

Independent Component Analysis

INTEGRANTES:

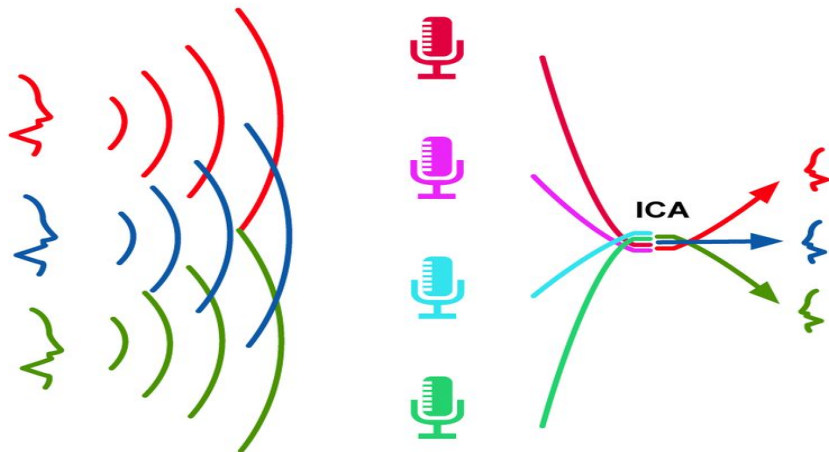
VIVIANA ANDREA GARCIA MONJE
VALERIA ROMERO PEREZ
YENIFER PATRICIA GUAJE NIÑO

A

Source Signals

Recorded Mixed Signals

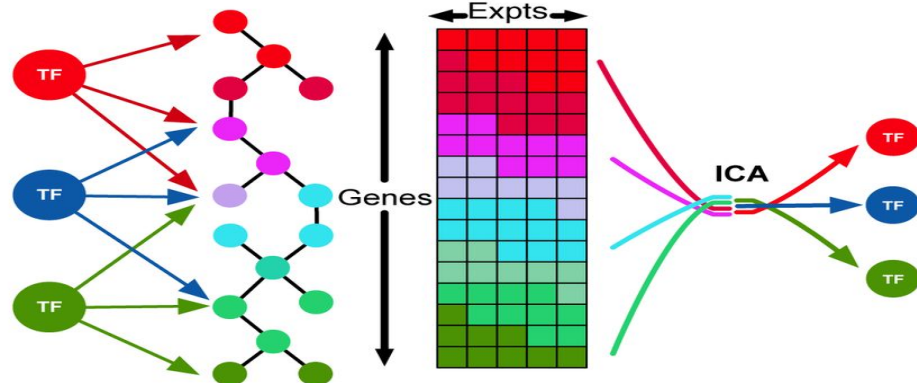
Derived Source Signals

**B**

Source Signals

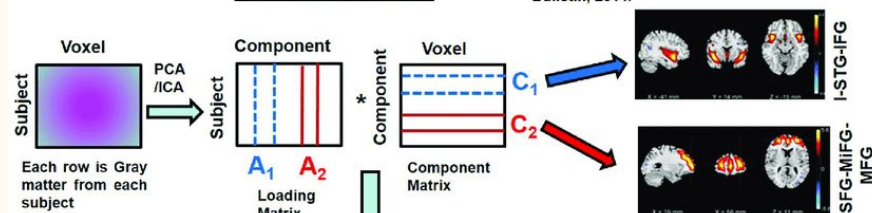
Recorded Mixed Signals

Derived Source Signals



B-ICA Algorithm

Step 1. ICA/SBM Module



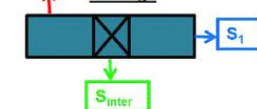
Two Components based on highest effect size from C.N.Gupta et al, Schizophrenia Bulletin, 2014.

Step 2. Sorting of loadings and subject name intersection



S₂

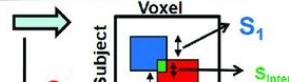
**Step 3. Subtypes
deciphered from
loadings**



Step 4. Subtype specific component maps using GIG-ICA (Y.Du et al, Neuroimage, 2013)

Maps presented in Figure 2

Step 5. Subject by voxel matrix (From Step 1) rearranged by B-ICA to detect biclusters (corresponding to Red, blue and green subtypes as in Step 3)

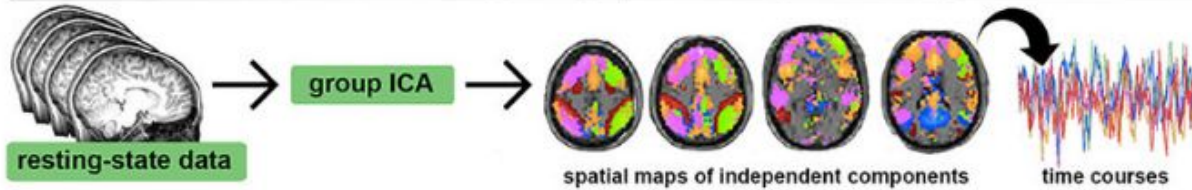


¿Que es?

