en todos los subconjuntos analizados [Hyvärinen et al., 2004].

Sin embargo, debido a las ambigüedades de ICA el signo y el orden de las CIs extraídas no será el mismo en todos los subconjuntos, lo que dificulta la identificación de dichas CIs significativas que hayan sido extraídas en todos los subconjuntos. Cuando ya se hayan extraído demasiadas CIs, éstas comenzarán a contener características de ruido asociadas a cada uno de los subgrupos y por tanto esas CIs ruidosas tendrán menores valores de correlación con las CIs extraídas de otros subgrupos. Análogamente cuando se obtenga un número insuficiente de CIs éstas se corresponderán con mezclas de las verdaderas señales independientes puras en diferentes proporciones, por lo que tendrán correlaciones menores con las CIs de otros subconjuntos.

NOCI se define como el mayor número de CIs para el cual todas las CIs extraídas de un bloque están altamente correladas con una CI de otro bloque.

En la figura 3.3 se muestra un ejemplo de aplicación del método 'ICA by two blocks' sobre un conjunto de 100 señales mezcladas obtenidas mediante una matriz de mezcla A generada aleatoriamente y aplicada sobre el conjunto completo de señales de prueba *Set 2*. En este caso se ha utilizado $maxCIs = 16$, por lo que se han computado 16 modelos ICA extrayendo CIs de 1 a $maxCIs$ sobre cada uno de los dos bloques generados, en total 32 modelos ICA.

En este caso, cada uno de los bloques contiene 50 señales mezcladas. Cada una de las series mostradas en la figura 3.3 se corresponde con los perfiles de correlación de cada uno de los pares de modelos ICA, uno de cada bloque. Así, para $maxCIs = 1$ se computará un modelo ICA para cada bloque, se obtendrán las correlaciones correspondientes y se representa el mayor valor de correlación; para $maxCIs = 2$ se computará un modelo ICA para cada bloque, se obtendrán las correlaciones correspondientes y se representarán los dos mayores valores de correlación. En general, para $maxCIs$ se obtendrá un modelo ICA para cada bloque, se computarán las correlaciones y se representarán las $maxCIs$ correlaciones de mayor valor.

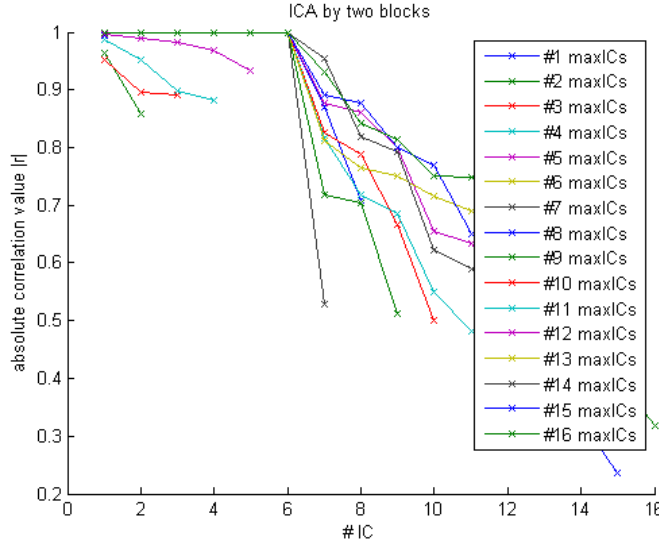


Figura 3.3: Ejemplo de aplicación del método 'ICA_by_two_blocks' a 100 señales de mezcla obtenidas a partir de las señales independientes del conjunto *Set 2*.

En la figura 3.3 se puede comprobar que los modelos ICA obtenidos para $maxICs = 6$ obtienen los mayores valores de correlación, de hecho están muy próximos a 1. Para $maxICs < 6$ los valores de correlación son altos pero algo menores que en el caso óptimo, y para $maxICs > 6$ los valores de correlación empiezan a disminuir de forma rápida. El resultado de aplicar este método indica que en este caso $NOCI = 6$.

3.2.1. 'Random ICA by Blocks'

El resultado obtenido con 'ICA by Blocks' depende principalmente de la distribución de las señales entre los dos (o más) bloques [Hyvärinen et al., 2004].

En este sentido, para evitar obtener distribuciones de bloques que pueden no representar la verdadera estructura de los datos, se puede aplicar una versión de ICA by Blocks que podemos llamar 'Random ICA by blocks' y consiste en repetir un cierto número de veces 'ICA by Blocks' sobre distribuciones aleatorias de las señales de muestra entre los bloques, todos los resultados de correlación son luego representados en una gráfica mostrando la repetibilidad de los resultados obtenidos.

3.3. Método 'RV ICA by blocks'

El coeficiente RV mide la similaridad entre dos matrices de datos. Dicho coeficiente RV se calcula entre dos matrices de señales y puede ser

interpretado como cualquier otro coeficiente de correlación, tomando valores entre 0 y 1. Cuando el coeficiente RV entre dos matrices toma valor 1, eso significa que las matrices de varianza/covarianza de las dos matrices de datos son muy similares y por tanto los datos que contienen son muy similares. Cuanto menor sera el valor del coeficiente RV menos similares serán las matrices de datos comparadas [Smilde et al., 2009]. El coeficiente RV entre dos matrices con señales almacenadas en las filas de la matriz se define en 3.2.

$$RV(\mathbf{X}_1, \mathbf{X}_2) = \frac{\text{trace}(\mathbf{X}_1^T \mathbf{X}_1 \mathbf{X}_2^T \mathbf{X}_2)}{\sqrt{\text{trace}(\mathbf{X}_1^T \mathbf{X}_1 \mathbf{X}_1^T \mathbf{X}_1) \text{trace}(\mathbf{X}_2^T \mathbf{X}_2 \mathbf{X}_2^T \mathbf{X}_2)}} \quad (3.2)$$

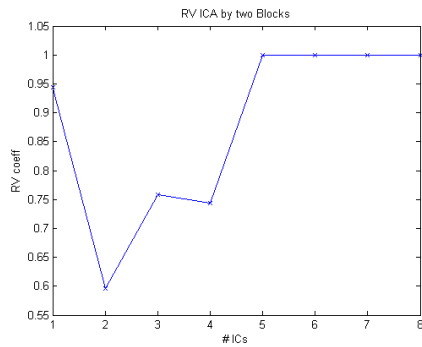
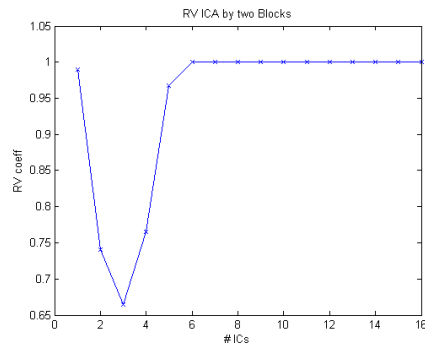
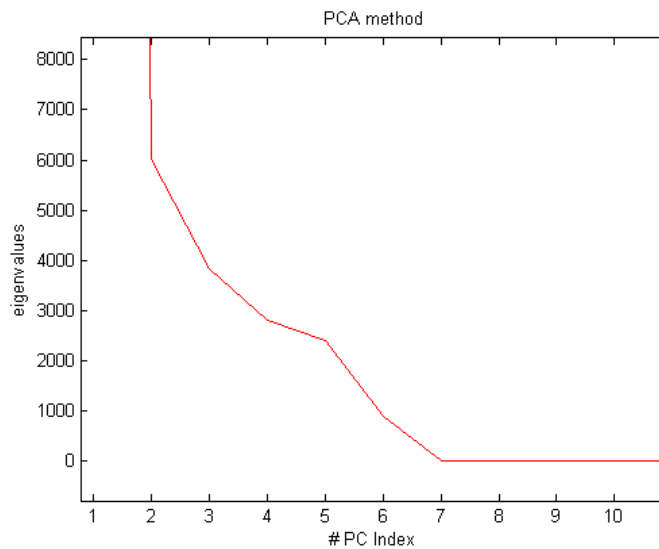
En el método '*RV ICA by Blocks*' los datos se dividen en dos subconjuntos (bloques) suficientemente representativos y se computan *maxCIs* modelos ICA de 1 a *maxCIs* CIs. Para cada modelo ICA se obtiene el coeficiente RV entre las matrices *S* obtenida de cada bloque, es decir entre las matrices de componentes independientes obtenidas de cada bloque [Hyvärinen et al., 2004]. Se obtiene por tanto un valor de coeficiente RV para cada número de CIs obtenida en el modelo. Se puede representar gráficamente el coeficiente RV respecto al número de CIs del modelo obtenido. Análogamente a lo expuesto en el apartado 3.2.1 se puede considerar una distribución aleatoria de las señales entre los bloques y repetir el análisis múltiples veces, obteniendo el método '*Random RV ICA by Blocks*'.

En la figura 3.4a se aplica este método a sobre 100 señales de mezcla obtenidas a partir de las señales independientes de prueba de '*Set 1*', mientras que en la figura 3.4b se realiza el mismo estudio sobre 100 señales de mezcla obtenidas a partir de las señales independientes de '*Set 2*'. En la figura de la izquierda el coeficiente RV alcanza el valor 1 a partir de 5 CIs, mientras que en la de la derecha ocurre lo mismo a partir de 6 CIs. Este método determina $\text{NOCI} = 5$ para el primer caso y $\text{NOCI} = 6$ para el segundo.

3.4. Método basado en PCA

Este algoritmo considera que NOCI es igual al número de Componentes Principales significativas. Este método es extremadamente sencillo, pero a su vez está basado en hipótesis que no son siempre ciertas [Mur et al., 2017]. Se trata de otro método puramente gráfico en el que se selecciona NOCI a partir de una gráfica en la que se muestran los autovalores de *X* con respecto al índice de su correspondiente Componente Principal. El NOCI estará en el valor del índice de la Componente Principal en el que se

produce un 'codo'. En la figura 3.5 se muestra la aplicación de este método sobre 100 mezclas generadas con una matriz de mezclas aleatoria sobre las señales de prueba *Set 2*. Se puede ver que no es nada sencillo determinar la posición exacta del 'codo', aunque en este caso se sabe que $\text{NOCI} = 6$.

(a) *RV ICA by Blocks* sobre '*Set 1*'.(b) *RV ICA by Blocks* sobre '*Set 2*'.Figura 3.4: Ejemplos de aplicación *RV ICA by blocks*.Figura 3.5: Ejemplo de aplicación del método basado en PCA a 100 señales de mezcla obtenidas a partir de las señales independientes del conjunto *Set2*.

3.5. Algoritmo Linear Correlations between Components (LCC)

Todos los métodos presentados anteriormente para calcular el NOCI se basan en la interpretación de una gráfica con las dificultades que esto implica en la mayoría de los casos. El algoritmo LCC proporciona automáticamente un valor numérico concreto para NOCI de una forma sencilla y rápida [Mur et al., 2017], al tratarse de un algoritmo no supervisado. El algoritmo hace uso de: (i) JADE para obtener los modelos ICA, (ii) la correlación lineal entre componentes, y (iii) el hecho de que aparece una CI descompuesta cuando el Número Componentes Independientes (NCI) utilizado en JADE es superior a NOCI.

Dado un conjunto de n señales mezcladas y un NCI, el algoritmo JADE proporciona un modelo con de NCI componentes. Mientras $\text{NCI} \leq \text{NOCI}$, el algoritmo JADE proporciona componentes absolutamente independientes y por tanto tendrán correlación cero entre ellas. Cuando $\text{NCI} > \text{NOCI}$, el algoritmo JADE es incapaz de proporcionar mas de NOCI componentes independientes, por lo que al menos alguna componente de las proporcionadas por JADE no será completamente independiente del resto de componentes. Desde un punto de vista cuantitativo se tendrá que la correlación lineal entre algunas de las componentes obtenidas por JADE será superior a cero.

Teniendo en cuenta lo anterior, LCC va calculando modelos ICA con JADE desde $\text{NCI} = 1$ e incrementado progresivamente NCI en una unidad y llegando como máximo hasta n . Para cada modelo obtiene NCI componentes y calcula la correlación lineal de todas las componentes entre sí. El algoritmo se detiene en el momento que alguna de las correlaciones entre componentes es suficientemente superior a cero. El NOCI será el mayor NCI para el cual todas las correlaciones entre componentes sea igual a cero, y por tanto todas las componentes sean independientes.

En [Mur et al., 2017] se detalla el algoritmo en los siguientes términos:

Entrada: Un conjunto de n señales independientes en una matriz \mathbf{X} de dimensiones $n \times t$, siendo t el número de muestras de cada una de las señales (duración de las señales).

Salidas: NOCI y las CI.

Para cada modelo de $\text{NCI} > 2$ repetir

1. Obtener las NCI componentes independientes utilizando JADE.
2. Calcular para cada par de componentes independientes (a, b) la $\text{Corr}(a, b)$ con $1 \leq a < b \leq \text{NCI}$, **hasta que para un modelo particular suceda que dos componentes particulares (p, m) ten-**

gan $|Corr(p, m)| > 0,1$ por lo que se estará en el caso de que $NCI = NOCI + 1$

3. Seleccionar NOCI como NCI inmediatamente anterior a la condición para terminar el algoritmo y sus CIs

Notas sobre el algoritmo:

- las letras a y b se refieren a números enteros para identificar un par cualquiera de componentes obtenidos por el modelo ICA;
- las letras p y m se refieren a números enteros para identificar el primer par de componentes detectadas donde sucede que $|Corr(p, m)| > 0,1$;
- el valor absoluto de $Corr(p, m)$ está limitado a 0,1 en lugar de a 0 para evitar que pequeñas cantidades muy próximas a cero puedan parar el algoritmo;
- JADE necesita que $n > NOCI$ para poder obtener el número óptimo de componentes independientes;
- el algoritmo LCC comienza considerando el modelo $NCI = 2$.

