

Separación ciega de mezclas convolutivas utilizando estadísticas de segundo orden y una máscara binaria

J. R. Villavicencio

*Instituto Tecnológico de Sonora, Av. Antonio Caso, Itson, C.P.85132,
Cd. Obregón, Sonora, México (e-mail:
jesus.villavicencio@itson.edu.mx).*

Resumen

En este trabajo se presenta un método híbrido (dominio del tiempo/frecuencia) para la solución de separación ciega de mezclas convolutivas. La presente metodología elimina el problema de permutación en la frecuencia, problema que inherente a los métodos de separación de mezclas convolutivas diseñados en el dominio de la frecuencia. Los resultados muestran un buen desempeño del algoritmo propuesto.

Keywords: Señales no estacionarias; Procesamiento de señales; Relación Señal a Ruido; Separación ciega de fuentes; Mezclas Convolutivas.

1. INTRODUCCIÓN

Durante las décadas pasadas se ha dado mucha importancia a la separación ciega de fuentes (BSS, por sus siglas en inglés) debido a las posibles aplicaciones dentro de las telecomunicaciones, procesamiento de señales biomédicas, audio, etc. [Romero et al., 2008, Hesse, 2008, Crespo-Garcia et al., 2008, Romo Vázquez et al., 2012, Comon and Jutten, 2010]. La importancia del BSS radica en recuperar fuentes originales que se encuentran latentes en observaciones adquiridas mediante múltiples sensores. La idea principal del BSS consiste en recuperar todas o alguna fuente original en particular, utilizando solamente estadísticas de los datos observados. Las funciones de costo utilizadas para recuperar las fuentes originales están basadas en estadísticas de alto-orden (HOS, por sus siglas en inglés) [Hyvärinen, 1999, Cardoso and Souloumiac, 1993] o estadísticas de segundo orden (SOS, por sus siglas en inglés) [Tong et al., 1991, Belouchrani et al., 1997, Choi and Cichocki, 2000].

El caso más sencillo de separación ciega de fuentes, BSS de mezclas instantáneas, supone que cada observación consiste en la suma ponderada de las fuentes originales. Este fenómeno puede representarse por el siguiente modelo:

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{V}(k), \quad k = 1, 2, \dots \quad (1)$$

donde $\mathbf{x}(k) \in \mathbb{R}^N$ son las observaciones, $\mathbf{s}(k) \in \mathbb{R}^M$ son las fuentes desconocidas, $\mathbf{A} \in \mathbb{R}^{N \times M}$ es la matriz que modela el proceso de mezclado, la cual también es desconocida, $\mathbf{V}(k) \in \mathbb{R}^M$ es ruido blanco gaussiano adherido y k es el índice de tiempo.

En varias aplicaciones de la vida real, tal como es el caso de la acústica y telecomunicaciones, el proceso de mezclado de las señales es más complejo. En dichos sistemas las señales pueden estar filtradas y retardadas, y además cada fuente puede contribuir con múltiples trayectorias por la cual la señal se propaga e incide sobre el sensor. A este tipo de

señales se le denominan mezclas convolutivas. El modelo de mezclas convolutivas se puede representar de la siguiente manera:

$$\mathbf{x}(n) = \sum_{k=0}^{K-1} \mathbf{A}_k \mathbf{s}(n-k), \quad n = 1, 2, \dots \quad (2)$$

donde \mathbf{A}_k es una matriz de $N \times M$ la cual contiene los filtros mezcladores de longitud K .

El problema de mezclas convolutivas se puede resolver al encontrar el filtro inverso de \mathbf{A}_k dado que éste exista. No es necesario estimar exactamente el filtro \mathbf{A}_k , a menudo es suficiente con estimar un filtro separador \mathbf{W}_l que elimine la diafonía introducida por el proceso de mezclado. Estos filtros separadores pueden ser de estructura respuesta infinita al impulso (IIR, por sus siglas en inglés) o respuesta finita al impulso (FIR, por sus siglas en inglés) Amari and Cichocki [1998].