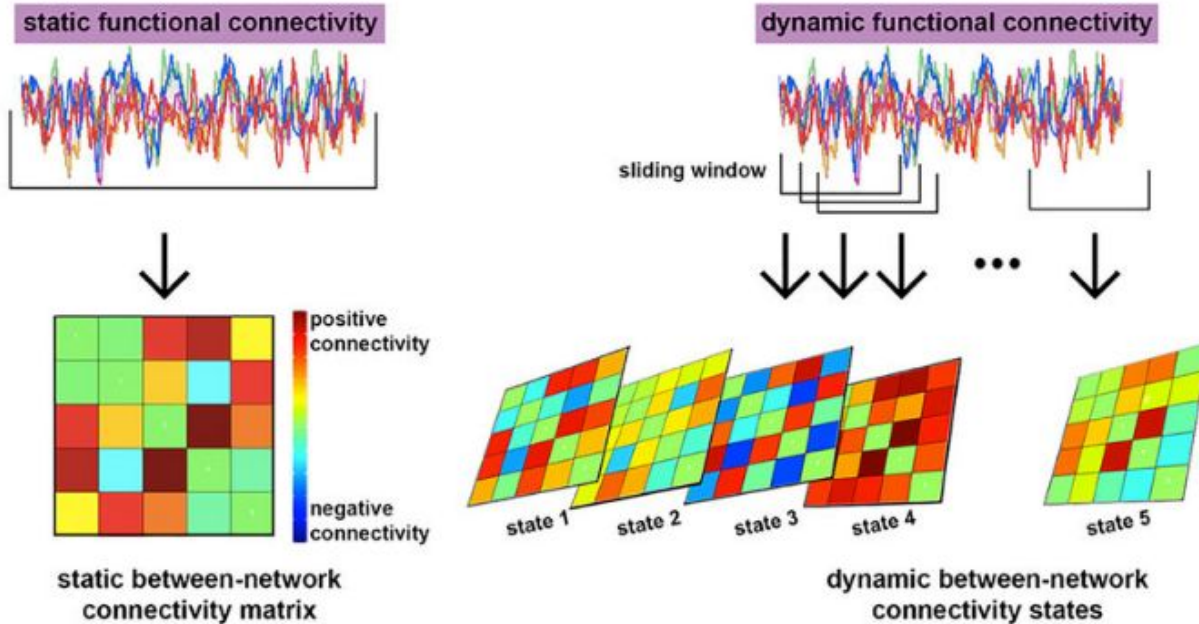
Es un método computacional que sirve para separar una señal multivariante en subcomponentes aditivos suponiendo que la señal de origen tiene una independencia estadística y es no-Gausiana. Éste es un caso especial de separación ciega de las señales.

El análisis de componentes independientes (ACI) es una generalización del análisis de componentes principales (ACP), en ambos casos se practica una transformación lineal de los datos originales, aunque la diferencia básica es que el ACI no requiere que las variables originales tengan una distribución gaussiana.

A. Identification of networks using independent component analysis



B. Estimation of static and dynamic functional network connectivity



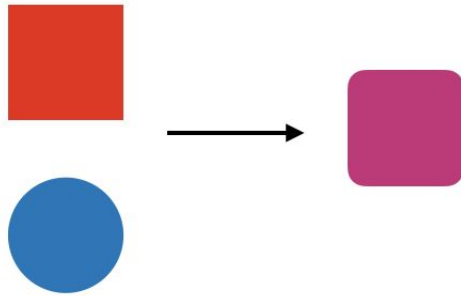
Representación esquemática del análisis de componentes independientes a nivel de grupo (ICA) y la estimación de conectividad funcional en estado de reposo estática y dinámica (rsFC). a) Conjuntos de datos en estado de reposo preprocesados y sin ruido

¿Cual es su Objetivo?

El objetivo fundamental del Análisis de Componentes Independientes (ICA) es el de proporcionar un método que permita encontrar una representación lineal de los datos no gaussianos de forma que las componentes sean estadísticamente independientes o lo más independiente posible. Una representación de este tipo permite obtener la estructura fundamental de los datos en muchas aplicaciones, incluidas la extracción de características y la separación de señales.

