



Universidad de Cantabria

Facultad de Ciencias

Máster en Física y Tecnología Física

Análisis de Espectros por
Resonancia Magnética usando
Métodos SVD

Autor: Alberto Peña Fernández

Director: Diego Herranz Muñoz

Fecha: 21 de Octubre de 2013

Resumen

La espectroscopía mediante resonancia magnética permite realizar espectros de tejidos y órganos del cuerpo humano que ayudan a esclarecer los componentes moleculares que los integran. La parte crucial del análisis de estos espectros consiste en aplicar un método que permita descifrar de forma precisa cada una de las componentes que forman parte del mismo, bien estudiando la señal en el dominio temporal o en el frecuencial.

Con este trabajo se pretende realizar una revisión de los métodos de análisis más importantes en ambos dominios, aplicarlos al estudio de un caso real y, si se obtienen los resultados esperados, presentar una nueva técnica mediante wavelets, usada ampliamente por grupos de otras áreas de la física, que se espera sea altamente competitiva respecto a las ya existentes en la identificación de señales débiles. Ésto supondría un avance en la detección de sustancias presentes en órganos y tejidos en bajas concentraciones, prácticamente indetectables con los métodos actuales, y abriría la posibilidad a que algunas de ellas puedan servir como biomarcadores de algún tipo de patología o lesión.

Específicamente, se va a intentar determinar un posible biomarcador para la Neurogénesis, que es un proceso que consiste en la generación de nuevas neuronas, aplicando los métodos sobre espectros de resonancia magnética del hipocampo cerebral para buscar algún pico de baja intensidad que pueda asociarse a alguna de las sustancias involucradas en dicho proceso.

Índice

1. Introducción	1
1.1. Espectroscopía por Resonancia Magnética	1
1.2. Análisis de datos MRS	5
1.2.1. Análisis en el dominio del tiempo	6
1.2.2. Análisis en el dominio de la Frecuencia	9
1.2.3. Otros métodos	11
1.2.4. Paquetes de Software	12
1.3. Caso Práctico: Neurogénesis	13
2. Método Experimental	16
2.1. Métodos SVD	17
2.1.1. LPSVD	20
2.1.2. HSVD	25
2.1.3. SELF-SVD	28
3. Discusión Y Resultados	36
3.1. Métodos SVD	36
3.1.1. Espectros simulados	36
3.1.2. Espectros reales	50
4. Conclusiones	56
5. Agradecimientos	60
6. Bibliografía	61
7. Anexo I: GUI para MATLAB	65

1. Introducción

1.1. Espectroscopía por Resonancia Magnética

La espectroscopía por Resonancia Magnética Nuclear (Nuclear Magnetic Resonance) (NMR) es una técnica que se emplea para obtener información sobre moléculas, como puede ser la elucidación de su estructura, cuantificar su presencia en un determinado entorno o estudiar sus comportamientos cinéticos y termodinámicos. Se usa en química inorgánica y orgánica, y en bioquímica. También se ha extendido su uso en medicina, donde incluso es posible conseguir imágenes. En casos como este último, cuando la espectroscopía se realiza *in vivo*, recibe el nombre de espectroscopía por resonancia magnética (Magnetic Resonance Spectroscopy) (MRS).

Se fundamenta en el hecho de que algunos núcleos atómicos sometidos a un campo magnético externo absorben radiación electromagnética en el rango de la radiofrecuencia. Como la frecuencia exacta de esta absorción depende del entorno de estos núcleos, se puede determinar la estructura de la molécula donde se encuentran éstos. Para que se pueda emplear esta técnica los núcleos atómicos deben tener un momento magnético distinto de cero, por lo que los núcleos con número atómico y número másico par no son adecuados para las observaciones MRS. Los núcleos más importantes para desarrollar este procedimiento son, 1H , ^{13}C y ^{31}P . Son preferibles los núcleos con número cuántico de espín nuclear igual a $1/2$, ya que carecen de un momento cuadrupolar eléctrico que produzca un ensanchamiento de las señales NMR. También es preferible que la abundancia del isótopo empleado sea lo mayor posible en la naturaleza, ya que la intensidad de señal que se obtiene depende de ésta. Debido a estas premisas, el núcleo atómico más usado es el 1H , dando lugar a la resonancia magnética nuclear de protón ($^1H - NMR$). A la hora de realizar espectros en química orgánica, el núcleo más usado es el ^{13}C , aunque se trata de un isótopo poco abundante y poco sensible.