Un enfoque para resolver el problema se ha planteado mediante el diseño de algoritmos en el dominio de la frecuencia. Tenemos que el modelo sin ruido blanco gaussiano en el dominio de la frecuencia se representa por la siguiente ecuación:

$$\mathbf{X}(f) = \mathbf{A}(f)\mathbf{S}(f) \quad (3)$$

Al trabajar en la frecuencia, el problema de mezclas convolutivas se reduce al problema de mezclas instantáneas. Ahora, el objetivo es en encontrar los coeficientes de $\mathbf{W}(f)$ para cada f . Donde las fuentes estimadas están dadas por:

$$\hat{\mathbf{S}}(f) = \mathbf{W}(f)\mathbf{X}(f) = \mathbf{W}(f)\mathbf{A}(f)\mathbf{S}(f) = \mathbf{G}(f)\mathbf{S}(f) \quad (4)$$

Sin embargo, a pesar de que al realizar ésta transformación se evita el proceso de la desconvolución surgen otros problemas. Entre los nuevos problemas que hay que resolver al utilizar éste planteamiento se encuentran la ambigüedad

de permutación en la frecuencia [Pedersen et al., 2007, Makino et al., 2005]. Esto se debe a que el filtro global

$$\mathbf{G}(f) = \mathbf{P}(f)\mathbf{A}(f) \quad (5)$$

puede representarse mediante la multiplicación de una matriz de permutación $\mathbf{P}(f)$ y una matriz con filtros en su diagonal $\mathbf{A}(f)$. El problema principal es que se corre el riesgo de que la matriz de permutación $\mathbf{P}(f)$ no sea consistente en todas las frecuencias, destruyendo de esta manera la separación que se logró en la frecuencia, al transformar $\hat{\mathbf{S}}(f)$ al dominio del tiempo, tal como se explica en Pedersen et al. [2007].

En este trabajo se presenta una metodología híbrida basada en el dominio del tiempo y en el dominio de la frecuencia evitando así el problema de permutación en la frecuencia. El método consiste en optimizar filtros inversos FIR en el dominio del tiempo mediante el método heurístico recocido simulado y calcular el costo en el dominio de la frecuencia. Mediante el uso de ésta metodología se evita la ambigüedad de permutación en la frecuencia. La función de costo que se propone es la suma de la coherencia (MSC, por sus siglas en inglés) en conjunto con una máscara binaria en el dominio de la frecuencia. La función de costo propuesta es una técnica basada en el enfoque de SOS.

2. METODOLOGÍA

En el desarrollo de este trabajo haremos las siguientes suposiciones:

1. La matriz de los filtros mezcladores \mathbf{A} es de rango pleno y tiene su inversa, \mathbf{W} .
2. Las señales $\mathbf{s}(t)$ están mutuamente descorrelacionadas y temporalmente autocorrelacionadas.
3. $\mathbf{V}(k)$ es ruido blanco gaussiano aditivo y descorrelacionado de las fuentes originales.
4. Existe f en donde $X(f) \gg V(f)$.

El sistema MIMO-LTI $(\mathbf{A}_k)_{k \in \mathbb{Z}}$, dado que exista, tiene su sistema inverso comúnmente llamado separador. Encontrar el sistema separador equivale a encontrar las fuentes originales y por consiguiente resolver el problema de BSS de mezclas convolutivas. Su respuesta al impulso está denotada por $(\mathbf{W}_l)_{l \in \mathbb{Z}}$, en donde las salidas separadas están dadas por:

$$\hat{\mathbf{s}}(n) = \sum_{l \in \mathbb{Z}} \mathbf{W}_l \mathbf{x}(n-l) \quad \forall l \in \mathbb{Z}. \quad (6)$$

El sistema mezclador y separador, con respectivas respuestas al impulso $(\mathbf{A}_k)_{k \in \mathbb{Z}}$ y $(\mathbf{W}_l)_{l \in \mathbb{Z}}$, están definidas en el dominio de la transformada z de la siguiente manera:

$$\mathbf{A}[z] \triangleq \sum_{k \in \mathbb{Z}} \mathbf{A}(k)z^{-k} \quad y \quad \mathbf{W}[z] \triangleq \sum_{k \in \mathbb{Z}} \mathbf{W}(k)z^{-k}. \quad (7)$$