Como ejemplo de aplicación del algoritmo LCC consideramos las tres señales independientes de la figura 3.6a, a partir de las cuales se han generado 8 señales de mezcla utilizando la matriz \mathbf{A} que se muestra en 3.3.

$$\mathbf{A} = \begin{pmatrix} 0,8715 & 0,9120 & 0,0580 \\ 0,9899 & 0,2671 & 0,9594 \\ 0,2729 & 0,9751 & 0,6441 \\ 0,0496 & 0,5051 & 0,5921 \\ 0,8817 & 0,9434 & 0,2697 \\ 0,5821 & 0,9545 & 0,6757 \\ 0,4420 & 0,3972 & 0,6368 \\ 0,3786 & 0,6416 & 0,8772 \end{pmatrix} \quad (3.3)$$

En la figura 3.6b se muestra el modelo JADE obtenido para $NCI = 2$ y como se puede ver la correlación entre las dos componentes obtenidas es 0. En la figura 3.6c se muestra el modelo JADE obtenido para $NCI = 3$ y se muestra además que la correlación máxima entre cada par de componentes extraídas sigue siendo cero. En la figura 3.6d se muestra el modelo JADE para $NCI = 4$, sin embargo en este caso la correlación máxima entre dos de las componentes extraídas es 0,8692 por lo que existe dependencia entre esas dos componentes y se produce la condición para detener el algoritmo. Se puede ver en esa figura que el máximo valor absoluto de la correlación se producirá entre las componentes S3 y S4. Se seleccionará $NOCI = 3$ como el mayor valor de NCI antes de que el máximo valor absoluto de la correlación entre las componentes extraídas por JADE sea mayor que 0,1 .

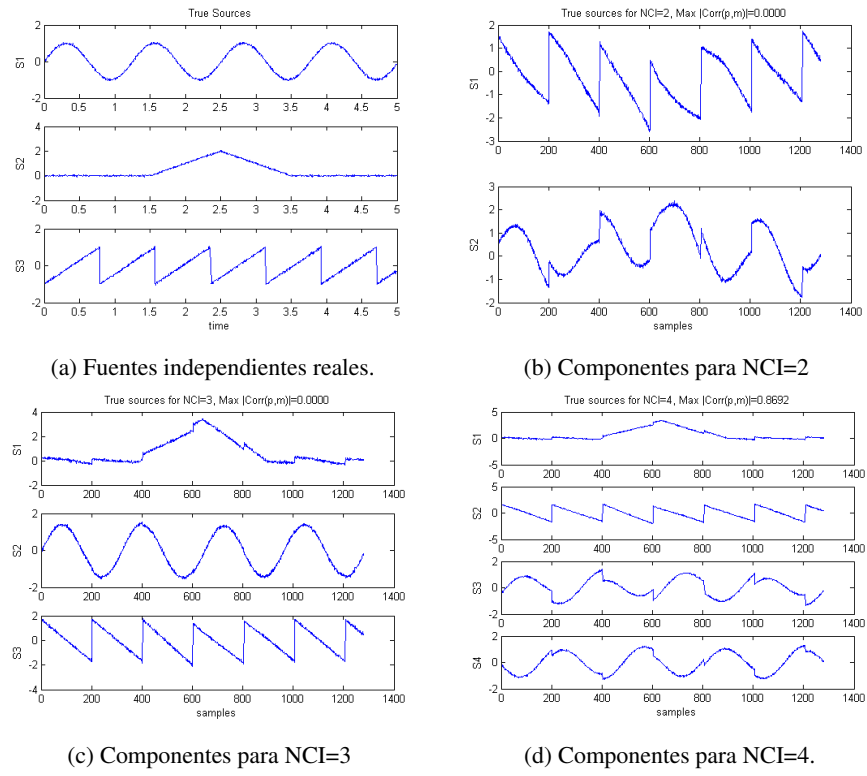


Figura 3.6: Representación algoritmo LCC.

Capítulo 4

Resultados del estudio

4.1. Señales de prueba

Para el estudio de los métodos de cálculo del NOCI presentados en el capítulo 3, se han utilizado dos conjuntos de señales diferentes que ya se introdujeron en dicho capítulo y un conjunto de señales adicional que se presenta más adelante. El conjunto *Set 1* es de producción propia, mientras que el conjunto *Set 2* es el conjunto de señales presentado y utilizado en [Mur et al., 2017] con la única diferencia de la componente aleatoria de cada una de las señales. El conjunto de señales *Set 1* se muestra en la figura 3.1a, y el conjunto de señales *Set 2* se muestra en la figura 3.1b.

Además de los conjuntos anteriores, se ha utilizado un subconjunto de *Set1* llamado *Set1a* que contiene todas las señales de *Set1* excepto la señal de ruido blanco, esto es las señales s_1 , s_2 , s_3 y s_4 , dejando fuera la señal s_5 .

Los conjuntos anteriores están creados a partir de señales generadas sintéticamente, por lo que el estudio quedaría más completo utilizando un conjunto adicional de señales reales. El conjunto adicional utilizado son cuatro señales de audio extraídas de [Jaakko Särelä and Ella Bingham, 2005]. Se trata de cuatro señales de audio con diferentes voces y sonidos de la vida cotidiana. En [Jaakko Särelä and Ella Bingham, 2005] se presenta una demostración para entender el Cocktail-Party Problem (CPP) y cómo ICA es una herramienta que puede resolverlo. Este conjunto de señales se denomina como *Set3* a lo largo de este texto.

En la tabla 4.1 se muestran las principales características de cada uno de los conjuntos de señales independientes de prueba.

4.1.1. Independencia de las señales de prueba

La *Distancia Correlación* ($dCorr$) [Mur et al., 2017] es una medida de la dependencia estadística entre dos variables aleatorias X e Y . Se define la *Distancia Correlación* según 4.1. Esta medida de dependencia es

Conjunto	CI	Tipo	F. Muestreo	Duración
Set1a	4	Sintéticas	256Hz	5 secs.
Set1	5	Sintéticas	256Hz	5 secs
Set2	6	Sintéticas	500Hz	2 secs.
Set3	4	Audio	8KHz	6,25 secs.

Tabla 4.1: Características de las señales de prueba.

cero sí y sólo sí X e Y son independientes. La correlación por sí sola no permite determinar la independencia de dos variables aleatorias, de forma que $Corr(X, Y) = 0$ no implica independencia entre X e Y . En las tablas 4.2, 4.3 y 4.4 se muestra la *Distancia Correlación* entre las señales que componen cada uno de los conjuntos de señales de prueba, lo que permite determinar el grado de independencia de cada una de las señales con respecto a sus compañeras.

$$dCorr^2(X, Y) = \frac{dCov^2(X, Y)}{(dVar^2(X) \cdot dVar^2(Y))^{1/2}} \quad (4.1)$$

$dCorr(s_i, s_j)$	s_2	s_3	s_4	s_5
s_1	0.0558	0.0798	0.0303	0.0468
s_2	-	0.1288	0.0187	0.0416
s_3	-	-	0.0326	0.0487
s_4	-	-	-	0.0326

Tabla 4.2: Distancia correlación de las señales *Set1*.

$dCorr(s_i, s_j)$	s_2	s_3	s_4	s_5	s_6
s_1	0.0767	0.0399	0.0725	0.0791	0.0186
s_2	-	0.0758	0.0971	0.0578	0.0151
s_3	-	-	0.0643	0.0987	0.0101
s_4	-	-	-	0.0188	0
s_5	-	-	-	-	0.0254

Tabla 4.3: Distancia correlación de las señales *Set2*.

Como se puede ver en dichas tablas, las señales s_3 y s_2 del conjunto *Set1* no son totalmente independientes, el resto de señales sí pueden considerarse independientes. En el caso de las señales del conjunto *Set2* y *Set3*, todas las señales son independientes.

Durante todo el estudio se han utilizado matrices de mezcla generadas aleatoriamente con *MATLAB*, con el fin de obtener resultados más generales.

$dCorr(s_i, s_j)$	s_2	s_3	s_4
s_1	0.0442	0.0295	0.0240
s_2	-	0.0419	0.0375
s_3	-	-	0.0407

Tabla 4.4: Distancia correlación de las señales *Set3*.

4.2. Ensayo sobre el conjunto de señales *Set1a*

En este apartado se muestra la aplicación de los métodos presentados en el capítulo 3 para la determinación del NOCI sobre diferentes conjuntos de señales de mezcla generadas a partir de las señales independientes del conjunto *Set1a*.

Se han generado aleatoriamente matrices de mezcla para generar 20, 40, 80, 100, 120, y 140 señales de mezcla u observaciones. En la tabla 4.5 se muestran los resultados obtenidos para la determinación del NOCI con los distintos métodos bajo estudio sobre las señales de mezcla generadas a partir del conjunto de señales de prueba *Set1a*.

Método / Mezclas	20	40	80	100	120	140
DW	?	4	4	4	4	4
ICA by blocks	7	4	4	4	4	4
RV ICA by blocks	4	4	4	4	4	4
PCA	2 a 5	2 a 5	2 a 5	2 a 5	2 a 5	2 a 5
LCC	4	4	4	4	4	4

Tabla 4.5: Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de *Set1a*.

4.2.1. Criterio DW sobre el conjunto de señales *Set1a*

En la figura 4.1 se muestran los resultados obtenidos mediante la aplicación de DW sobre las señales de mezcla generadas a partir del conjunto de señales *Set1a*. Se trata de un subconjunto de las señales presentadas en la figura 3.1a en el que se ha eliminado la señal de ruido blanco. Como se puede apreciar en todos los casos se adivina que el valor de NOCI recomendado por el método es 4, ya que a partir de este valor la matriz de residuos X_r empieza a contener mucho más ruido que información. Sin embargo, en el caso de 20 señales de mezcla que se presenta en la figura 4.1a, resulta mucho más complicado determinar este resultado. El principal problema de DW es que una vez generadas las gráficas comienza la parte de interpretación de las mismas y no en todas las ocasiones es igual de sencillo. A tenor de los resultados presentados en la figura 4.1 se podría deducir además que en términos generales, cuanto mayor es el número de

mezclas disponibles, mayor claridad para interpretar los resultados de DW y determinar el NOCI.

4.2.2. Método 'ICA by blocks' sobre el conjunto de señales Set1a

En la figura 4.2 se muestran los resultados obtenidos durante la aplicación de 'ICA by blocks' sobre las señales de mezcla generadas a partir del conjunto de señales Set1a.

Al igual que ocurría en el caso anterior, de acuerdo al resultado obtenido en la figura 4.2a no se puede determinar de manera precisa el NOCI. Según dicha gráfica quizás 7 sería el valor estimado para este caso concreto. En el resto de gráficas presentadas en la misma figura, se determina claramente que 4 es el NOCI estimado. De nuevo parece claro que cuanto mayor es el número de mezclas disponibles, mayor facilidad para interpretar el resultado de 'ICA by blocks'. En todos los casos excepto el de 20 señales de mezcla, a partir de considerar un máximo de 4 CIs en adelante, la líneas de correlación empiezan a decrecer en el mismo punto, que se corresponde con 4 CIs.

4.2.3. Método 'RV ICA by blocks' sobre el conjunto de señales Set1a

En la figura 4.3 se muestran los resultados obtenidos durante la aplicación de 'RV ICA by blocks' sobre las señales de mezcla generadas a partir del conjunto de señales Set1a.

En todos los casos el NOCI determinado es 4, y en este caso no parece influir demasiado en los resultados el número de mezclas disponibles.

4.2.4. Método 'PCA' sobre el conjunto de señales Set1a

En la figura 4.4 se muestran los resultados obtenidos durante la aplicación del método 'PCA' sobre las señales de mezcla generadas a partir del conjunto de señales Set1a.

En todos los casos el 'codo' se produce entre 2 y 5, pero resulta difícil determinar el valor de NOCI estimado por el método.