La aplicación química de la NMR fue descubierta a principios de los años cincuenta del siglo pasado, cuando se observó que la frecuencia de resonancia de un núcleo dependía fuertemente de su entorno, lo que se conoce como desplazamiento químico, en inglés *chemical shift*. A partir de los años setenta, el desarrollo de nuevas técnicas y mayores campos magnéticos, lo que incrementa la resolución y sensibilidad de las señales, permitió estudiar moléculas cada vez mayores. Además, el descubrimiento de la NMR multidimensional

y el uso de los marcajes ^{13}C y ^{15}N , permitieron el nacimiento de la NMR biológica y posteriormente de la MRS.

Los espectrómetros actuales utilizan la técnica NMR de Transformada de Fourier (FT-NMR). La FT-NMR consiste en tener una muestra sometida a un campo magnético externo constante. A continuación se irradia la muestra con un pulso electromagnético de muy corta duración en la región de la radiofrecuencia. Estos pulsos suelen tener forma rectangular, lo que hace que la intensidad de la radiofrecuencia oscile entre un máximo y un mínimo que es constante mientras dura el pulso. La descomposición de Fourier de una onda rectangular contiene contribuciones de todas las frecuencias cercanas a la frecuencia central del pulso, por tanto el pulso que se genera es policromático y, cuanto más corto sea, será capaz de excitar un mayor rango de frecuencias, de acuerdo con el principio de incertidumbre de Heisenberg. Antes de la aplicación del pulso, el vector de polarización neta de cada uno de los espines nucleares se encuentra en situación de equilibrio alineado en la dirección del campo magnético. Durante el tiempo que se aplica el pulso, éste introduce un segundo campo magnético en una dirección perpendicular al campo principal del imán, lo que genera que el vector polarización realice un determinado movimiento de precesión. Cuando cesa el pulso, el vector polarización de todos los espines afectados formará un cierto ángulo con el eje del campo magnético principal. Entonces, como si fuesen pequeños imanes polarizados, los espines comienzan a precesionar con su frecuencia característica en torno al campo magnético externo lo que hace que se induzca una pequeña corriente oscilante de radiofrecuencia en una bobina receptora situada en las inmediaciones de la muestra. A medida que los núcleos van regresando a la situación inicial de equilibrio, alineados con el campo magnético principal, la señal detectada va disminuyendo en intensidad hasta hacerse cero. Esta caída de la señal se conoce como decaimiento de inducción libre (Free Induction Decay) (FID) y es el tipo de señal que se registra en cada realización. La FID es una onda que contiene todas las señales del espectro en una forma dependiente del tiempo. Ésta puede convertirse en un espectro de señales en función de su frecuencia, lo que se conoce como espectro de frecuencias, a partir de la Transformada de Fourier (Fourier Transform) (FT).