La salida global del sistema combinado mezclador-separador está dada por:

$$\hat{\mathbf{s}}(n) = \sum_{k \in \mathbb{Z}} \mathbf{G}(k)\mathbf{s}(n-k) \quad \forall n \in \mathbb{Z}, \quad (8)$$

donde la respuesta al impulso y la transformada- z del sistema global $(\mathbf{G}(n))_{n \in \mathbb{Z}}$ están dadas respectivamente por las ecuaciones:

$$\mathbf{G}(n) = \sum_{k \in \mathbb{Z}} \mathbf{B}(n-k)\mathbf{A}(k) \quad y \quad \mathbf{G}[z] = \mathbf{B}[z]\mathbf{A}[z]. \quad (9)$$

2.1 La coherencia en el modelo convolutivo BSS

La MSC es una función *normalizada* que mide la correlación entre dos señales en el dominio de la frecuencia, y puede tomar valores entre $[0 \leq |\gamma_{\hat{s}_l \hat{s}_k}(f)|^2 \leq 1]$. Cuando $|\gamma_{\hat{s}_l \hat{s}_k}(f)|^2 = 0$ indica que \hat{s}_l está totalmente descorrelacionada de \hat{s}_k en la frecuencia f . Cuando $|\gamma_{\hat{s}_l \hat{s}_k}(f)|^2 = 1$ indica que \hat{s}_l está totalmente correlacionada con \hat{s}_k en la frecuencia f . Si seleccionamos tal f en donde la potencia de las fuentes no-estacionarias es mucho mayor que la del ruido blanco gaussiano, $X(f) \gg V(f)$, entonces podemos desestimar el ruido del modelo y la coherencia entre las salidas del sistema global quedaría de la siguiente manera

$$|\gamma_{\hat{s}_1 \hat{s}_2}(f)|^2 = \frac{|P_{\hat{s}_1 \hat{s}_2}(f)|^2}{P_{\hat{s}_1 \hat{s}_1}(f)P_{\hat{s}_2 \hat{s}_2}(f)}, \quad (10)$$

donde $\hat{s}_1(n)$ y $\hat{s}_2(n)$ son las salidas del sistema mezclador-separador para dos fuentes, representado por el sistema de la Figura 1. Si calculamos la MSC entre las salidas del sistema de la Figura 1

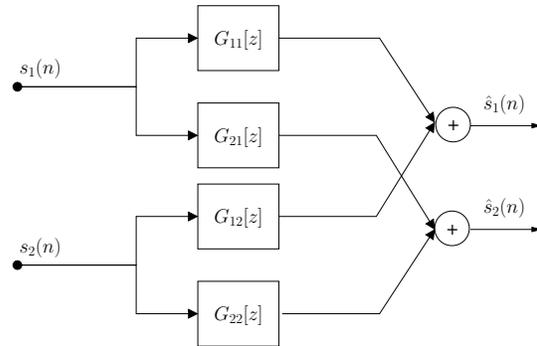


Figura 1. Sistema global para 2 fuentes.

$$P_{\hat{s}_1 \hat{s}_1} = |G_{11}(f)|^2 P_{s_1 s_1}(f) + |G_{12}(f)|^2 P_{s_2 s_2}(f), \quad (11)$$

$$P_{\hat{s}_2 \hat{s}_2} = |G_{21}(f)|^2 P_{s_1 s_1}(f) + |G_{22}(f)|^2 P_{s_2 s_2}(f), \quad (12)$$

$$P_{\hat{s}_1 \hat{s}_2} = G_{11}(f)G_{21}^*(f)P_{s_1 s_1}(f) + G_{12}(f)G_{22}^*(f)P_{s_2 s_2}(f), \quad (13)$$

tenemos por resultado la siguiente ecuación:

$$|\gamma_{\hat{s}_1 \hat{s}_2}(f)|^2 = \frac{|G_{11}(f)G_{21}(f)^*|^2 P_{s_1 s_1}^2(f) + |G_{12}(f)G_{22}(f)^*|^2 P_{s_2 s_2}^2(f)}{[|G_{11}(f)|^2 P_{s_1 s_1}(f) + |G_{12}(f)|^2 P_{s_2 s_2}(f)][|G_{21}(f)|^2 P_{s_1 s_1}(f) + |G_{22}(f)|^2 P_{s_2 s_2}(f)]}. \quad (14)$$

Podemos observar que el numerador de la ecuación (14) es cero si y solo si el sistema mezclador-separador $\mathbf{G}[z]$ es una matriz diagonal, i.e., $G_{12}[z] = G_{21}[z] = 0$. El minimizar la ecuación (14) será idéntico a encontrar el filtro inverso $\mathbf{W}[z]$ y, por consiguiente, a recuperar las fuentes originales. En este escenario se pueden recuperar los componentes independientes hasta las ambigüedades de filtrado y escalamiento mencionadas anteriormente. Como se observa en la ecuación (14), el factor de escalamiento no es distinguible en la función de costo, es decir, la función de coherencia no es sensible a los valores de la diagonal del filtro mezclador-separador $\mathbf{G}[z]$. Por lo tanto, los componentes recuperados serán una versión filtrada de los componentes originales.