4.2.5. Método 'LCC' sobre el conjunto de señales Set1a

En la figura 4.5 se representan los resultados obtenidos durante la aplicación del método 'LCC' sobre las señales de mezcla generadas a partir del conjunto de señales Set1a. En el eje x se representa cada uno de los conjuntos de señales de mezcla y en el eje y se representa el valor de NOCI estimado por el LCC. Como se puede ver clara e inequívocamente en la figura 4.5, el valor de NOCI estimado por este método es 4 en todos los

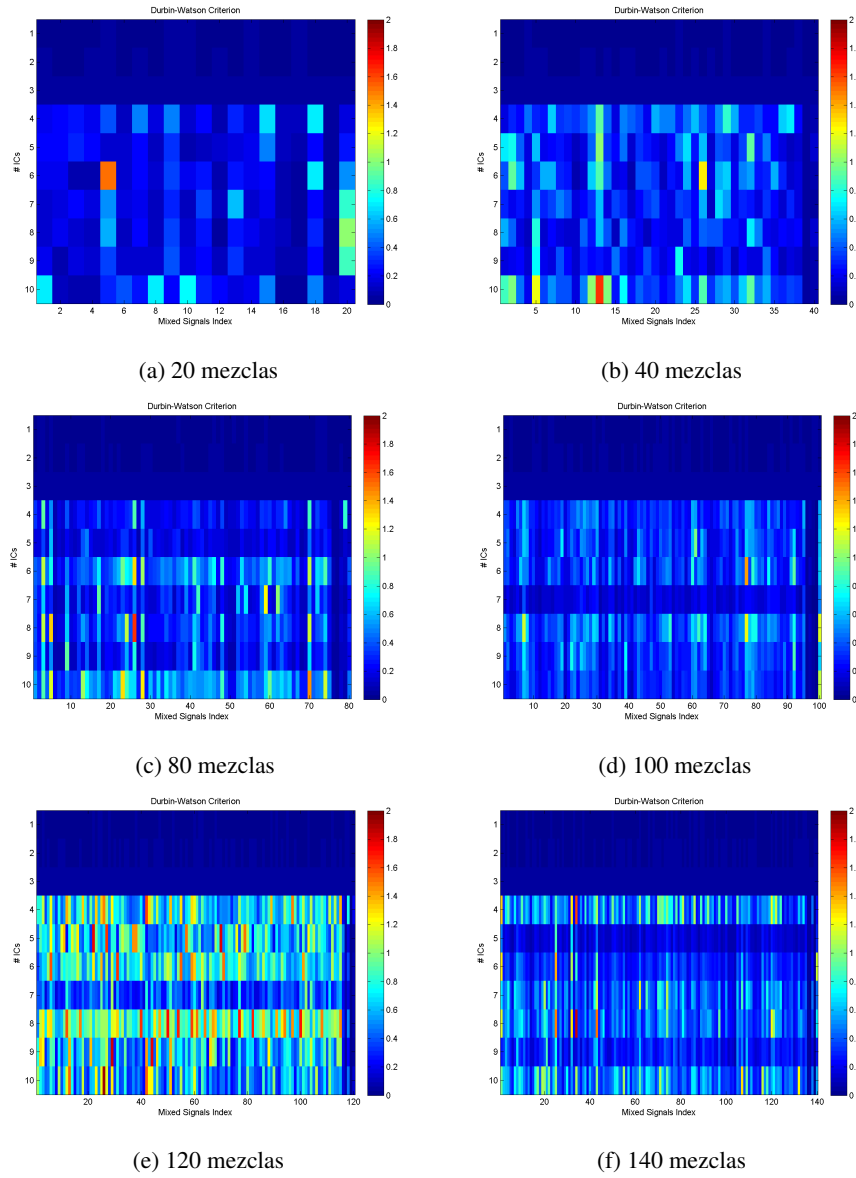
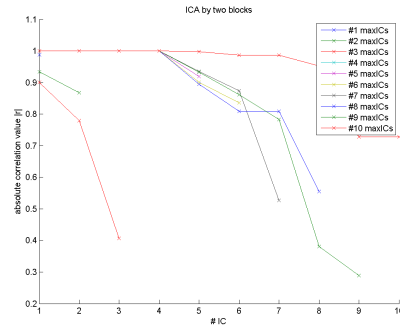
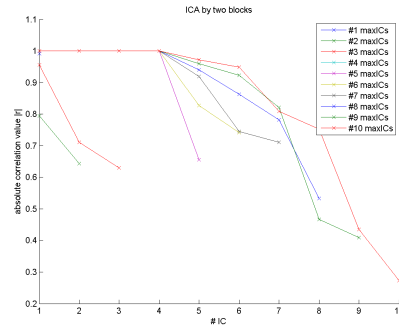


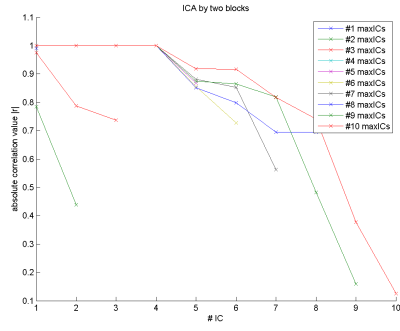
Figura 4.1: Aplicación del criterio DW sobre mezclas de señales *Set1a*.



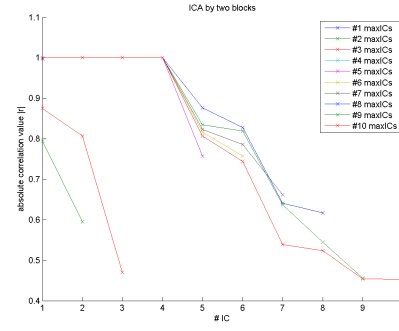
(a) 20 mezclas



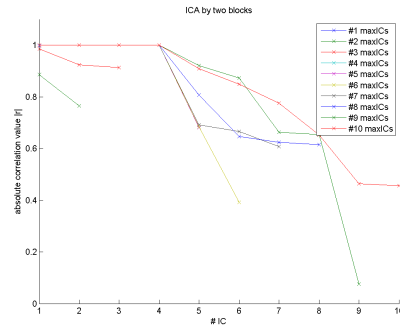
(b) 40 mezclas



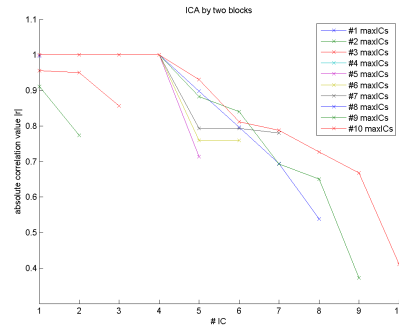
(c) 80 mezclas



(d) 100 mezclas

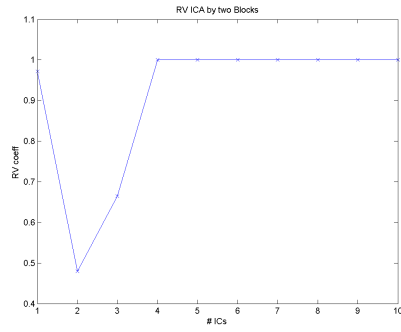


(e) 120 mezclas

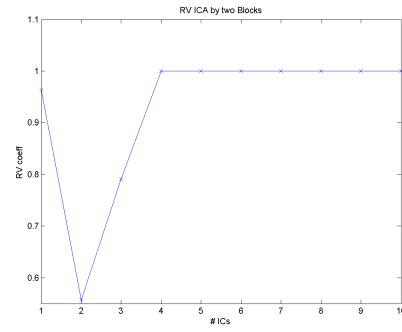


(f) 140 mezclas

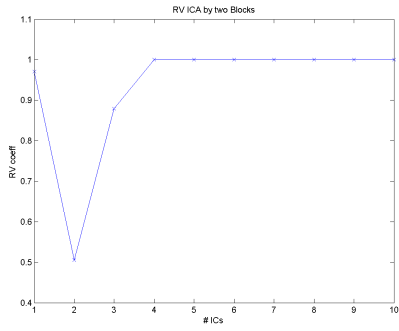
Figura 4.2: Aplicación del método 'ICA by blocks' sobre mezclas de señales *Set1a*.



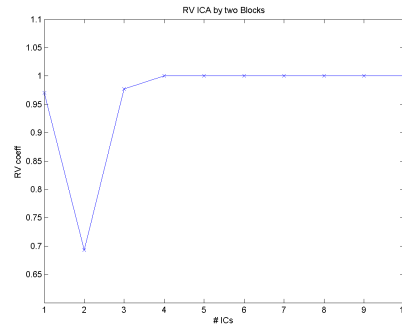
(a) 20 mezclas



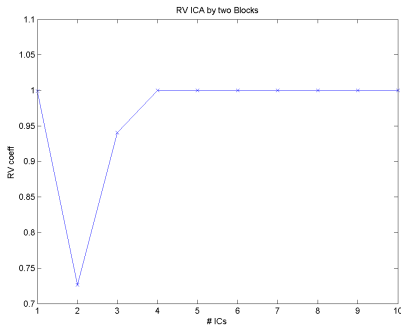
(b) 40 mezclas



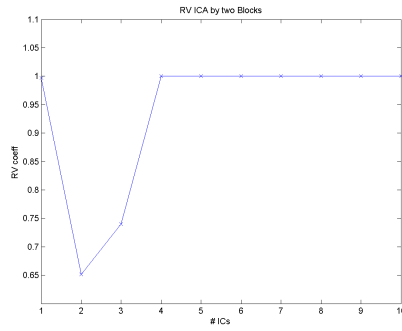
(c) 80 mezclas



(d) 100 mezclas

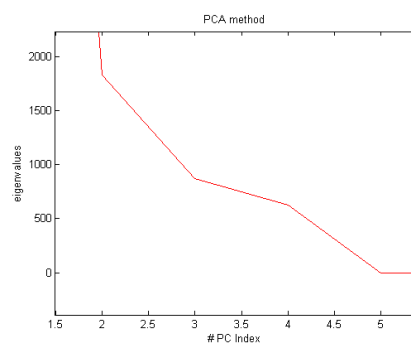


(e) 120 mezclas

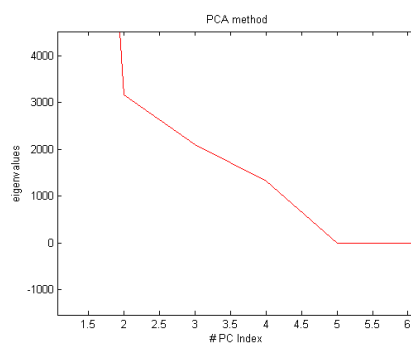


(f) 140 mezclas

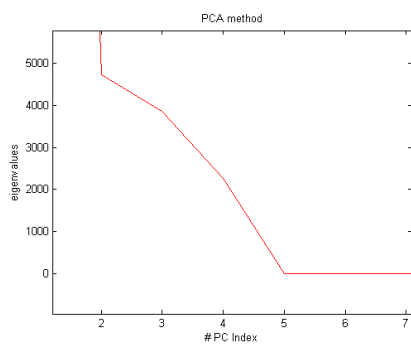
Figura 4.3: Aplicación del método 'RV ICA by blocks' sobre mezclas de señales *Set1a*.



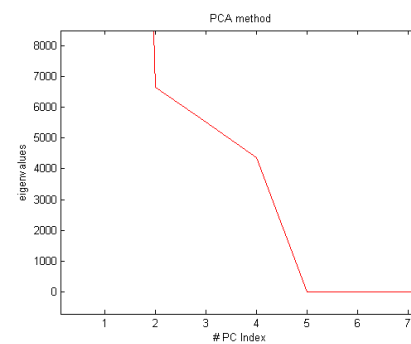
(a) 20 mezclas



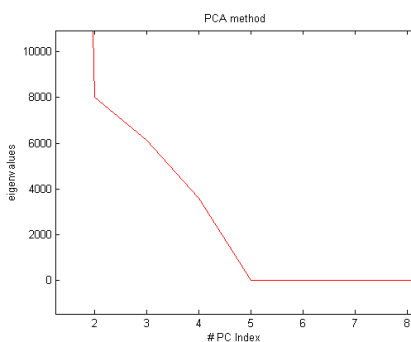
(b) 40 mezclas



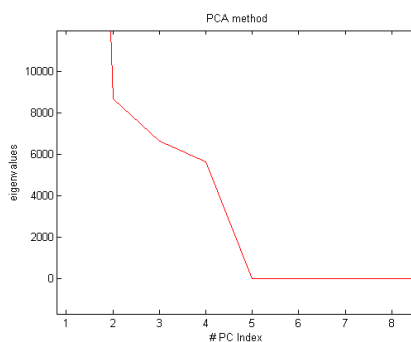
(c) 80 mezclas



(d) 100 mezclas



(e) 120 mezclas



(f) 140 mezclas

Figura 4.4: Aplicación del método 'PCA' sobre mezclas de señales *Set1a*.

casos. Debe recordarse que el método *LCC* proporciona automáticamente y de modo no supervisado un valor numérico para NOCI.

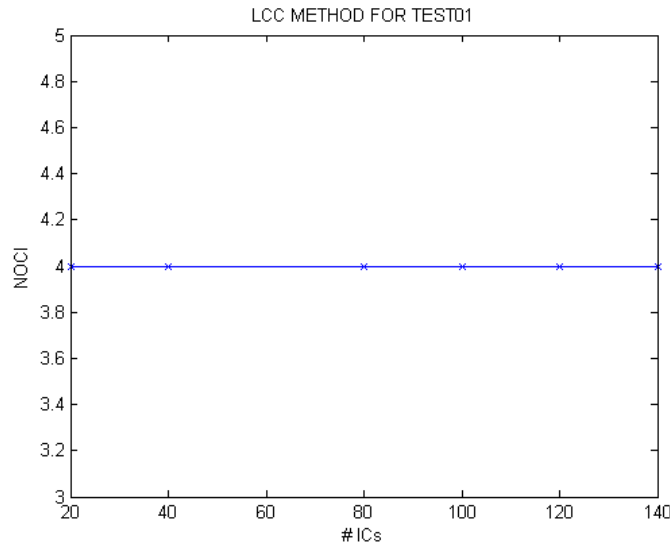


Figura 4.5: Aplicación del método '*LCC*' sobre mezclas de señales *Set1a*.

4.3. Ensayo sobre el conjunto de señales *Set1*

En este apartado se muestra la aplicación de los métodos presentados en el capítulo 3 para la determinación del NOCI sobre diferentes conjuntos de señales de mezcla generadas a partir de las señales independientes del conjunto *Set1*. Se trata del conjunto de 5 señales independientes que se muestran en la figura 3.1a, una de ellas es ruido blanco.

Se han generado aleatoriamente matrices de mezcla para generar 20, 40, 80, 100, 120, y 140 señales de mezcla u observaciones. En la tabla 4.6 se muestran los resultados obtenidos para la determinación del NOCI con los métodos bajo estudio sobre las señales de mezcla generadas a partir del conjunto de señales de prueba *Set1*.

Método / Mezclas	20	40	80	100	120	140
DW	5	5	5	5	5	5
ICA by blocks	6	5	5	5	5	5
RV ICA by blocks	5	5	5	5	5	5
PCA	2 a 6	2 a 6	2 a 6	2 a 6	2 a 6	2 a 6
LCC	5	5	5	5	5	5

Tabla 4.6: Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de *Set1*.