La posibilidad de excitar la muestra con uno o más pulsos de radiofrecuencia, cada uno de ellos aplicado con una potencia, duración, frecuencia, forma y fase particulares, e introducirlos en momentos específicos de tiempo

durante el experimento, generalmente antes de que el sistema haya regresado al equilibrio por relajación, permite diseñar toda una gama de secuencias de pulsos que permiten extraer información molecular muy variada. Ésto, permitió desarrollar la NMR multidimensional. Una secuencia de pulsos es una distribución en el tiempo de alguno o varios de los siguientes elementos: un cierto número de pulsos RF que afecten a uno o más tipos de núcleos; tiempos de espera en los que no se hace nada sino esperar a que el sistema evolucione de una determinada forma. Estos tiempos pueden ser fijos o bien incrementables, si su duración se va aumentando a medida que se repite el experimento; gradientes de campo magnético y una etapa final en la que se adquiere la FID. Todo esto permitió expandir la NMR a la bioquímica, donde se pudo determinar la estructura en disolución de biopolímeros como proteínas o incluso ácidos nucleicos de pequeño tamaño.

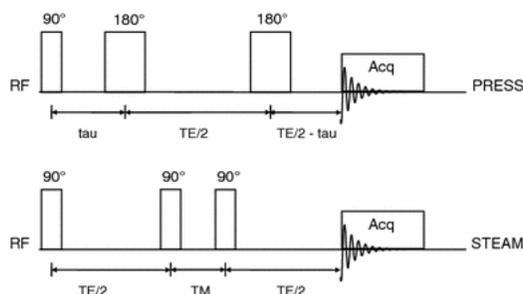


Figura 1: Representación esquemática de las secuencias de pulsos PRESS y STEAM [1].

En las últimas décadas, y gracias a todos estos avances, se comenzó a usar la MRS como técnica biofísica en biología y medicina. En medicina, la MRS permite realizar diferentes procesos, como la determinación *in vitro* de la estructura de proteínas y su interacción con medicamentos, la obtención de espectros o imágenes no-invasivas que permiten detectar metabolitos en diferentes partes del cuerpo o visualizar *in vivo* diferentes regiones de las mismas con gran contraste entre tejidos normales y afectados. Todo esto hace que la MRS proporcione valiosa información, tanto en el diagnóstico de enfermedades o lesiones, como en la monitorización terapéutica. Hoy en día, mediante esta técnica es posible obtener tanto espectros como imágenes, incluso una combinación de ambos. La diferencia en el proceder a la hora de realizar la adquisición de datos para conseguir un espectro o una imagen, radica en las secuencias de pulsos empleadas. Para que la MRS sea una herramienta de diagnóstico confiable, se han de poder realizar cuantificaciones de las señales

con confianza. Generar datos MRS de calidad varía en función del órgano estudiado, ya que los diferentes grosores y los movimientos pulsátiles de los mismos, llevan a distorsiones que incrementan la anchura de los picos de los metabolitos. La aplicación de la MRS en el cerebro es más simple debido a varias circunstancias. Por un lado, su perfil esférico y sus menores movimientos biológicos de tejidos lo hacen más fácil de focalizar y centrar. Por otro, la separación de los picos referentes a lípidos de los referentes a metabolitos está muy bien resuelta en el cerebro. Por tanto, las condiciones y los parámetros experimentales para generar datos MRS de calidad implican: una buena focalización, para evitar el efecto de la susceptibilidad magnética debido a cada órgano o tejido en diferentes personas; compensación de movimientos biológicos; calibración del ángulo del pulso; establecer un sistema de tiempos apropiado, basado en los tiempos de relajación longitudinal (T_1) y transversal (T_2), para el tiempo de *echo* (TE) y para el tiempo de repetición (TR). Esto se ha de realizar para cada tejido sobre el que se va a llevar a cabo un estudio. Las técnicas más aplicadas en medicina para realizar estudios *in vivo* son el $^1H - MRS$ y el $^{31}P - MRS$. El más ampliamente usado es el $^1H - MRS$, con las secuencias de pulsos PRESS (Point-RESolve Spectroscopy)[2] o STEAM (SimulaTed Echo Adquisition Mode)[3].