2.2 Criterio de separación

Basándonos en la ecuación (14), podemos construir una función de costo para separar las fuentes originales (componentes independientes) de las mezclas convolutivas. El criterio adoptado es la suma de la MSC tal como se realizó en Fancourt and Parra [2001] a excepción de una modificación que se ha realizado. Aquí se añadió una máscara binaria de una dimensión.

$$\mathbf{J} = \sum_{l,k} \sum_{f=0}^P |\gamma_{\hat{s}_l \hat{s}_k}(f)|^2 \varphi(f) \quad (15)$$

donde $|\gamma_{\hat{s}_l \hat{s}_k}(f)|^2$ es la MSC y $\varphi(f)$ es la máscara binaria de una dimensión. De esta manera, la ecuación (15) se centra en calcular la suma de la coherencia en aquellas frecuencias en donde hay una relación de señal a ruido relativamente grande, seleccionada por la función $\varphi(f)$.

2.3 Máscara binaria en el dominio de la frecuencia $\varphi(f)$

La máscara binaria $\varphi(f)$ selecciona las frecuencias en donde la relación señal-ruido de las observaciones es relativamente mayor que la del ruido blanco gaussiano adherido. El algoritmo 1 muestra la creación de la máscara binaria. Debido a que el espectro del ruido blanco gaussiano tiene una distribución uniforme, es relativamente sencillo obtener una estimación de él en la frecuencia y por lo tanto crear un algoritmo que estime el umbral del ruido. La densidad espectral de las observaciones se calcula mediante el método de Welch y los componentes frecuenciales en donde la SNR es grande se seleccionan de acuerdo a un umbral ψ calculado de la siguiente manera

$$\psi = \mu_{P_{xx}} + \lambda \sigma_{P_{xx}} \quad (16)$$

donde $\mu_{P_{xx}}$, y $\sigma_{P_{xx}}$ son la media y la desviación estándar de la densidad espectral de potencia (PSD, por sus siglas en inglés) del vector de observación, respectivamente. λ es una constante de ajuste del umbral.

3. EXPERIMENTOS

En esta sección se muestran experimentos del método propuesto para dos fuentes.

Algorithm 1 Máscara Binaria $\varphi(f)$

Require: vector de observaciones \mathbf{x}

Ensure: Función $\varphi(f)$

- 1: Calcula la PSD, $P_{xx}(f)$, del vector de observación $\mathbf{x}(n)$ mediante el método de Welch.
 - 2: Calcula la media y la desviación estándar de $P_{xx}(f)$, $\mu_{P_{xx}}$ y $\sigma_{P_{xx}}$ respectivamente.
 - 3: Calcula ψ , mediante la ecuación (16).
 - 4: **for all** f **do**
 - 5: **if** $P_{xx}(f) \geq \psi$ **then**
 - 6: $\varphi(f) = 1$
 - 7: **else**
 - 8: $\varphi(f) = 0$
 - 9: **end if**
 - 10: **end for**
 - 11: $\varphi(f)$ selecciona el ancho de banda con la mayor SNR.
-

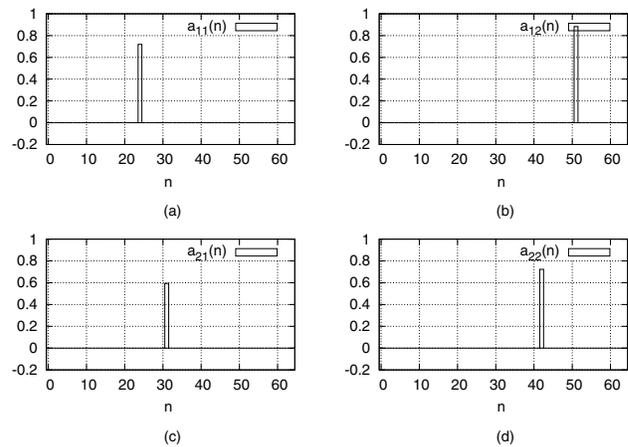


Figura 2. Filtros mezcladores.