PCA

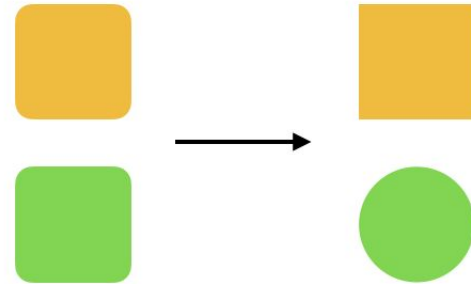
Compresses information



Requires preprocessing: autoscaling

ICA

Separates information

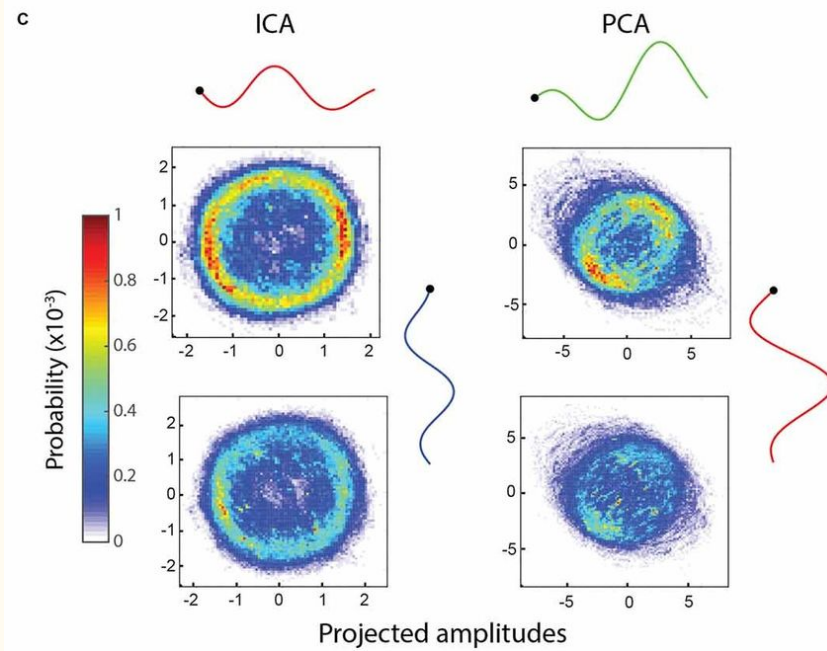
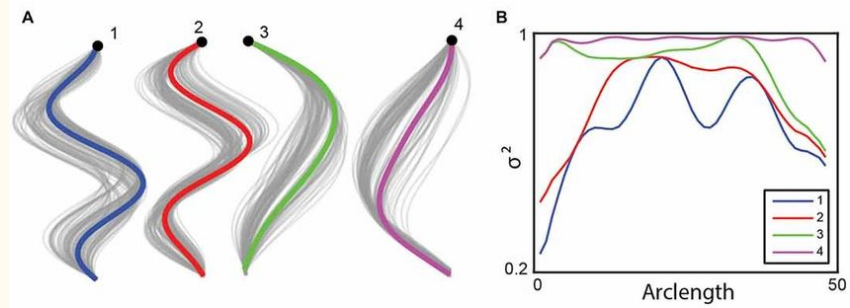


Requires preprocessing: autoscaling

Often benefits from first applying PCA

¿De donde provienen los datos analizados por ICA?

Los datos analizados por ICA podrían provenir de muchos tipos diferentes de campos de aplicación, incluidas imágenes digitales, bases de datos de documentos, indicadores económicos y mediciones psicométricas. En muchos casos, las medidas se dan como un conjunto de señales paralelas o series de tiempo; el término separación ciega de la fuente se utiliza para caracterizar este problema. Ejemplos típicos son mezclas de señales de voz simultáneas que han sido captadas por varios micrófonos, ondas cerebrales registradas por múltiples sensores, señales de radio interferentes que llegan a un teléfono móvil o series de tiempo paralelas obtenidas de algún proceso industrial.



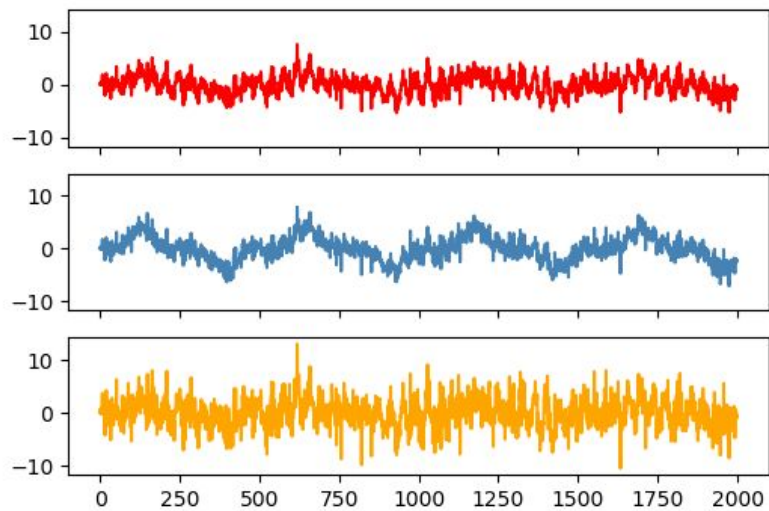
Características

El método ICA tiene múltiples características pero las más destacadas son:

- El número de entradas es igual al número de salidas
- Asume que los componentes independientes son estadísticamente independientes
- Asume que los componentes independientes no son gaussianos (medida de no gaussianidad)
- Las entradas deben ser valores autoescalados (restar cada columna por su media y dividir por su desviación estándar)

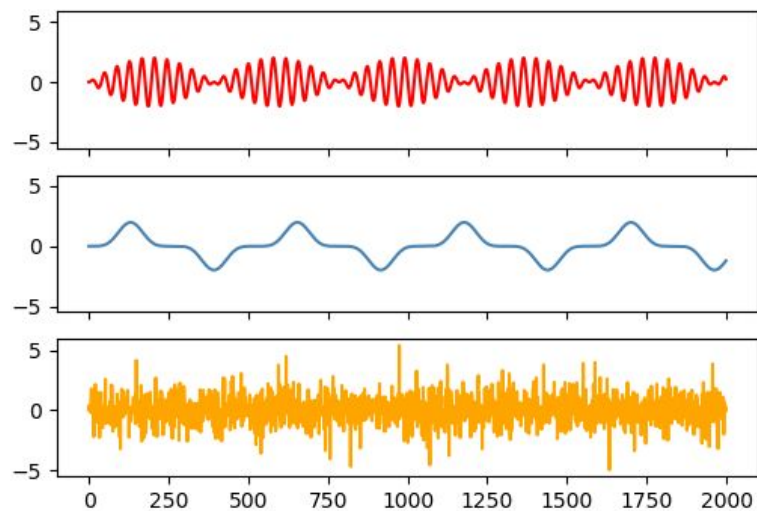
Se diferencia del método PCA por su diferencia de enfoque. El PCA se usa para comprimir información (reducción de dimensionalidad) mientras que el ICA se encarga de separar la información.