Las concentraciones de los metabolitos compuestos por diferentes sustancias se representan, bien como amplitudes de la FID en el dominio temporal, o bien como el área bajo los picos del espectro de potencias en el dominio de la frecuencia. Generalmente, diferentes técnicas MRS, así como diferentes secuencias de pulsos, se usan para detectar diferentes metabolitos. La amplitud del campo magnético juega un papel fundamental en la calidad de los datos. Los magnetómetros mayormente disponibles son los de 3 Teslas, 3T, haciendo referencia a la magnitud del campo magnético externo que emplean, y permiten obtener datos adecuados para su uso clínico. Asimismo, la resolución espacial de los espectros aumenta al incrementar la intensidad del campo magnético externo. La relación señal-ruido de los metabolitos se incrementa aproximadamente de forma lineal con la intensidad del campo magnético, por lo que una mayor intensidad de campo permite tener que realizar menos escáneres y, por tanto, ahorrar tiempo clínico. Hay que tener en cuenta que el aumento de la magnitud del campo magnético externo hace variar los tiempos de relajación de los metabolitos, alargándose T_1 y disminuyendo T_2 , a la vez que se incrementan los efectos de la susceptibilidad. Actualmente existen espectrómetros MRS de 7T [1,4].

1.2. Análisis de datos MRS

El análisis de datos MRS se divide en dos etapas: pre-procesado y cuantificación. En el pre-procesado se intentan eliminar las imperfecciones que presentan los datos debido a motivos experimentales como, imperfecciones en el receptor que impidan recolectar datos exactamente al comienzo de la FID o del centro exacto del *echo*, movimientos fisiológicos de las diferentes partes del cuerpo bajo estudio, señales de decaimiento rápido procedentes de componentes inmóviles, superposición o solapamiento de señales de componentes relativamente móviles, truncamiento de datos antes de que la FID decaiga al nivel del ruido o presencia de picos residuales del agua. Las técnicas de pre-procesado son diferentes dependiendo de la naturaleza del proceso de obtención de los datos MRS y con ellas se pretende mejorar tanto su visualización como permitir una estimación más precisa de sus parámetros. En este trabajo no se va a profundizar en estas técnicas, ya que no se han empleado en el mismo. Una descripción de las mismas puede encontrarse en [2,6,7].

Por otro lado, en la cuantificación se trata de caracterizar las diferentes componentes que forman la señal para obtener toda la información física posible sobre cada una de ellas.

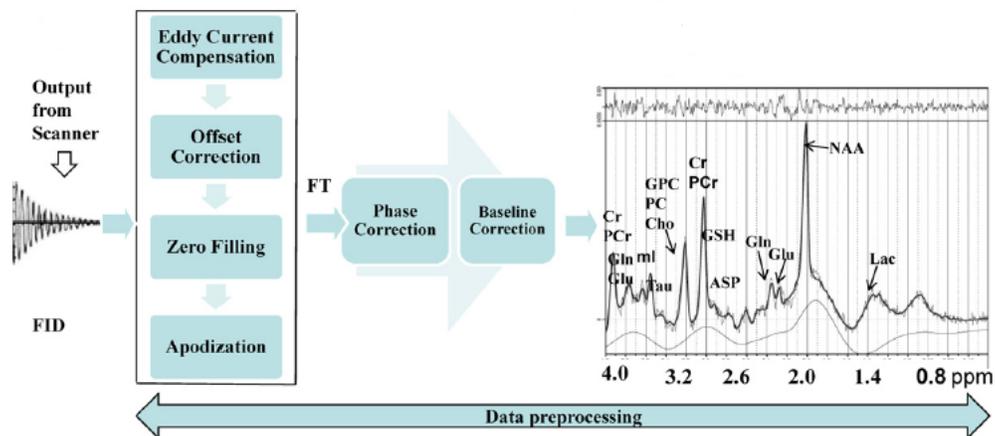


Figura 2: Esquema del proceso a seguir para procesar datos MRS [2].