3.1 Simulación

Dos señales estadísticamente independientes $s_1(n)$ y $s_2(n)$, Figura 6 a) y b), han sido mezcladas por el modelo convolutivo (17) generando el vector $x_1(n)$ y $x_2(n)$, Figura 6 c) y d). El modelo convolutivo está dado por

$$x_i(n) = \sum_{j=1}^N \sum_{k=0}^{M-1} a_{ij}(k) s_j(n-k), \quad (17)$$

donde N es el número de fuentes y sensores, M es la longitud del filtro mezclador, y D_{ij} es el retardo entre la i -ésima fuente y j -ésimo sensor. En notación de transformada z , el sistema queda de la siguiente manera:

$$\begin{aligned} A_{11}[z] &= 0,7194z^{-25} \\ A_{12}[z] &= 0,8828z^{-52} \\ A_{21}[z] &= 0,5934z^{-32} \\ A_{22}[z] &= 0,7228z^{-43} \end{aligned} \quad (18)$$

4. RESULTADOS

Como se observa en la Figura 3, después de 8,000 iteraciones el algoritmo encontró el sistema separador. La función de costo $\mathbf{J} = 0,054378$ indica una buena separación. Las señales recuperadas se muestran en la Figura 6 e) y 6 f). En la Figura 7 vemos la MSC entre las fuentes estimadas y la máscara binaria $\varphi(f)$ (en color azul). La MSC cercana

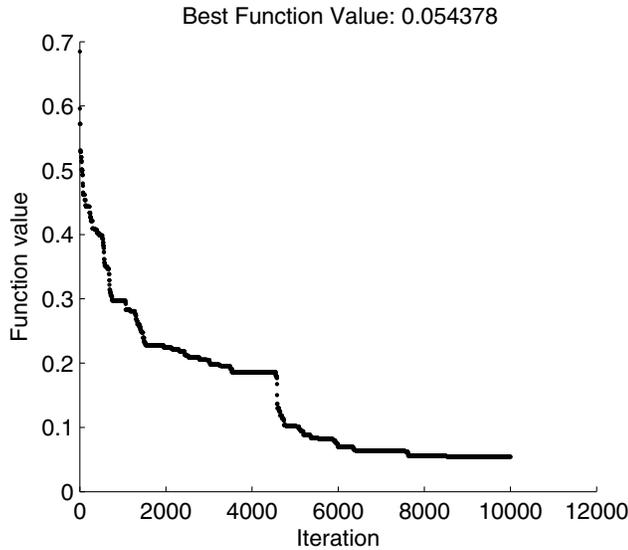


Figura 3. Evolución de la función de costo en el algoritmo de optimización propuesto.

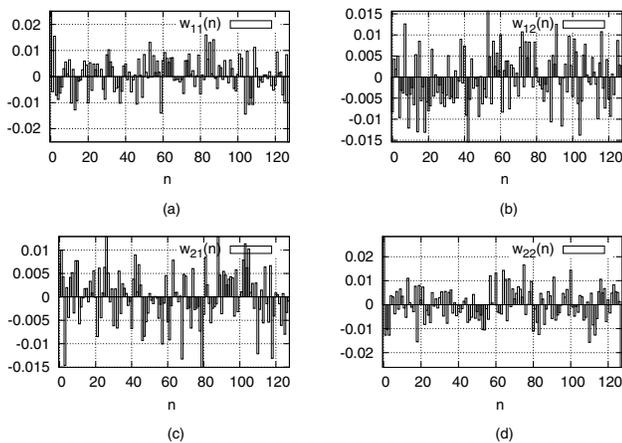


Figura 4. Parámetros del sistema separador final.

a cero significa que las señales están descorrelacionadas en ese intervalo de frecuencias. El intervalo de frecuencias es seleccionada por la máscara binaria. Fuera del ancho de banda de interés hay muy poca energía, aunque las fuentes estimadas estén altamente correlacionadas, esas frecuencias pueden ignorarse.