Observations (mixed signal)



ICA
⇒

True Sources



Metodología Algoritmo ICA

El análisis de componentes independientes (ICA) es otro método efectivo para la extracción de señales y reducción de dimensión de los datos. el método ICA se buscan componentes que sean estadísticamente independientes

Para aplicar los métodos clásicos es necesario realizar un preprocesado de los datos con la finalidad de que nuestra base de datos sea estacionaria en media y en varianza.

Reducción de la dimensión:

Se selecciona la medida de desempeño ante la reducción para cada dimensión del espacio.

En los procesos industriales la dimensión del vector de mediciones (el número de sensores), puede ser muy grande, y muchos de sus elementos pueden ser redundantes o incluso irrelevantes con respecto al proceso de clasificación.

Existen múltiples razones para reducir la dimensión de un vector de mediciones a un mínimo suficiente. La complejidad computacional es una razón obvia. Otra razón muy importante es que un incremento excesivo de la dimensión causa un decremento del desempeño del clasificador cuando no se poseen suficientes datos de cada variable.

Análisis de componentes independientes.

ICA es una técnica emergente para la reducción de dimensión en diagnóstico de fallos, la cual tiene como objetivo encontrar, a partir de una transformación lineal, una representación de un conjunto de variables donde se minimice la dependencia estadística entre las nuevas variables que la forman, llamadas componentes independientes

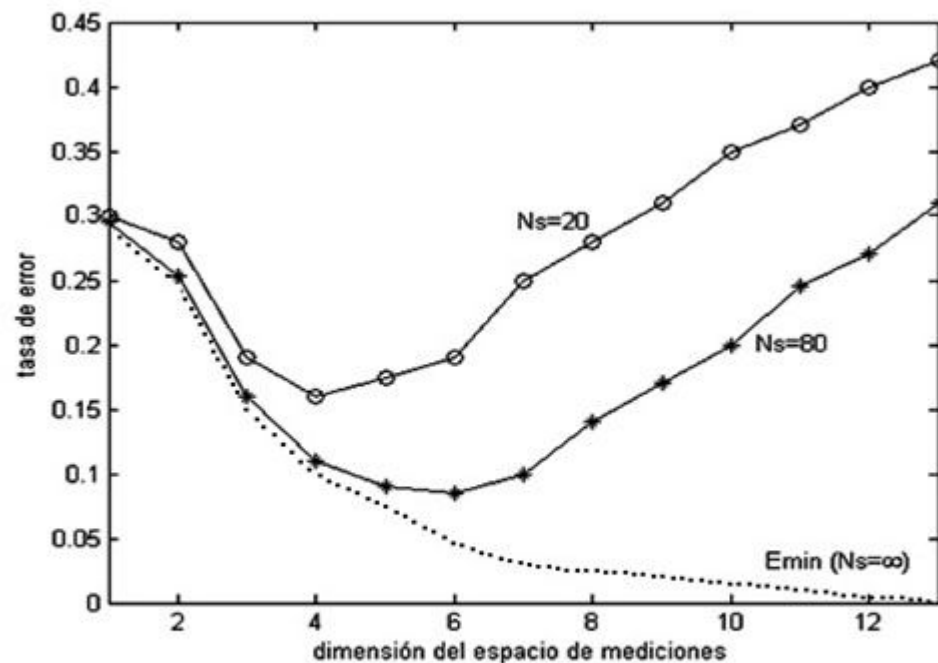


Figura 1. Tasa de error vs dimensión del espacio de mediciones

La Figura ilustra como el uso de técnicas de reducción de dimensión puede simplificar y mejorar positivamente el procedimiento de diagnóstico en estos casos. El eje de las abscisas muestra el número de variables utilizadas para la clasificación, el eje de las ordenadas muestra la tasa de error y Ns indica la cantidad de muestras de cada variable para cada experimento. En el caso ideal, cuando Ns! « y al utilizar todas las variables, se observa que el error de clasificación se aproxima a cero. Sin embargo, cuando se poseen pocas observaciones respecto al número de variables utilizadas el error de clasificación puede aumentar en la medida en que se incorporen nuevas variables.

Configuración del algoritmo FastICA.

El algoritmo FastICA se basa en estimar las componentes independientes a partir del denominado modelo ICA libre de ruido y que se presenta a continuación:

Definición 1. El modelo ICA de un conjunto de n variables aleatorias $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, consiste en:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (1)$$

donde $\mathbf{s} = (s_1, s_2, \dots, s_n)^T$ es un conjunto de variables aleatorias estadísticamente independientes y \mathbf{A} es una matriz cuadrada, llamada matriz de mezcla.

Las dos versiones del algoritmo más aplicadas son FastICA basado en maximización de la no gaussianidad y FastICA basado en estimación de la máxima probabilidad de independencia.

Estrategias para la selección de las componentes independientes al aplicar ICA.