4.3.1. Criterio DW sobre el conjunto de señales Set1

En la figura 4.6 se muestran los resultados obtenidos durante la aplicación de DW sobre las señales de mezcla generadas a partir del conjunto de señales *Set1*. Como se puede apreciar en todos los casos se adivina que el valor de NOCI recomendado por el método es 5, ya que a partir de este valor la matriz de residuos X_r empieza a contener mucho más ruido que información. En esta ocasión incluso en el caso de 20 señales de mezcla que se presenta en la figura 4.6a se puede determinar el NOCI correcto con suficiente claridad. El principal problema de DW es que una vez generadas las gráficas comienza la parte de interpretación de las mismas y no en todas las ocasiones es igual de sencillo. Al igual que en el caso anterior, considerando las gráficas presentadas en la figura 4.6 se podría deducir que, en términos generales, cuanto mayor es el número de mezclas disponibles, mayor claridad para interpretar los resultados de DW y determinar el NOCI. Estas conclusiones son consistentes con las obtenidas en las pruebas realizadas sobre las señales generadas a partir del conjunto *Set1a*.

4.3.2. Método 'ICA by blocks' sobre el conjunto de señales Set1

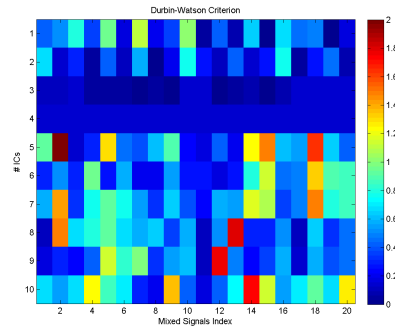
En la figura 4.7 se muestran los resultados obtenidos durante la aplicación de 'ICA by blocks' sobre las señales de mezcla generadas a partir del conjunto de señales *Set1*.

De acuerdo al resultado obtenido en la figura 4.7a el NOCI es 6. En el resto de gráficas presentadas en la misma figura, se determina claramente que 5 es el NOCI estimado. De nuevo parece claro que cuanto mayor es el número de mezclas disponibles, mayor facilidad para interpretar el resultado de 'ICA by blocks'. En todos los casos excepto el de 20 señales de mezcla, a partir de considerar un máximo de 5 CIs en adelante, las líneas de correlación empiezan a decrecer en el mismo punto, que se corresponde con 5 CIs. Sin embargo, la pendiente a partir de 5 CIs es mas abrupta cuanto mayor número de mezclas hay disponibles.

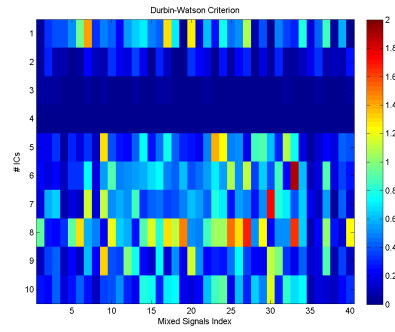
4.3.3. Método 'RV ICA by blocks' sobre el conjunto de señales Set1

En la figura 4.8 se muestran los resultados obtenidos durante la aplicación de 'RV ICA by blocks' sobre las señales de mezcla generadas a partir del conjunto de señales *Set1*.

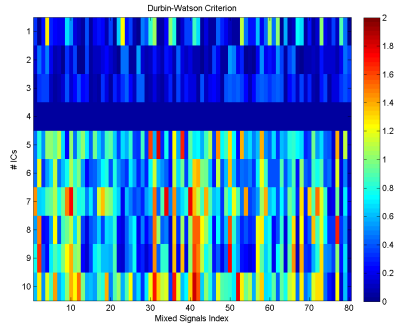
En todos los casos el NOCI determinado es 5, y en este caso no parece influir demasiado en los resultados el número de mezclas disponibles. Sin embargo la interpretación de estas gráficas proporcionan muy poca información de valor añadido.



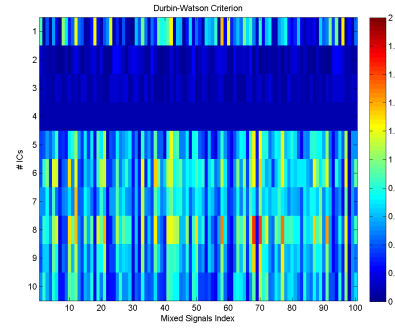
(a) 20 mezclas



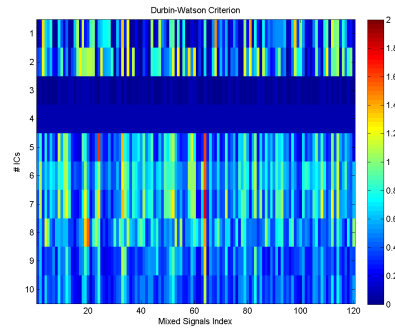
(b) 40 mezclas



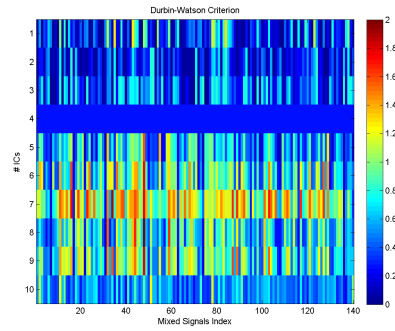
(c) 80 mezclas



(d) 100 mezclas

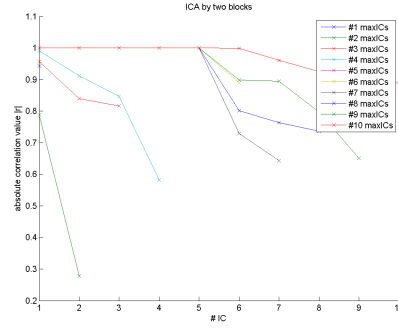


(e) 120 mezclas

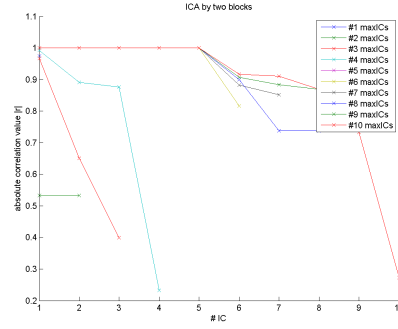


(f) 140 mezclas

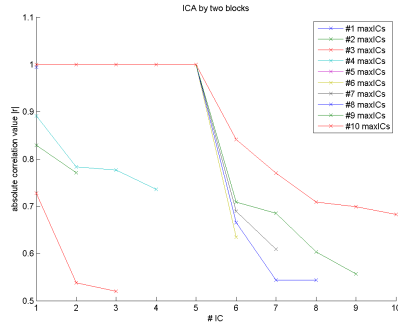
Figura 4.6: Aplicación del criterio DW sobre mezclas de señales *Set1*.



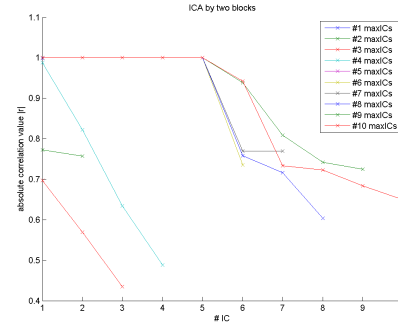
(a) 20 mezclas



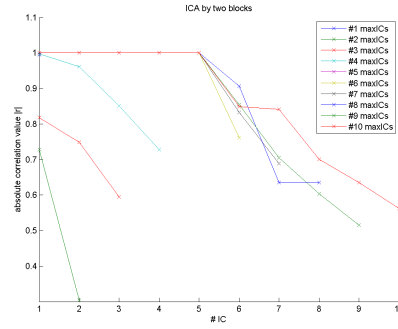
(b) 40 mezclas



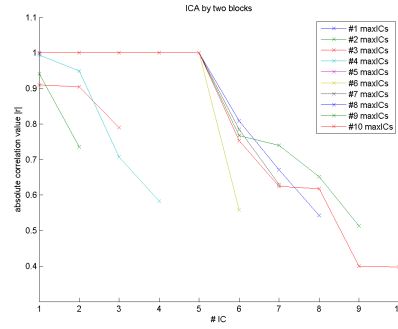
(c) 80 mezclas



(d) 100 mezclas

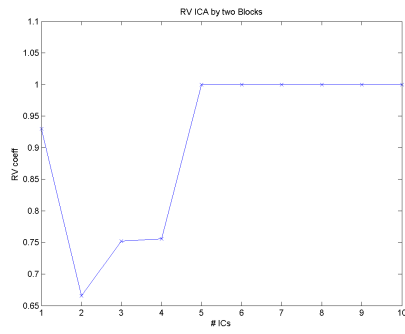


(e) 120 mezclas

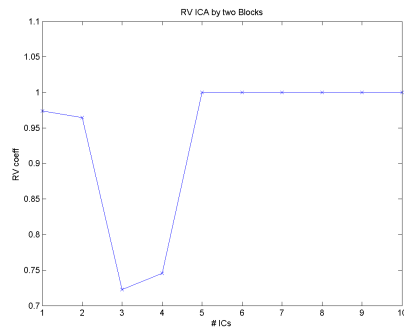


(f) 140 mezclas

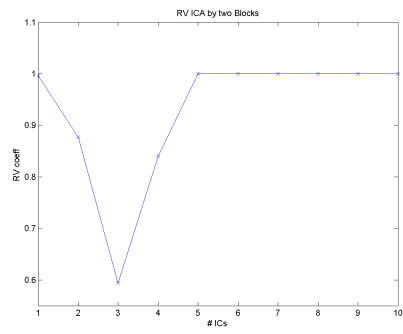
Figura 4.7: Aplicación del método 'ICA by blocks' sobre mezclas de señales *Set1*.



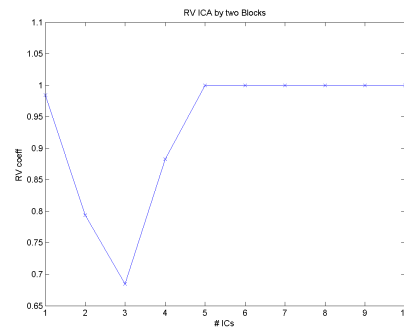
(a) 20 mezclas



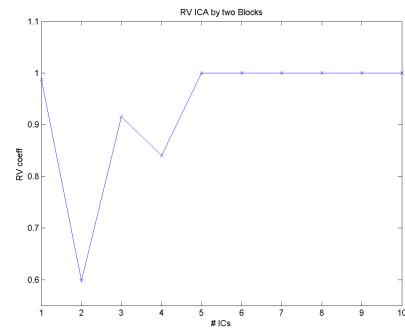
(b) 40 mezclas



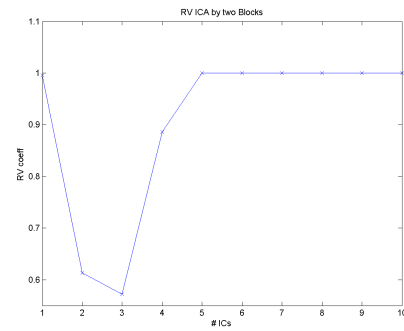
(c) 80 mezclas



(d) 100 mezclas



(e) 120 mezclas



(f) 140 mezclas

Figura 4.8: Aplicación del método 'RV ICA by blocks' sobre mezclas de señales *Set1*.

4.3.4. Método 'PCA' sobre el conjunto de señales Set1

En la figura 4.9 se muestran los resultados obtenidos durante la aplicación del método 'PCA' sobre las señales de mezcla generadas a partir del conjunto de señales *Set1*. En todos los casos el 'codo' se produce entre 2 y 6, pero resulta difícil determinar el valor de NOCI estimado por el método. Este método genera un resultado gráfico difícil de interpretar.

4.3.5. Método 'LCC' sobre el conjunto de señales Set1

En la figura 4.10 se representan los resultados obtenidos durante la aplicación del método 'LCC' sobre las señales de mezcla generadas a partir del conjunto de señales *Set2*. En el eje x se representa cada uno de los conjuntos de señales de mezcla y en el eje y se representa el valor de NOCI estimado por el LCC. Como se puede ver clara e inequívocamente en la figura 4.10, el valor de NOCI estimado por este método es 5 en todos los casos. Debe recordarse que el método LCC proporciona automáticamente y de modo no supervisado un valor numérico para NOCI, sin necesidad de analizar ningún gráfico.

4.4. Ensayo sobre el conjunto de señales Set2

En este apartado se muestra la aplicación de los métodos presentados en el capítulo 3 para la determinación del NOCI sobre diferentes conjuntos de señales de mezcla generadas a partir de las señales independientes del conjunto *Set2*. Se trata de un conjunto de 6 señales independientes, siendo las mismas señales que se utilizaron en [Mur et al., 2017].

Se han generado aleatoriamente matrices de mezcla para generar 20, 40, 80, 100, 120, y 140 señales de mezcla u observaciones. En la tabla 4.7 se muestran los resultados obtenidos para la determinación del NOCI con los métodos bajo estudio sobre las señales de mezcla generadas a partir del conjunto de señales de prueba *Set2*.

Método / Mezclas	20	40	80	100	120	140
DW	6	6	6	6	6	6
ICA by blocks	6	6	6	6	6	6
RV ICA by blocks	6	6	6	6	6	6
PCA	2 a 6	2 a 6	2 a 6	2 a 6	2 a 6	2 a 6
LCC	6	6	6	6	6	6

Tabla 4.7: Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de *Set2*.