5. CONCLUSIONES

En este trabajo se presentó un método híbrido (dominio del tiempo/frecuencia) para la solución de separación ciega de mezclas convolutivas. La presente metodología elimina el problema de permutación en la frecuencia, problema que inherente a los métodos de separación de mezclas convolutivas en el dominio de la frecuencia. Como se muestran en los resultados, el algoritmo propuesto tuvo un buen desempeño. Para trabajo futuro se plantea buscar nuevos métodos para reducir la complejidad de los filtros separadores así como buscar nuevos algoritmos heurísticos de optimización para mejorar la rapidez de convergencia del método.

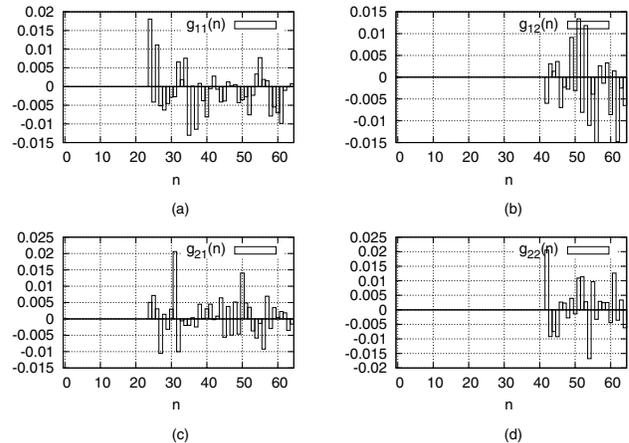


Figura 5. Parámetros del sistema global.

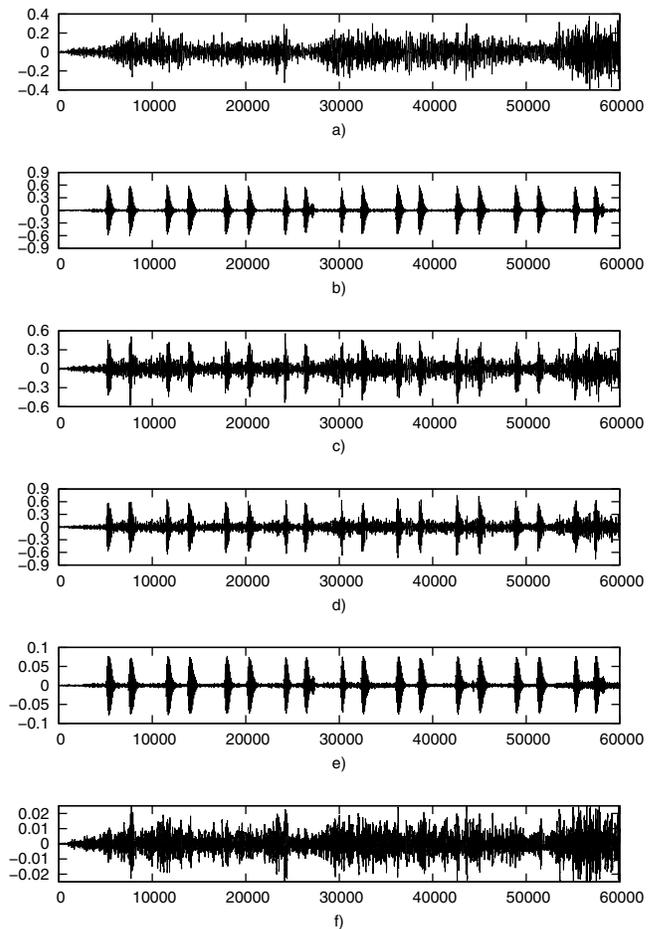


Figura 6. Ejemplo de desconvolución de 2 señales: a) y b) Señales originales. c) y d) Señales mezcladas y convolucionadas. e) y f) Señales recuperadas después de 10,000 iteraciones.

REFERENCIAS

Shun-ichi Amari and Andrzej Cichocki. Adaptive blind signal processing-neural network approaches. *Proceedings of the IEEE*, 86(10):2026–2048, 1998.