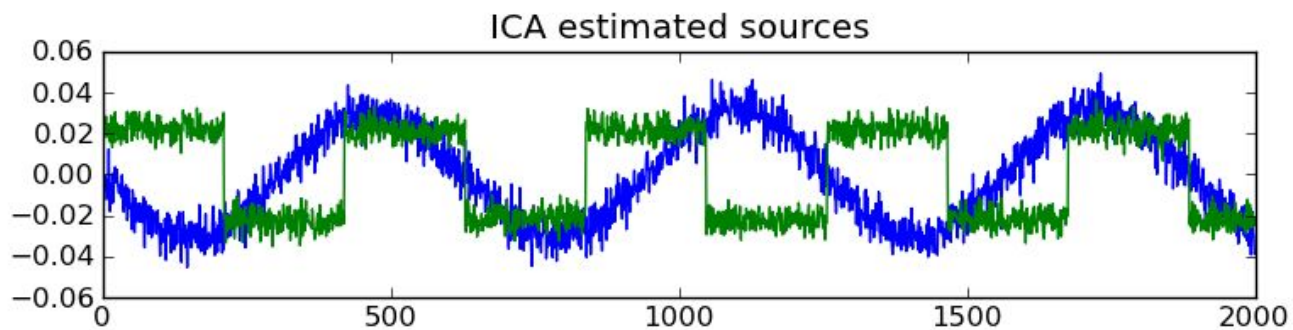
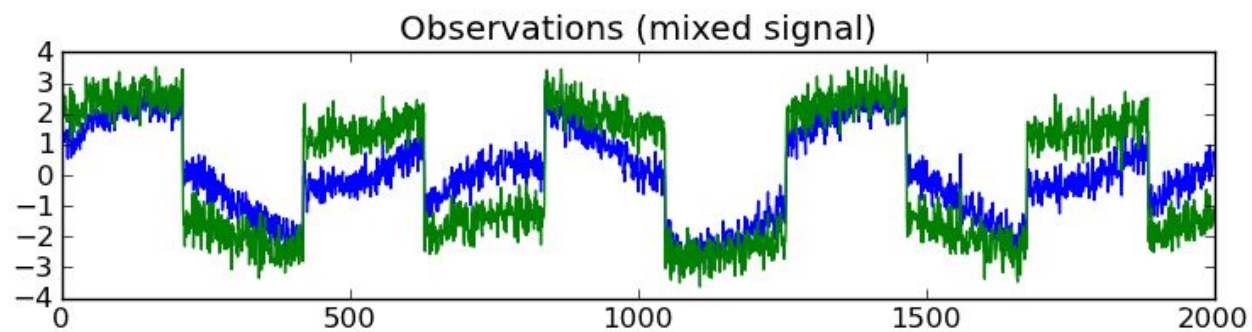
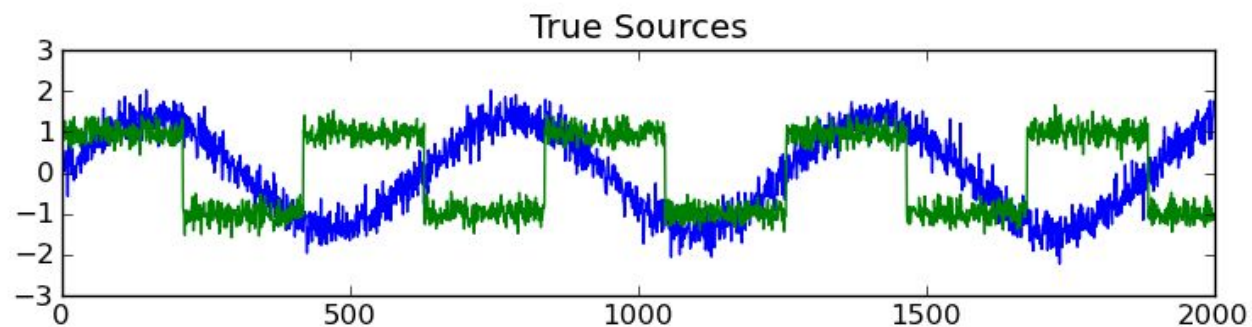
Las variantes del algoritmo FastICA no estiman las componentes independientes siguiendo algún criterio para su disposición. Sin embargo, posterior a la estimación existen varias estrategias a seguir para ordenar las componentes independientes y seleccionar así un número menor de estas para realizar la tarea de clasificación.

Una estrategia muy popular, y una de las que se va a utilizar en este trabajo, es privilegiar la selección de las componentes independientes que más variabilidad capturen de las variables originales.

**Algoritmos que estiman los
componentes independientes**

FastICA

Es un algoritmo computacionalmente eficiente que utiliza estimadores simples de la entropía negativa para la búsqueda de una matriz W que, al ser aplicada a las mezclas, maximiza esta propiedad en las componentes resultantes, permitiendo así la estimación de fuentes con distribuciones de probabilidad no-Gaussianas. Para trabajar, FastICA calcula las fuentes por deflación o por aproximación simétrica utilizando funciones no-lineales que el usuario puede elegir (e.g. \tanh o y^3).



Infomax

Este algoritmo encuentra la matriz W usando como criterio la minimización de la información mutua entre las fuentes estimadas, con lo que la entropía negativa conjunta se maximiza. Con este criterio, esta implementación hace posible el descomponer señales x en fuentes con distribuciones de probabilidad sub y súper-Gaussianas.