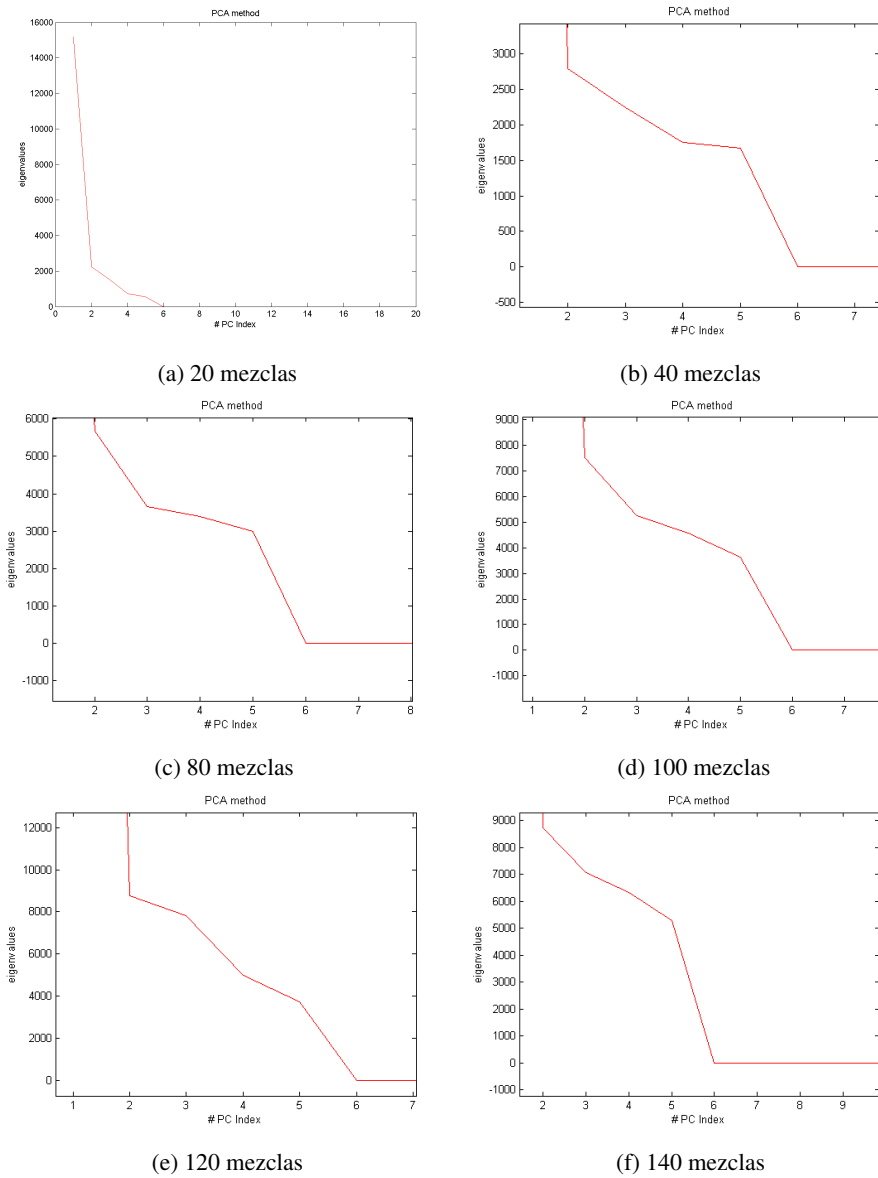


Figura 4.9: Aplicación del método 'PCA' sobre mezclas de señales *Set1*.

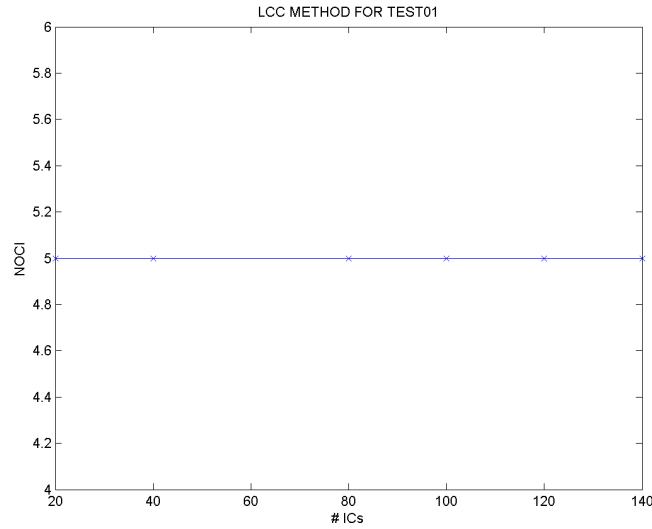


Figura 4.10: Aplicación del método 'LCC' sobre mezclas de señales *Set1*.

4.4.1. Criterio DW sobre el conjunto de señales *Set2*

En la figura 4.11 se muestran los resultados obtenidos durante la aplicación de DW sobre las señales de mezcla generadas a partir del conjunto de señales *Set2*. Como se puede apreciar en todos los casos se adivina que el valor de NOCI recomendado por el método es 6, ya que a partir de este valor la matriz de residuos X_r empieza a contener mucho más ruido que información. En esta ocasión en el caso de 20 señales de mezcla que se presenta en la figura 4.11a resulta más complicado determinar el NOCI correcto. Al igual que en los casos anteriores, considerando las gráficas presentadas en la figura 4.11 se podría deducir que, en términos generales, cuanto mayor es el número de mezclas disponibles, mayor facilidad para interpretar los resultados de DW y determinar el NOCI. Sin embargo con estas señales las gráficas obtenidas no son tan claramente interpretables como en los casos estudiados anteriormente, esto se debe probablemente a las características particulares de las señales del conjunto *Set2*.

4.4.2. Método 'ICA by blocks' sobre el conjunto de señales *Set2*

En la figura 4.12 se muestran los resultados obtenidos durante la aplicación de 'ICA by blocks' sobre las señales de mezcla generadas a partir del conjunto de señales *Set2*.

En todas las gráficas de la figura 4.12 se determina claramente que 6 es el NOCI estimado. De nuevo parece claro que cuanto mayor es el número de mezclas disponibles, mayor facilidad para interpretar el resultado de

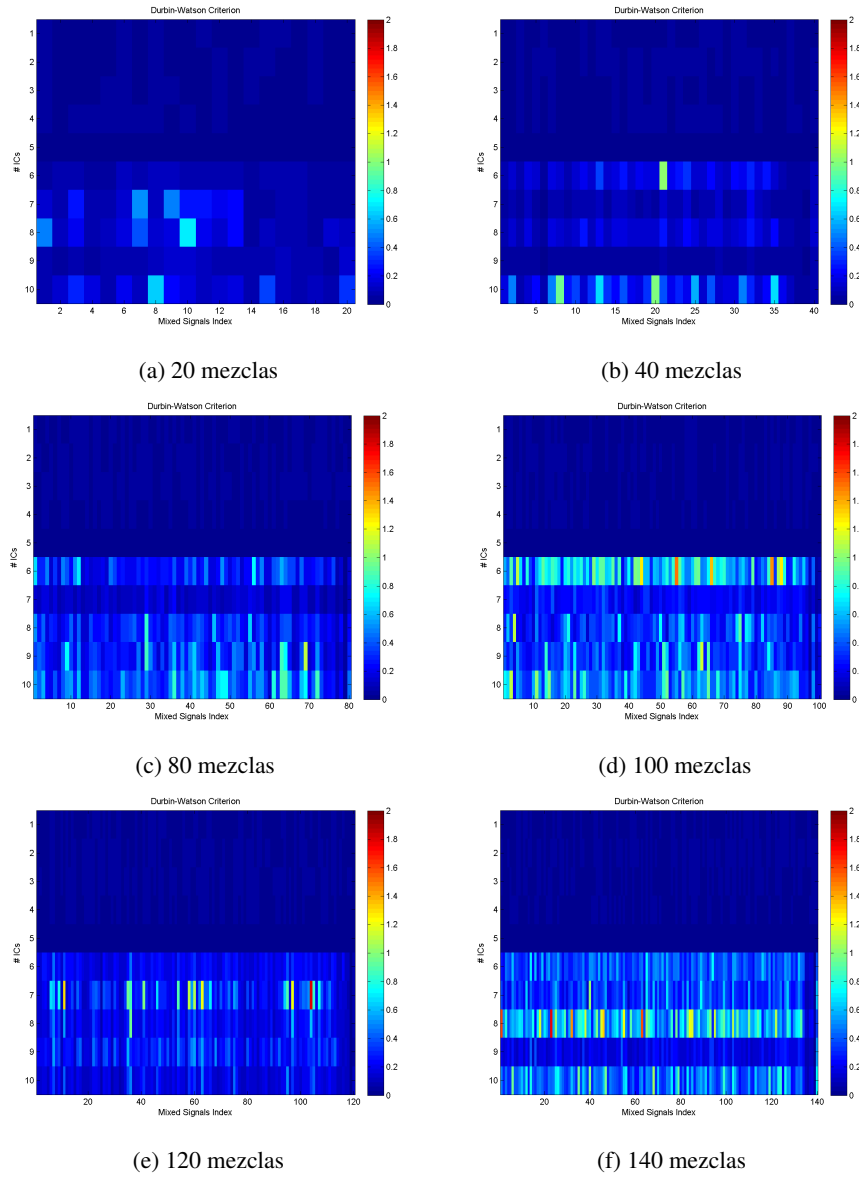


Figura 4.11: Aplicación del criterio DW sobre mezclas de señales *Set2*.

'ICA by blocks'. En todos los casos a partir de considerar un máximo de 6 CIs en adelante, las líneas de correlación empiezan a decrecer en el mismo punto, que se corresponde con 6 CIs. En esta ocasión se puede determinar correctamente el NOCI a partir de las gráficas generadas en todos los casos.

4.4.3. Método 'RV ICA by blocks' sobre el conjunto de señales Set2

En la figura 4.13 se muestran los resultados obtenidos durante la aplicación de 'RV ICA by blocks' sobre las señales de mezcla generadas a partir del conjunto de señales Set2.

En todos los casos el NOCI determinado es 6, y en este caso no parece influir demasiado en los resultados el número de mezclas disponibles. Sin embargo este método parece necesitar bastante mayor tiempo de ejecución que los métodos anteriores. Esta circunstancia se puede constatar mediante las pruebas de tiempo de ejecución presentadas en el apartado 4.6.

4.4.4. Método 'PCA' sobre el conjunto de señales Set2

En la figura 4.14 se muestran los resultados obtenidos durante la aplicación del método 'PCA' sobre las señales de mezcla generadas a partir del conjunto de señales Set2.

En todos los casos el 'codo' se produce entre 2 y 7, pero resulta difícil determinar el valor de NOCI estimado por el método. Este método genera un resultado gráfico difícil de interpretar.

4.4.5. Método 'LCC' sobre el conjunto de señales Set2

En la figura 4.15 se representan los resultados obtenidos durante la aplicación del método 'LCC' sobre las señales de mezcla generadas a partir del conjunto de señales Set2. En el eje x se representa cada uno de los conjuntos de señales de mezcla y en el eje y se representa el valor de NOCI estimado por el LCC. Como se puede ver clara e inequívocamente en la figura 4.15, el valor de NOCI estimado por este método es 6 en todos los casos. Debe recordarse que el método LCC proporciona automáticamente y de modo no supervisado un valor numérico para NOCI.

4.5. Ensayo sobre el conjunto de señales Set3

En este apartado se muestra la aplicación de los métodos presentados en el capítulo 3 para la determinación del NOCI sobre diferentes conjuntos de señales de mezcla generadas a partir de las señales independientes del conjunto Set3. Los conjuntos de señales independientes anteriores son conjuntos de señales generadas sintéticamente mientras que en este caso se