A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines. A blind source separation technique

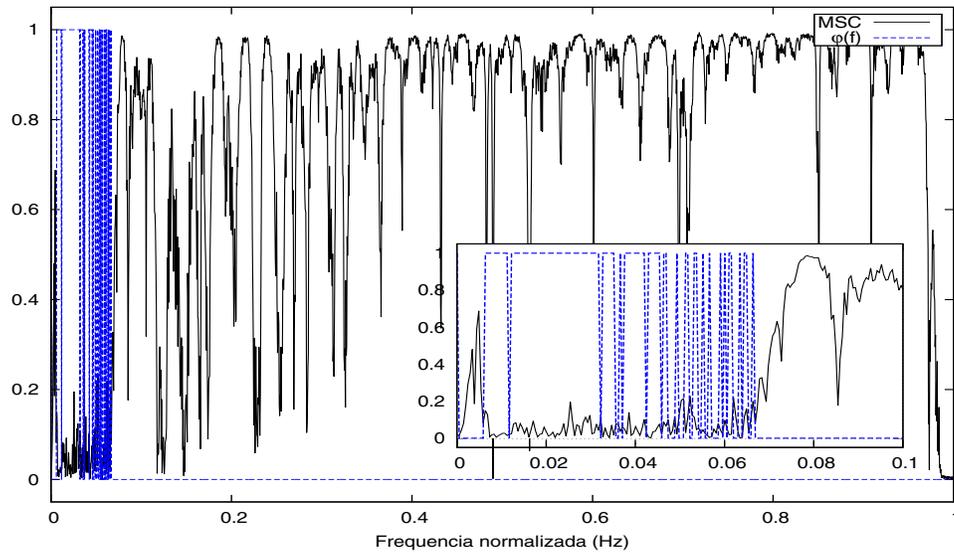


Figura 7. Coherencia entre las señales originales estimadas. La coherencia es aproximadamente cero en el ancho de banda seleccionado por la función $\varphi(f)$ (en color azul-línea punteada), indicando descorrelación entre las señales recuperadas.

- using second-order statistics. *Signal Processing, IEEE Transactions on*, 45(2):434–444, feb 1997. ISSN 1053-587X. doi: 10.1109/78.554307.
- J.F. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. *Radar and Signal Processing, IEE Proceedings F*, 140(6):362–370, dec 1993. ISSN 0956-375X.
- S. Choi and A. Cichocki. Blind separation of nonstationary sources in noisy mixtures. *Electronics Letters*, 36(9):848–849, apr 2000. ISSN 0013-5194. doi: 10.1049/el:20000623.
- P. Comon and C. Jutten. *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Academic Press, 1st edn edition, 2010. ISBN 0123747260, 9780123747266.
- Maite Crespo-García, Mercedes Atienza, and Jose L Cantero. Muscle artifact removal from human sleep eeg by using independent component analysis. *Annals of biomedical engineering*, 36(3):467–475, 2008.
- C.L. Fancourt and L. Parra. The coherence function in blind source separation of convolutive mixtures of nonstationary signals. In *Proc. NNSP*, pages 303–312, 2001.
- Christian W Hesse. Model order estimation for blind source separation of multichannel magnetoencephalogram and electroencephalogram signals. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 3348–3351. IEEE, 2008.
- Michael S Pedersen, Jan Larsen, Ulrik Kjems, and Lucas C Parra. A survey of convolutive blind source separation methods. *Multichannel Speech Processing Handbook*, pages 1065–1084, 2007.
- Sergio Romero, Miguel A Mañanas, and Manel J Barbanj. A comparative study of automatic techniques for ocular artifact reduction in spontaneous eeg signals based on clinical target variables: a simulation case. *Computers in biology and medicine*, 38(3):348–360, 2008.
- Rebeca Romo Vázquez, Hugo Vélez-Pérez, Radu Ranta, Valérie Louis Dorr, Didier Maquin, and Louis Maillard. Blind source separation, wavelet denoising and discriminant analysis for eeg artefacts and noise cancelling. *Biomedical Signal Processing and Control*, 7(4):389–400, 2012.
- L. Tong, R.-w. Liu, V.C. Soon, and Y.-F. Huang. Indeterminacy and identifiability of blind identification. *Circuits and Systems, IEEE Transactions on*, 38(5):499–509, may 1991. ISSN 0098-4094. doi: 10.1109/31.76486.
- A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, 10(3):626–634, may 1999. ISSN 1045-9227. doi: 10.1109/72.761722.
- Shoji Makino, Hiroshi Sawada, Ryo Mukai, and Shoko Araki. Blind source separation of convolutive mixtures of speech in frequency domain. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 88(7):1640–1655, 2005.