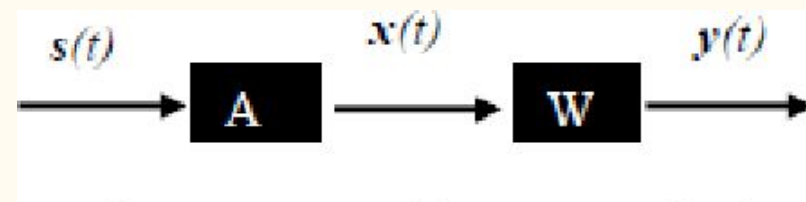
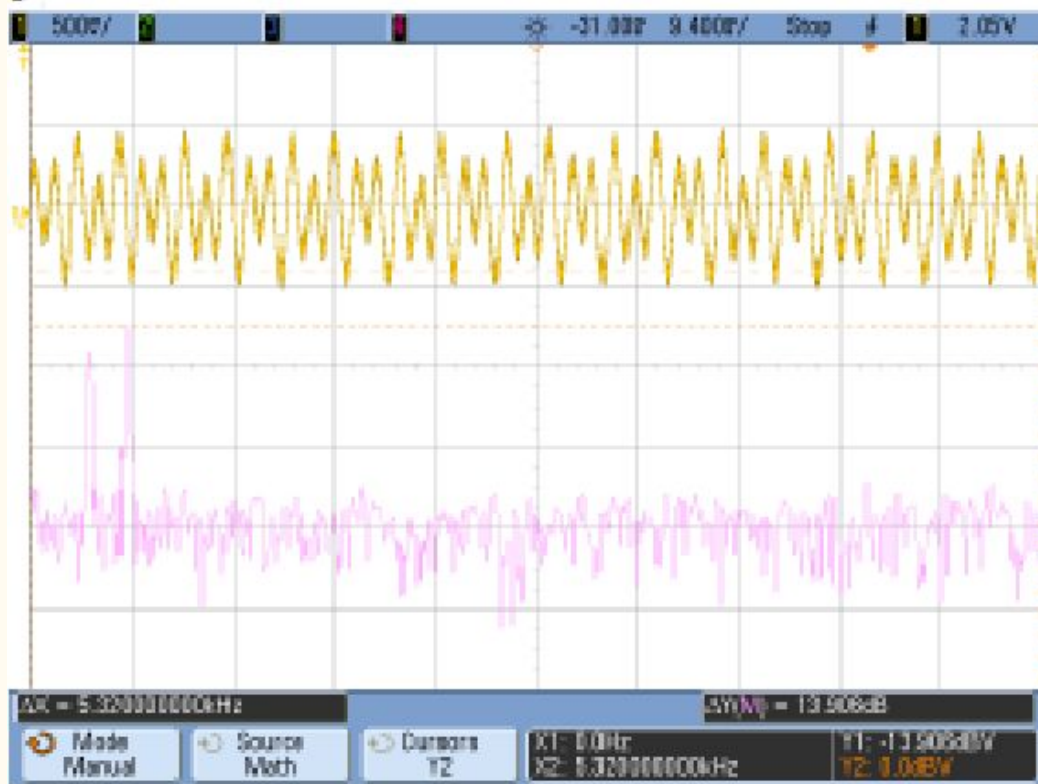
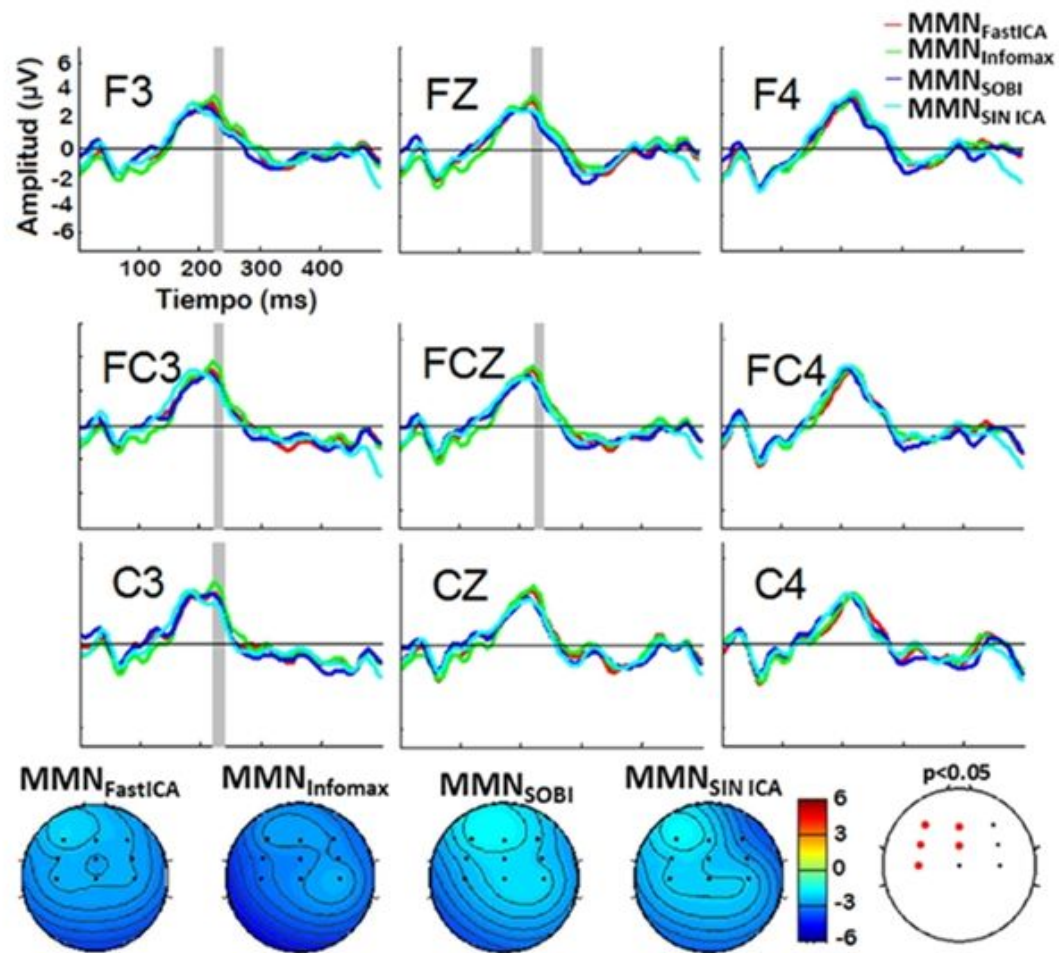