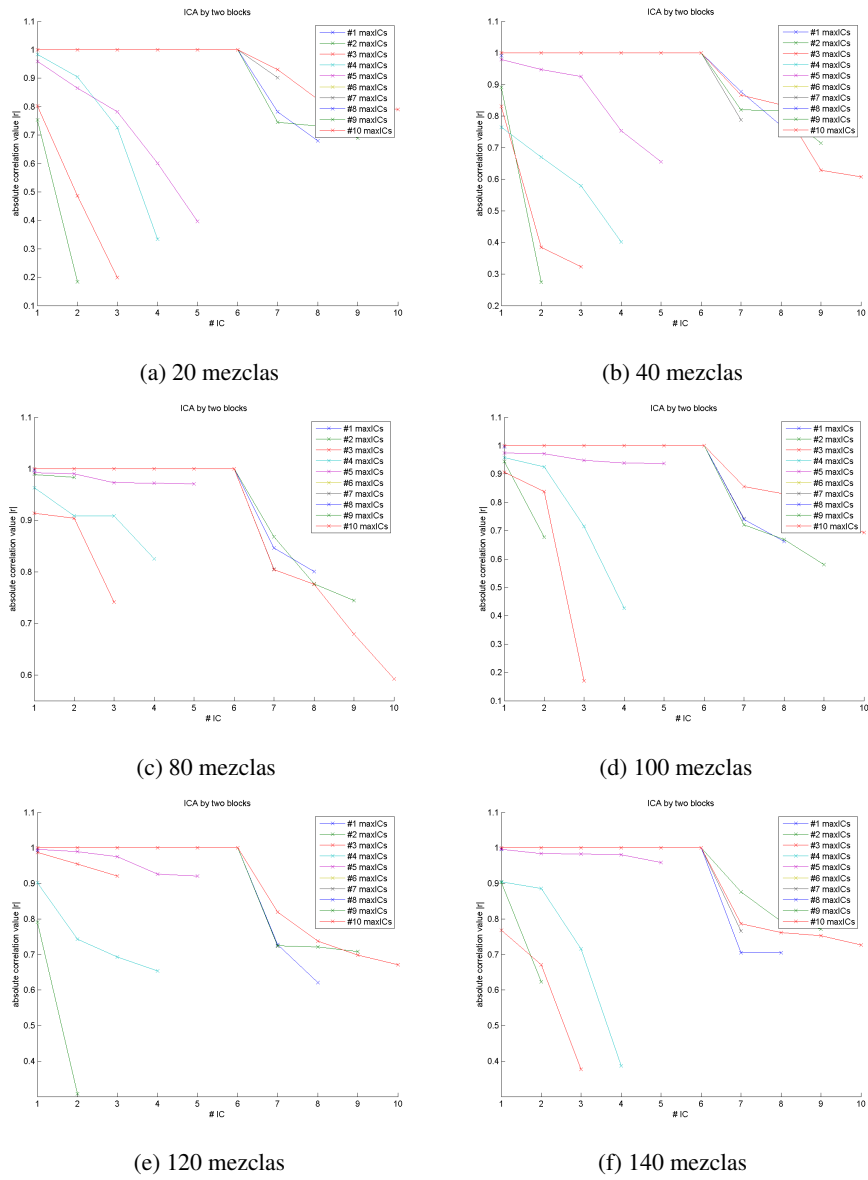
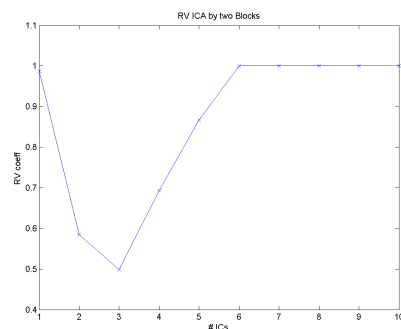
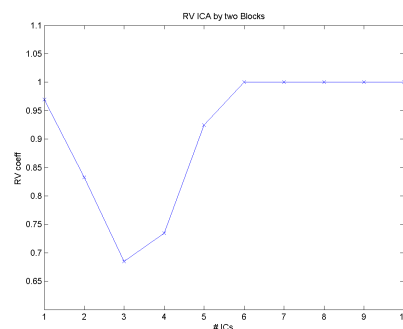


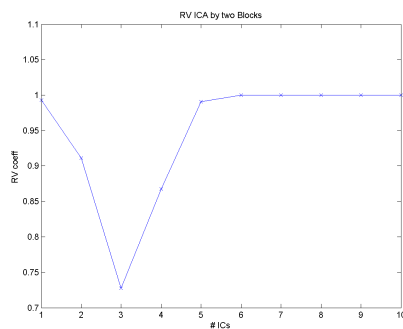
Figura 4.12: Aplicación del método 'ICA by blocks' sobre mezclas de señales Set2.



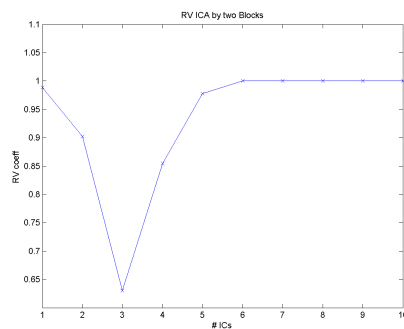
(a) 20 mezclas



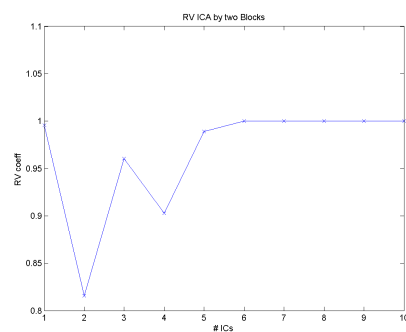
(b) 40 mezclas



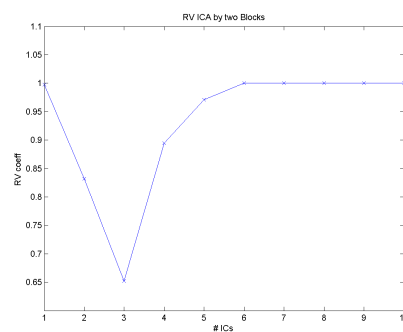
(c) 80 mezclas



(d) 100 mezclas

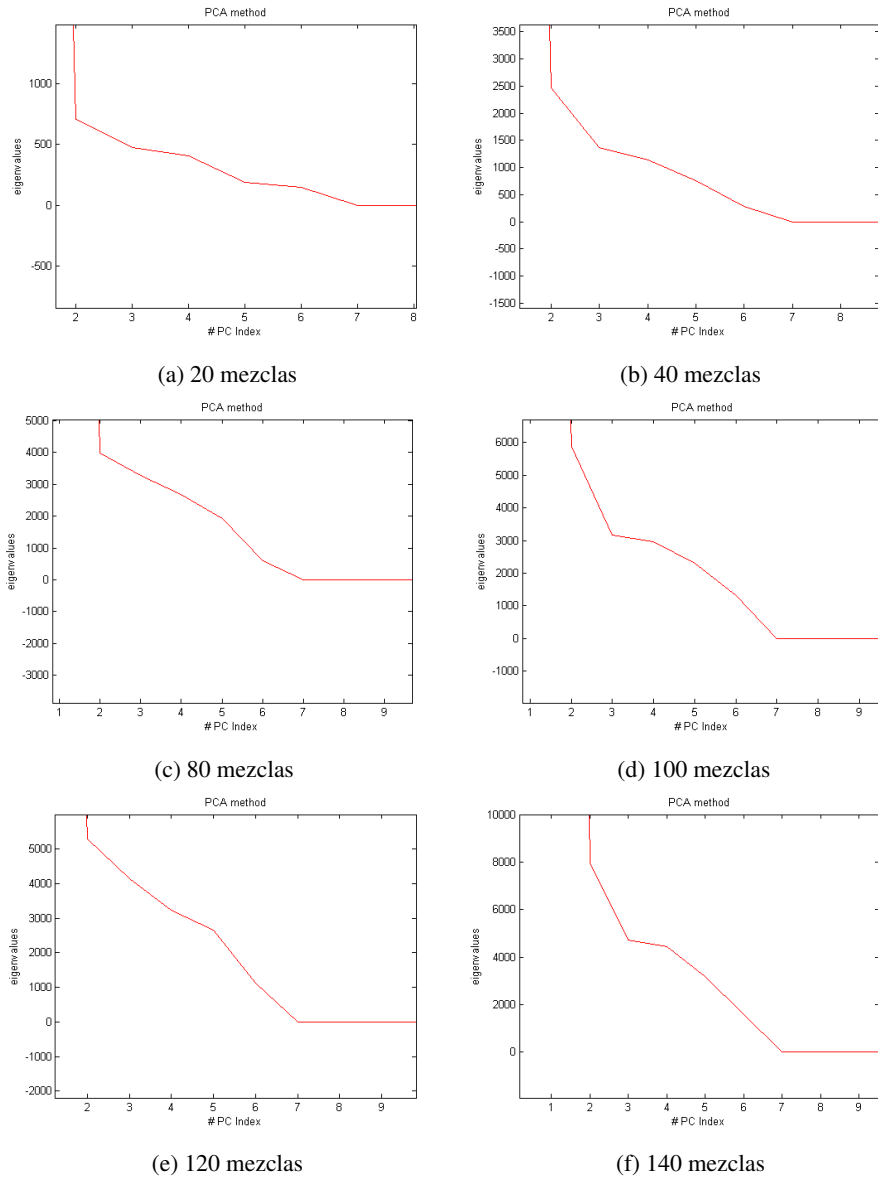


(e) 120 mezclas



(f) 140 mezclas

Figura 4.13: Aplicación del método 'RV ICA by blocks' sobre mezclas de señales *Set2*.

Figura 4.14: Aplicación del método 'PCA' sobre mezclas de señales *Set2*.

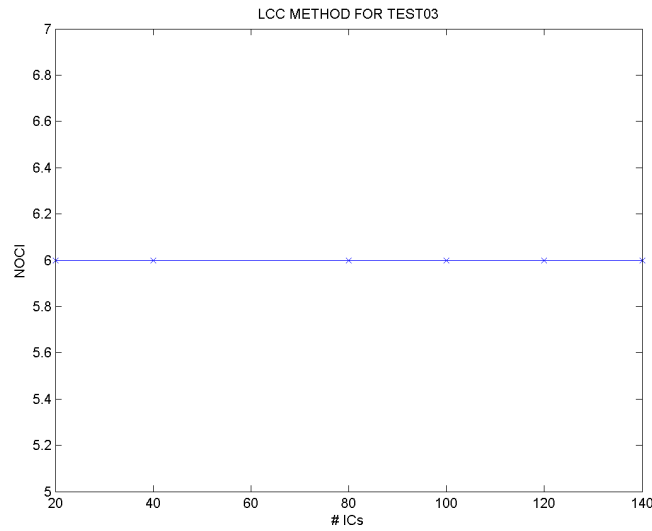


Figura 4.15: Aplicación del método 'LCC' sobre mezclas de señales *Set2*.

trata de aplicar los métodos de forma análoga a lo realizado en las pruebas anteriores pero con 4 señales reales de audio obtenidas de [Jaakko Särelä and Ella Bingham, 2005]. Se trata de señales de 6,25 segundos de duración y una frecuencia de muestreo de 8KHz, por lo que la longitud de los datos es considerablemente mayor que en las pruebas anteriores.

Debido a la longitud de los datos, dos de los métodos bajo estudio no han podido obtener resultados por alcanzar el error *Out of Memory* en *MATLAB*. Se trata del método '*RV ICA by blocks*' y el método basado en '*PCA*'. En el caso del primero, el cálculo del coeficiente RV de acuerdo a 3.2 requiere la operación con matrices demasiado grandes, lo que hace que este método sea muy lento. En el caso del segundo, el error se produce al intentar obtener los autovalores de la matriz de señales de mezcla \mathbf{X} .

Se utilizarán las mismas matrices de mezcla que se utilizaron para el análisis de las señales *Set1*. En la tabla 4.8 se muestran los resultados obtenidos para la determinación del NOCI con los métodos bajo estudio sobre las señales de mezcla generadas a partir del conjunto de señales de prueba *Set3*.

Método / Mezclas	20	40	80	100	120	140
DW	4	4	4	4	4	4
ICA by blocks	-	4	4	4	4	4
RV ICA by blocks	-	-	-	-	-	-
PCA	-	-	-	-	-	-
LCC	4	4	4	4	4	4

Tabla 4.8: Tabla resumen de resultados obtenidos sobre señales de mezcla generadas a partir de *Set3*.

4.5.1. Criterio DW sobre el conjunto de señales Set3

En la figura 4.16 se muestran los resultados obtenidos durante la aplicación de DW sobre las señales de mezcla generadas a partir del conjunto de señales *Set3*. Como se puede apreciar en todos los casos se adivina que el valor de NOCI recomendado por el método es 4, ya que a partir de este valor la matriz de residuos X_r empieza a contener mucho más ruido que información. Al igual que en el caso anterior, considerando las gráficas presentadas en la figura 4.11 se podría deducir que, en términos generales, cuanto mayor es el número de mezclas disponibles, mayor claridad para interpretar los resultados de DW y determinar el NOCI. Además, a pesar de haber utilizado las mismas matrices de mezcla que con las señales de *Set1*, en este caso todas las gráficas proporcionadas por este método son fácilmente interpretables. Esto es debido a las características de las propias señales, como su independencia o incluso la duración de las mismas.

4.5.2. Método 'ICA by blocks' sobre el conjunto de señales Set3

En la figura 4.17 se muestran los resultados obtenidos durante la aplicación de 'ICA by blocks' sobre las señales de mezcla generadas a partir del conjunto de señales *Set3*.

En todas las gráficas de la figura 4.12 se determina claramente que 4 es el NOCI estimado, excepto en el caso de 20 señales de mezcla donde es mucho más complicado llegar a este resultado sin un conocimiento a priori del NOCI. De nuevo parece claro que cuanto mayor es el número de mezclas disponibles, mayor facilidad para interpretar el resultado de 'ICA by blocks' porque mayor es la pendiente de caída de las líneas de correlación a partir de 4 CIs.