Fig. 6. Señales mezcladas representadas en tiempo y en frecuencia.

SOBI (Second Order Blind Separation)

Aunque éste no es estrictamente un algoritmo para el ACI, es una técnica de BBS muy utilizada, por eso se incluyó en este trabajo. Es una implementación que utiliza la ausencia de correlación temporal y espacial entre las fuentes como criterio para definir su independencia, por lo que explota la estructura temporal de las señales para calcular W . Para ello, el algoritmo trabaja sobre una pila de matrices de corrimiento (que construye a partir de las mezclas) y las diagonaliza simultáneamente mediante una matriz de transformación que resulta ser A , la matriz de mezcla e inversa de W .

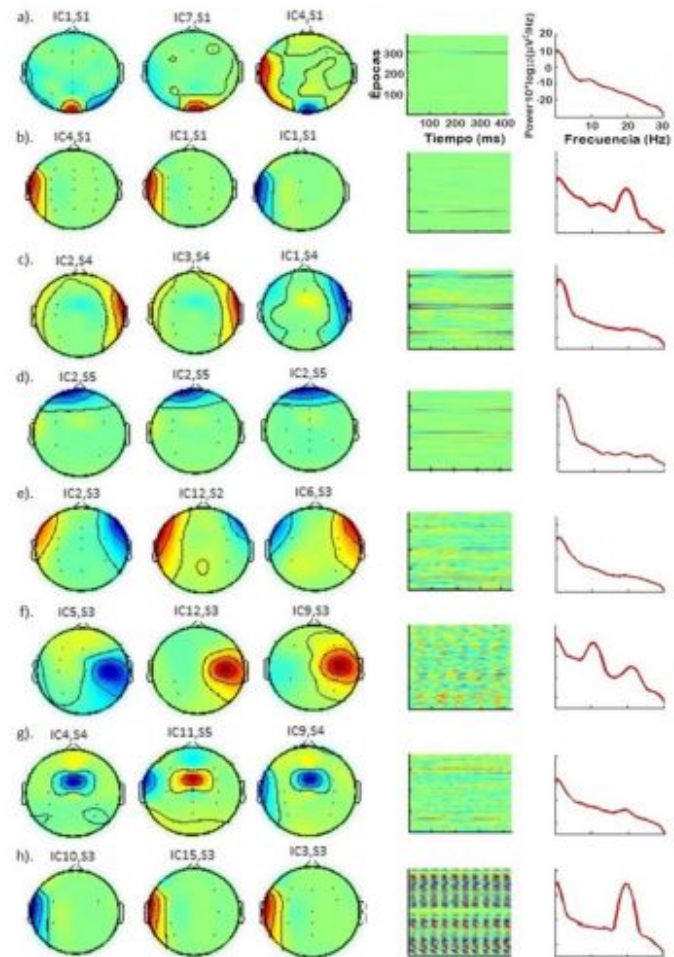


Aplicación del Independent Component Analysis (ICA)

ICA se puede ampliar para analizar señales no físicas. Por ejemplo, ICA se ha aplicado para descubrir temas de discusión en una bolsa de archivos de listas de noticias.

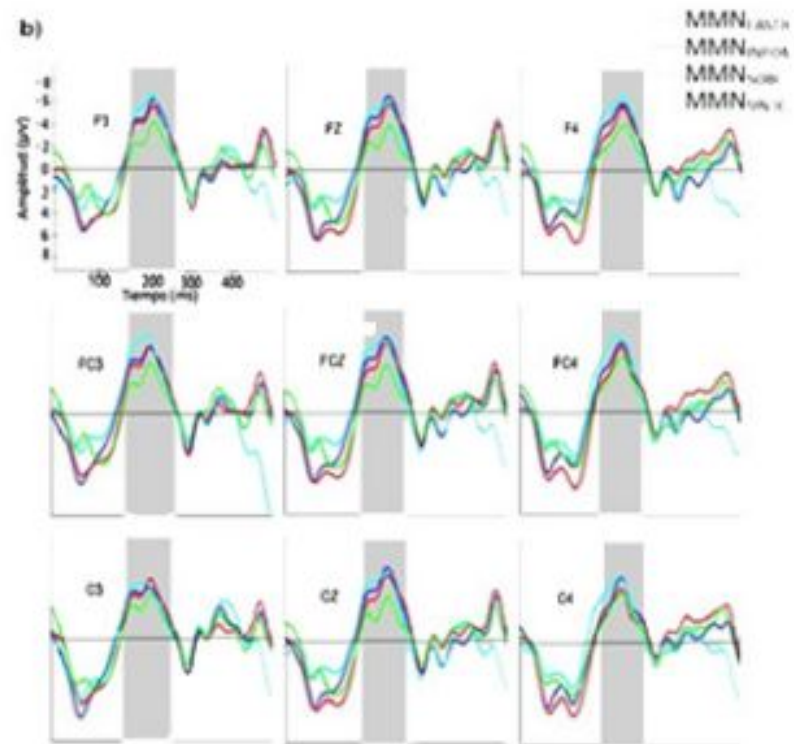
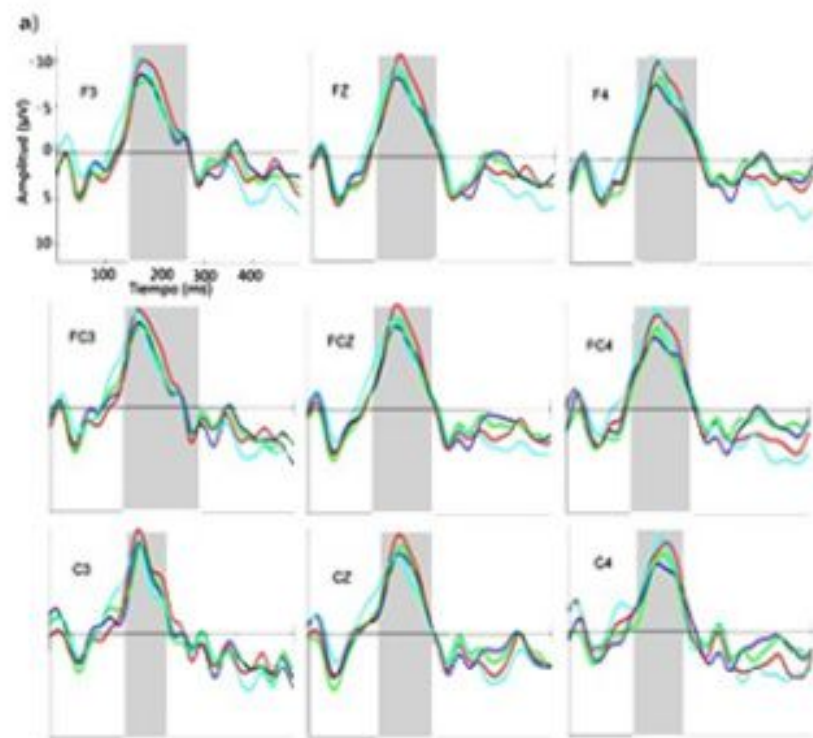
Algunas aplicaciones ICA se enumeran a continuación:

- Imágenes ópticas de neuronas
- Clasificación de picos neuronales
- Reconocimiento facial
- Modelado de campos receptivos de neuronas visuales primarias
- Predecir los precios del mercado de valores
- Comunicaciones por telefonía móvil
- Detección basada en el color de la madurez de los tomates
- Eliminar artefactos, como parpadeos, de los datos del EEG .
- Análisis de los cambios en la expresión génica a lo largo del tiempo en experimentos de secuenciación de ARN de una sola célula .
- Estudios de la red en estado de reposo del cerebro.
- Astronomía y cosmología



Aplicación del análisis de componentes independientes (ICA) en la interpretación de señales electroencefalográficas en epilepsia

El análisis de componentes independientes (ICA) es una técnica novedosa en el estudio de la señal electroencefalográfica. Los estudios previos son escasos y parciales. Los objetivos de este trabajo han sido determinar de forma cuantificada la validez de ICA en la eliminación de artefactos, comparar ICA y filtros digitales en la mejoría obtenida en la visualización del inicio de crisis, estudiar la fisiopatología de descargas epileptiformes intercríticas y crisis de distintos tipos de epilepsia y comprobar la capacidad de ICA para diferenciar entre epilepsias uni y multifocal es. Para ello se aplicó ICA a muestras de eeg intercrítico (muestras artificiales y muestras reales) Y crítico, y se reconstruyeron con los componentes de interés. Los resultados fueron valorados mediante técnicas de análisis frecuencial, proyección de topografías y cálculo de correlaciones, los resultados indican que ICA elimina de forma selectiva los artefactos en registros de EEG. Su eficacia depende de la amplitud y morfología del artefacto, del número de componentes calculados y el montaje utilizado. ICA obtiene mejores resultados que los filtros digitales para mejorar la visualización del inicio de las crisis epilépticas; la combinación de ambas técnicas fue la mejor opción. ICA descompone las grafoelementos intercríticos y las crisis en componentes con una secuencia temporal y topográfica. La separación de las descargas intercríticas varía dependiendo del tipo de epilepsia mientras que la de las crisis depende del tipo de patrón electroencefalográfico. ICA diferencia entre epilepsias uni y multifocales. Estos datos indican que ICA es una herramienta muy útil para el estudio de distintos aspectos del EEG.



GRACIAS!