4.5.3. Método 'LCC' sobre el conjunto de señales Set3

En la figura 4.18 se representan los resultados obtenidos durante la aplicación del método 'LCC' sobre las señales de mezcla generadas a partir

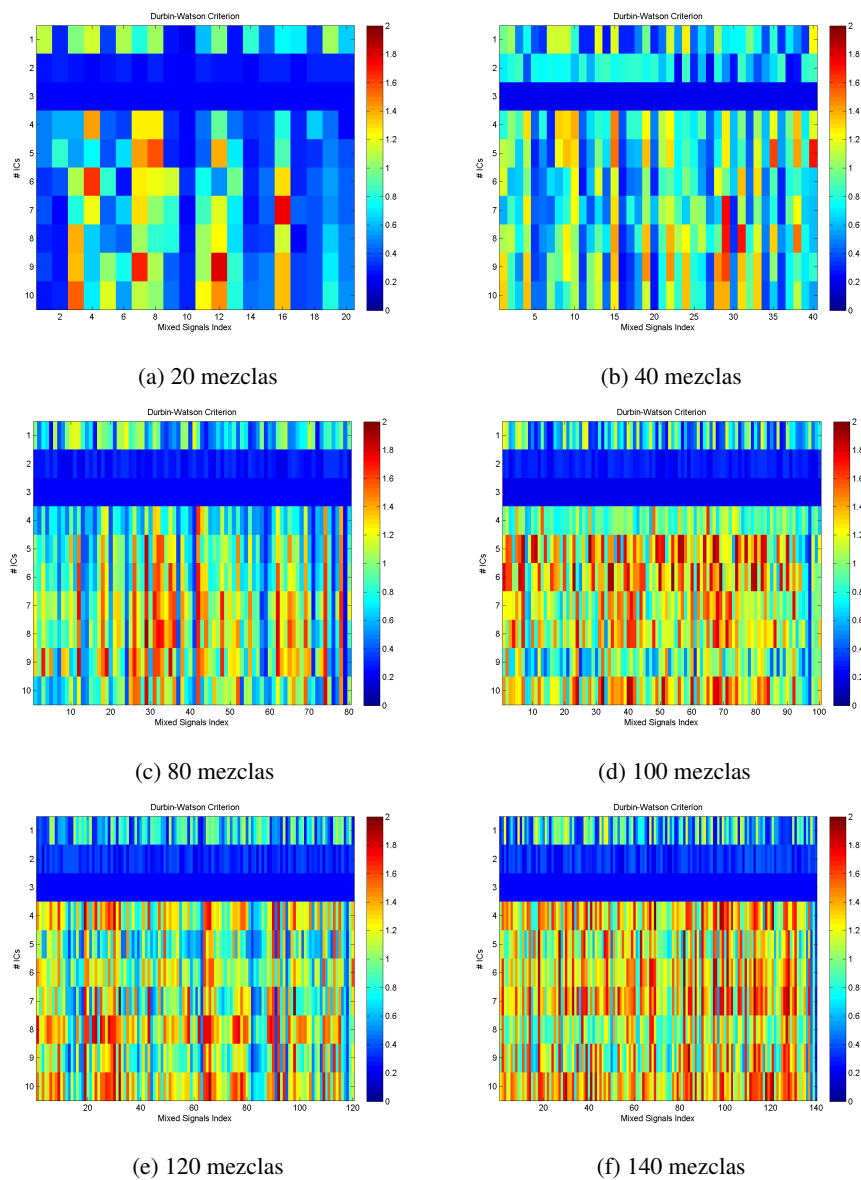


Figura 4.16: Aplicación del criterio DW sobre mezclas de señales *Set3*.

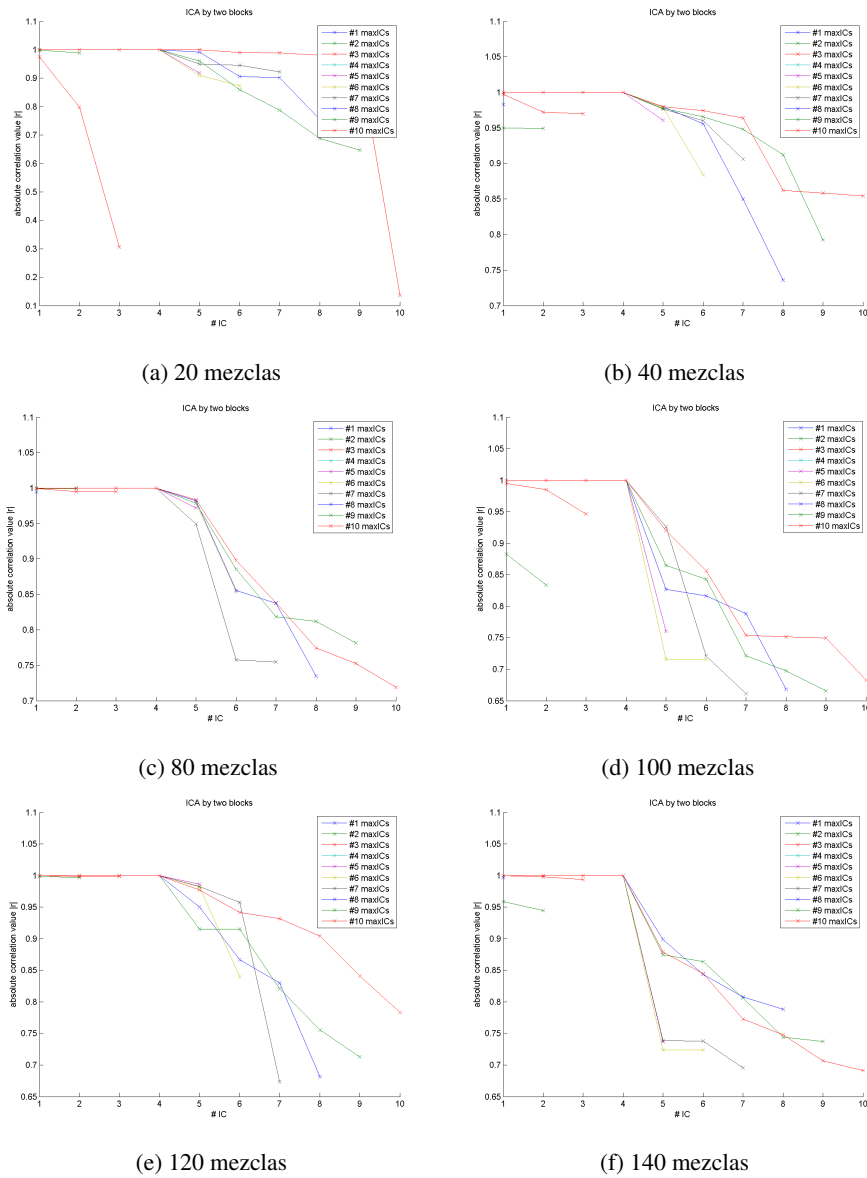


Figura 4.17: Aplicación del método 'ICA by blocks' sobre mezclas de señales Set3.

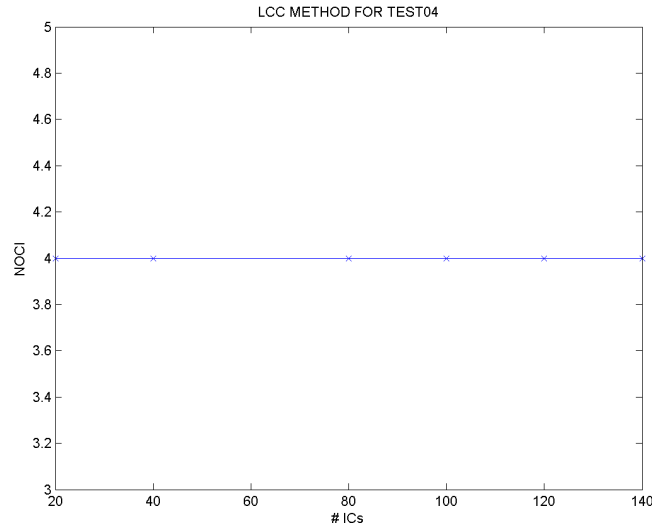


Figura 4.18: Aplicación del método 'LCC' sobre mezclas de señales *Set3*.

del conjunto de señales *Set3*. En el eje x se representa cada uno de los conjuntos de señales de mezcla y en el eje y se representa el valor de NOCI estimado por el LCC. Como se puede ver clara e inequívocamente en la figura 4.18, el valor de NOCI estimado por este método es 4 en todos los casos. En todos los conjuntos de datos estudiados el método LCC se presenta como un método robusto que obtiene buenos resultados.

4.6. Análisis de tiempo de ejecución

En este apartado se presenta un análisis del tiempo de ejecución de cada uno de los métodos presentados en el capítulo 3. El estudio se ha realizado sobre diferentes conjuntos de señales de mezcla generadas a partir de las señales independientes del conjunto *Set1*. Se trata de un conjunto de 5 señales independientes (ver figura 3.1a), una de ellas es ruido blanco. Se han generado aleatoriamente matrices de mezcla para producir diversos números de señales de mezcla/observaciones m y considerando diferentes números máximos de CIs a estudiar $maxICs$. La medición de tiempos se ha realizado durante 50 iteraciones en las que se han generado diferentes matrices de mezcla aleatoriamente. Se presentan los datos de tiempo en forma de tiempos medios de ejecución de los diferentes métodos durante las 50 iteraciones para cada uno de los algoritmos y en los diferentes casos estudiados. No se han considerado los tiempos empleados en representaciones gráficas de resultados e interpretación de los datos.

En la tabla 4.9 se muestran los tiempos medios de ejecución del método

de DW para cada uno de los casos estudiados. Como se puede apreciar, en el caso de DW, el tiempo de ejecución no es excesivamente alto, sin embargo crece exponencialmente conforme aumenta $maxICs$ debido a que se deben realizar más análisis ICA y también aumenta, aunque no tanto, conforme aumenta el número de señales de mezcla.

[secs.]	$m = 14$	$m = 20$	$m = 100$	$m = 200$
$maxICs = 6$	0,0258	0,0278	0,0458	0,0918
$maxICs = 10$	0,2164	0,1823	0,2137	0,3054
$maxICs = 20$	-	6,3832	5,7151	5,9199

Tabla 4.9: Tiempos de ejecución con el algoritmo DW en segundos.

En la tabla 4.10 se muestran los tiempos medios de ejecución del algoritmo 'ICA by blocks' para cada uno de los casos bajo estudio. Se puede observar que el comportamiento es similar al caso anterior y los tiempos son razonablemente parecidos, ambos algoritmos son equiparables en cuanto a tiempos de ejecución.

[secs.]	$m = 14$	$m = 20$	$m = 100$	$m = 200$
$maxICs = 6$	0,0536	0,0528	0,0600	0,0756
$maxICs = 10$	-	0,5589	0,3889	0,5923
$maxICs = 20$	-	-	11,2474	11,3106

Tabla 4.10: Tiempos de ejecución con el algoritmo 'ICA by blocks' en segundos.

En la tabla 4.11 se muestran los tiempos medios de ejecución del algoritmo 'RV ICA by blocks' para cada uno de los casos bajo estudio. Se puede observar que el comportamiento es similar a los casos anteriores pero los tiempos de ejecución son considerablemente mayores. Esta diferencia se debe fundamentalmente al tiempo empleado en calcular los coeficientes RV de correlación.

[secs.]	$m = 14$	$m = 20$	$m = 100$	$m = 200$
$maxICs = 6$	1,8466	1,8628	1,8501	1,9736
$maxICs = 10$	-	5,0117	4,2893	4,5673
$maxICs = 20$	-	-	19,3752	19,6602

Tabla 4.11: Tiempos de ejecución con el algoritmo 'RV ICA by blocks' en segundos.

En la tabla 4.12 se muestran los tiempos medios de ejecución del algoritmo basado en PCA para cada uno de los casos bajo estudio. Este algoritmo vuelve a ser un algoritmo relativamente rápido, equiparable en velocidad a los dos primeros algoritmos estudiados.

[secs.]	$m = 14$	$m = 20$	$m = 100$	$m = 200$
$maxICs = 6$	0,0521	0,0686	0,2899	0,0771
$maxICs = 10$	0,0479	0,0724	0,2998	0,0867
$maxICs = 20$	-	0,0660	0,2912	0,0771

Tabla 4.12: Tiempos de ejecución con el algoritmo 'PCA' en segundos.

Por último, en la tabla 4.13 se muestran los tiempos medios de ejecución del algoritmo *LCC* para cada uno de los casos bajo estudio. Este algoritmo es el más rápido de todos en términos generales. Además se comprueba que el tiempo de ejecución para nada depende de $maxICs$, de hecho es un parámetro que no necesita el algoritmo ya que irá buscando el NOCI secuencialmente y detendrá su ejecución en cuanto detecte el valor óptimo.

[secs.]	$m = 14$	$m = 20$	$m = 100$	$m = 200$
$maxICs = 6$	0,0228	0,0226	0,0293	0,0623
$maxICs = 10$	0,0244	0,0202	0,0299	0,0674
$maxICs = 20$	-	0,0261	0,0315	0,0648

Tabla 4.13: Tiempos de ejecución con el algoritmo 'LCC' en segundos.

Capítulo 5

Conclusiones

5.1. Características de los métodos

En este apartado se presentan las características principales de cada uno de los métodos, extraídas de la experimentación realizada aplicándolos sobre los diferentes conjuntos de señales de prueba.

5.1.1. Características principales del método DW

En general este método ha presentado resultados satisfactorios con los conjuntos de señales que se han estudiado, las características a destacar de este método son:

- El método proporciona un resultado gráfico que debe posteriormente interpretarse, no resultando siempre una tarea sencilla.
- Necesita proporcionar a priori un número máximo de CIs a considerar *maxICs*.
- Cuanto mayor sea *maxICs* mayor será el tiempo necesario para obtener la gráfica de resultados, ya que se deberán realizar *maxICs* análisis JADE.
- Cuanto mayor es el número de señales de mezcla disponibles respecto al número de CIs, en términos generales es más sencillo interpretar los resultados.

5.1.2. Características principales del método 'ICA by blocks'

En general este método también ha presentado resultados satisfactorios sobre los conjuntos de señales de prueba bajo estudio. Las características principales de este método son:

- El método proporciona un resultado gráfico que debe posteriormente interpretarse.
- Necesita proporcionar a priori un número máximo de CIs a considerar $maxICs$.
- Cuanto mayor sea $maxICs$ mayor será el tiempo necesario para obtener la gráfica de resultados, ya que se deberán realizar $b \cdot maxICs$ análisis JADE, siendo b el número de bloques considerado.
- Cuanto mayor es el número de señales de mezcla disponibles respecto al número de CIs, en términos generales es más sencillo interpretar los resultados.
- Se necesitan como mínimo $b \cdot maxICs$ señales de mezcla para poder aplicar el método.
- El número mínimo de bloques a considerar es $b = 2$.

5.1.3. Características principales del método 'RV ICA by blocks'

Las principales características de este método son:

- El método proporciona un resultado gráfico que debe posteriormente interpretarse.
- Necesita proporcionar a priori un número máximo de CIs a considerar $maxICs$.
- Cuanto mayor sea $maxICs$ mayor será el tiempo necesario para obtener la gráfica de resultados, ya que se deberán realizar $b \cdot maxICs$ análisis JADE, siendo b el número de bloques considerado.
- Cuanto mayor es el número de señales de mezcla disponibles respecto al número de CIs, en términos generales es más sencillo interpretar los resultados.
- Se necesitan como mínimo $b \cdot maxICs$ señales de mezcla para poder aplicar el método.
- El número mínimo de bloques a considerar es $b = 2$.
- Es el método con más consumo de tiempo de todos los estudiados, debido al tiempo empleado en calcular el coeficiente RV.
- Con conjuntos de señales de mayor duración o mayor número de señales/mezclas se han producido problemas de 'Out of Memory' que han impedido obtener resultados con algunos de los conjuntos de datos de prueba.

5.1.4. Características principales del método basado en 'PCA'

Las principales características que se han observado en este método son:

- El método proporciona un resultado gráfico que debe posteriormente interpretarse.
- En general las gráficas proporcionadas no son fácilmente interpretables porque se producen diversos 'codos' en los autovalores.
- No es necesario proporcionar a priori un número máximo de CIs a considerar.
- Con conjuntos de señales de mayor duración o mayor número de señales/mezclas se han producido problemas de '*Out of Memory*' que han impedido obtener resultados con algunos de los conjuntos de datos de prueba.
- El análisis PCA busca la extracción de componentes principales, es decir eliminar la redundancia de los datos. No es lo mismo que la búsqueda de señales independientes que persigue ICA, por lo que los resultados obtenidos con este método pueden ser erróneos.
- Con este método se han obtenido significativamente peores resultados sobre los conjuntos de datos analizados que con el resto de métodos.

5.1.5. Características principales del método 'LCC'

Las principales características de este método son:

- Se trata de un método no supervisado que proporciona un resultado numérico concreto, dejando fuera cualquier tipo de error de interpretación.
- Al obtener un resultado numérico directo no se necesita consumir tiempo adicional en la interpretación de resultados.
- No se necesita proporcionar a priori un número máximo de CIs a considerar.
- El número de análisis JADE a realizar depende del NOCI real de las señales de los datos y una vez alcanzado el NOCI no se continúa ejecutando el algoritmo. Esto supone un ahorro de tiempo considerable.
- Este método obtiene resultados homogéneos y consistentes para diferentes números de mezclas disponibles.
- Es el método con el que se han obtenido los mejores resultados en cuanto a la determinación del NOCI correctamente y el tiempo/recursos empleados para hacerlo.

5.2. Conclusiones

A lo largo de este estudio se han presentado diversos métodos para determinar el NOCI a emplear cuando se necesita aplicar ICA a un conjunto de señales de las cuales no se tiene información previa sobre número de CI que pueden haber sido la fuente de dichas señales. Los distintos métodos analizados han obtenido resultados positivos tanto con conjuntos de señales de origen sintético como con conjuntos de señales del mundo real, como las señales de audio del conjunto *Set3*. En base a las características presentadas en el apartado anterior y los resultados presentados en el capítulo 4, hay tres métodos que presentan ventajas sobre el resto en cuanto a resultados y tiempo de ejecución. Estos métodos son '*LCC*', '*ICA by blocks*' y '*Criterio de DW*'. El método '*RV ICA by blocks*' presenta resultados muy consistentes pero el tiempo de ejecución es mucho mas elevado y no siempre se puede obtener un resultado por la forma en que se calcula el coeficiente RV. Cuando se presente la situación de tener que determinar el NOCI en un conjunto de señales sobre las que no se tiene información previa, el procedimiento más óptimo sería utilizar primeramente '*LCC*' para obtener un resultado numérico concreto para NOCI y posteriormente comprobar el resultado obtenido con alguno de los otros métodos presentados o una combinación de varios de ellos. Una vez obtenido un valor numérico concreto con '*LCC*', éste nos permite establecer de forma óptima el valor para *maxICs* que necesitan otros algoritmos para trabajar. Esta método de proceder permitirá ahorrar un tiempo considerable en la determinación del NOCI y establecer *maxICs* para el uso de los otros algoritmos, que de otra forma no quedaría mas remedio que acudir al método de 'prueba y error' para determinarlo. Además, el valor obtenido por el método '*LCC*' nos proporciona una pista importante a la hora de interpretar los resultados gráficos obtenidos mediante el resto de métodos.

5.3. Futuros Trabajos

Los posibles trabajos que se proponen a partir de las conclusiones obtenidas en este método son:

- Mejorar las limitaciones que pueden presentar algunos métodos, como por ejemplo las presentadas por '*RV ICA by blocks*' y el método basado en '*PCA*' cuando el conjunto de datos a analizar comienza a ser un poco grande.
- Búsqueda de métodos para automatizar la interpretación de resultados de aquellos métodos que proporcionan resultados gráficos como el *Criterio de Durbin-Watson*, '*ICA by blocks*' y otros.

- Estudio de la robustez de los métodos atendiendo al nivel de ruido presente en las señales.
- Nuevos estudios con otros conjuntos de señales.
- Optimización de los métodos, principalmente destinadas a reducir el alto coste computacional de JADE.
- Aunque en principio para el análisis ICA el número de mezclas tiene que ser al menos igual al número de CIs, sería conveniente estudiar si en algunos casos pueden existir condiciones (a determinar) que permitan obtener el NOCI cuando el número de mezclas es inferior al número de CIs.

Apéndice A

Guía rápida de funciones

A.1. Guía rápida de funciones

A.1.1. DW.m

```
function dw = DW(R)
% dw = DW(R) is an n*1 matrix with Durbin-Watson (DW) criteria for
% temporal series in rows of nxT matrix R. Each row of R is considered
% as a temporal serie.
%
% Durbin-Criterion. Version 1 Oct 2016
%
% Usage:
% * If R is an nxT data matrix (n sensors, T samples) then
% dw=DW(R) is a nx1 matrix such that every element is the DW
% criteria of each row in R.
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.2. RV_coeff.m

```
function RV=RV_coeff(X,Y)
% RV = RV_coeff(X,Y) is a matrix with RV coefficients for X and Y
% matrices
%
% RV_coeffs. Version 1 Oct 2016
%
% Usage:
% * If X is an NxT data matrix (N sensors, T samples) and Y is other
% data
% matrix (M sensors, T samples) then RV=RV_coeff(X,Y) is a NxM matrix
% with
% RV coeffs.
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.3. matrixdcorr.m

```
function C = matrixdcorr( S )
% matrixdcorr Distance Correlation Matrix
```

```
%
% matrixdcorr Version 1 Oct 2016
%
% Usage:
% * If S is an NxT data matrix (N sensors, T samples) then
%   C = matrixdcorr(S) is a NxN Distance Correlation Matrix
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.4. DWCriterion.m

```
% DWCriterion Optimal Number of Independant Components determination by
% Durbin-Watson Criteria
%
% Input parameters:
%   X: MxN matrix. N samples of M sensors.
%   maxICs: maximun ICs for calculations
%   verbose: 0 No verbose, 1 plot DWMMatrix, 2 plot residuals
%
% Output parameters:
%   DWMMatrix: M x maxICs matrix with Durbin Watson of residual
%   matrices
%   elapsedTime: For timing purpouses
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.5. ICA_by_two_blocks.m

```
function [correlation_data, elapsedTime] = ICA_by_two_blocks(X,maxICs,
    random,verbose)
% ICA_by_two_blocks Optimal Number of Independant Components determination
% with ICA by blocks method
%
% Input parameters:
%   X: MxN matrix. N samples of M sensors.
%   maxICs: maximun ICs for calculations
%   random: true or false to apply random order to sensor data for
%   block creation
%   verbose: 0 no verbose, 1 plot correlation_data
%
% Output parameters:
%   correlation_data: Cell Array with correlation data
%   elapsedTime: For timing purpouses
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.6. RV_ICA_by_two_blocks.m

```
function [correlation_data, elapsedTime] = rv_ICA_by_two_blocks(X,maxICs,
    random,verbose)
% rv_ICA_by_two_blocks Optimal Number of Independant Components
% determination with Rv ICA by blocks method
%
% Input parameters:
%   X: MxN matrix. N samples of M sensors.
%   maxICs: maximun ICs for calculations
%   random: true or false to apply random order to sensor data for
%   block creation
```

```
%      verbose: 0 no verbose, 1 plot RV correlation_data
%
%      Output parameters:
%      correlation_data: mxICs x 1 Vector with RV coeffs
%      elapsedTime: For timing purposes
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.7. PCA_NOCI.m

```
function [autv, elapsedTime] = PCA_NOCI(X,random,verbose)
%PCA_NOCI Optimal Number of Independant Components
% determination with PCA based method
%
%      Input parameters:
%      X: MxN matrix. N samples of M sensors.
%      maxICs: maximun ICs for calculations
%      verbose: 0 no verbose, 1 plot eigenvalues
%
%      Output parameters:
%      cautov: Vector with eigenvalues
%      elapsedTime: For timing purposes
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.8. LCC_NOCI.m

```
function [ICs, NOCI, elapsedTime] = LCC_NOCI(X,verbose)
%LCC_NOCI Optimal Number of Independant Components
% determination with Linear Component Correlation method
%
%      Input parameters:
%      X: MxN matrix. N samples of M sensors.
%      verbose: 0 no verbose, 2 plot sucesive estimated ICs
%
%      Output parameters:
%      ICs: Independant Components Extracted
%      NOCI: Optimal Number of Independant Components
%      elapsedTime: For timing purposes
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.9. Test01.m

```
% Test01.m
% Script for tests over observations created from
% Set1a independant signals
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.10. Test02.m

```
% Test02.m
% Script for tests over observations created from
% Set1 independant signals
%
```

```
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.11. Test03.m

```
% Test03.m
% Script for tests over observations created from
% Set2 independant signals
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

A.1.12. Test04.m

```
% Test04.m
% Script for tests over observations created from
% Set3 independant signals
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
%
```

A.1.13. Test05.m

```
% Test05.m
% Script for timing tests over observations created from
% Set1 independant signals
%
% Author: Daniel Mercader (mercaderd@yahoo.es)
% Universidad Nacional de Educacion a Distancia
```

Bibliografía

- [Bouveresse et al., 2012] Bouveresse, D. J.-R., Moya-González, A., Ammari, F., and Rutledge, D. N. (2012). Two novel methods for the determination of the number of components in independent components analysis models. *Chemometrics and Intelligent Laboratory Systems*, 112:24–32.
- [Bouveresse and Rutledge, 2016] Bouveresse, D. J.-R. and Rutledge, D. (2016). Chapter 7 - independent components analysis: Theory and applications. In Ruckebusch, C., editor, *Resolving Spectral Mixtures-With Applications from Ultrafast Time-Resolved Spectroscopy to Super-Resolution Imaging*, volume 30 of *Data Handling in Science and Technology*, pages 225 – 277. Elsevier.
- [Cardoso and Souloumiac, 1993] Cardoso, J.-F. and Souloumiac, A. (1993). Blind beamforming for non Gaussian signals. *IEE Proceedings-F*, 140(6):362–370.
- [Haykin and Chen, 2005] Haykin, S. and Chen, Z. (2005). The cocktail party problem. *Neural computation*, 17(9):1875–1902.
- [Hyvärinen et al., 2004] Hyvärinen, A., Karhunen, J., and Oja, E. (2004). *Independent component analysis*, volume 46. John Wiley & Sons.
- [Hyvärinen and Oja, 2000] Hyvärinen, A. and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430.
- [Hyvrinen, 1999] Hyvrinen, A. (1999). Sparse code shrinkage: Denoising of nongaussian data by maximum likelihood estimation. *Neural Computation*, 11(7):1739–1768.
- [Jaakko Särelä and Ella Bingham, 2005] Jaakko Särelä, P. H. and Ella Bingham, P. S. (2005). Cocktail party demo. https://research.ics.aalto.fi/ica/cocktail/cocktail_en.cgi.

- [MATLAB, 2014] MATLAB (2014). *version 8.3.0 (R2014a)*. The Math-Works Inc., Natick, Massachusetts.
- [Mur et al., 2017] Mur, A., Dormido, R., Duro, N., and Mercader, D. (2017). An unsupervised method to determine the optimal number of independent components. *Expert Systems with Applications*, 75:56 – 62.
- [Rutledge and Barros, 2002] Rutledge, D. and Barros, A. (2002). Durbin–watson statistic as a morphological estimator of information content. *Analytica Chimica Acta*, 454(2):277 – 295.
- [Smilde et al., 2009] Smilde, A. K., Kiers, H. A. L., Bijlsma, S., Rubingh, C. M., and van Erk, M. J. (2009). Matrix correlations for high-dimensional data: the modified rv-coefficient. *Bioinformatics*, 25(3):401.
- [Vigário, 1997] Vigário, R. N. (1997). Extraction of ocular artefacts from eeg using independent component analysis. *Electroencephalography and clinical neurophysiology*, 103(3):395–404.

Acrónimos

CPP: Cocktail-Party Problem.

BSS: Blind Source Separation.

DW: Criterio de Durbin-Watson.

EEG: Electroencefalografía.

CI: Componente Independiente.

ICA: Análisis de Componentes Independientes.

CI: Componentes Independientes.

JADE: Joint Approximate Diagonalization of Eigenmatrices.

LCC: Linear Correlations between Components.

NCI: Número de Componentes Independientes.

NOCI: Número Óptimo de Componentes Independientes.

PCA: Análisis de Componentes Principales.

SNR: Relación Señal-Ruido.