Ambas etapas pueden llevarse a cabo tanto en el dominio del tiempo como en el de la frecuencia. En el dominio del tiempo la concentración de la

componente de un metabolito es proporcional a la amplitud de la señal. En el dominio de la frecuencia es proporcional al área bajo el pico del metabolito. Si la señal medida es ideal, con una alta tasa señal-ruido y sin truncamientos ni distorsiones de base, tanto el análisis temporal como el frecuencial son equivalentes, dando los mismos resultados y con los mismos márgenes de error. Cualquier imperfección en la señal adquirida, debido a condiciones experimentales, puede generar diferencias en las concentraciones medidas en uno y otro dominio y favorecer la cuantificación en un dominio sobre otro. En el dominio del tiempo es posible seleccionar un rango particular de buenos datos para el análisis, mientras que en el dominio de la frecuencia la transformada de la Fourier se lleva a cabo sobre todo el conjunto de datos, lo que impide seleccionar los mejores datos [4,5,6].

1.2.1. Análisis en el dominio del tiempo

Al dominio del tiempo se le denomina dominio de medida, ya que es en él donde se muestra la verdadera representación de los datos obtenidos experimentalmente. La señal en este dominio es, como ya se ha comentado anteriormente, una FID. La adquisición de datos experimentales en los espectrómetros actuales para generar la FID se lleva a cabo de forma cuadrática, lo que hace que la señal adquirida presente dos partes, referidas como parte real e imaginaria debido a que se hace uso de los números complejos para describir matemáticamente la señal, que se corresponden con las dos componentes del vector rotacional de magnetización, \vec{M} . Dado que \vec{M} rota alrededor del eje z, las partes real e imaginaria de la FID están relacionadas entre ellas como seno y coseno. Tanto a partir de la parte real como de la imaginaria se pueden obtener los parámetros que caracterizan la señal. Una FID, entonces, consiste en una combinación de diferentes sinusoides complejas exponencialmente amortiguadas, cada una de ellas única y caracterizada por ciertos parámetros físicos: la frecuencia de resonancia, que sirve para identificar el metabolito; el factor de amortiguación, que caracteriza la movilidad del mismo; la fase inicial, que está relacionada con la posición del vector magnetización nuclear al inicio de la adquisición y que es equivalente al desplazamiento de la FID con el eje del tiempo; y la amplitud para $t = 0s$, que es directamente proporcional a la concentración del metabolito. El factor de amortiguación, α , es inversamente proporcional al tiempo de relajación transversal aparente, T_2^* , mediante la relación $\alpha = 1/T_2^*$, que depende del T_2 de la componente y de las inhomogeneidades del campo externo. Si la

muestra o tejido sobre la que se aplica la MRS presenta más de una señal, la FID resultante es la suma de las FID's individuales [4,5].

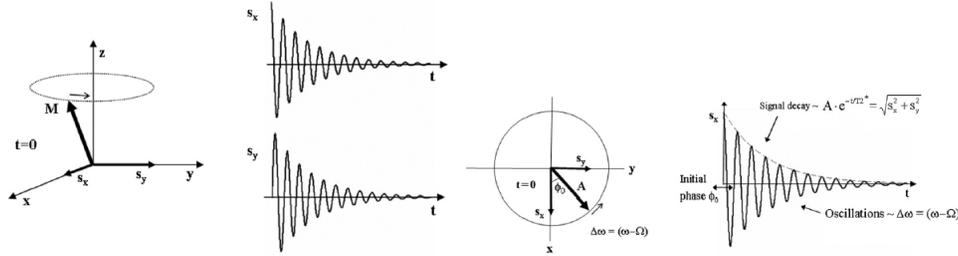


Figura 3: Representación gráfica de la FID y sus componentes [7].

La cuantificación precisa y eficiente de estas señales se consigue con unas funciones de modelado adecuadas. Por ello, la selección de una función de modelado adecuada es un paso crucial para estimar los parámetros intrínsecos a la señal y convertir estos parámetros estimados en cantidades bioquímicas. Teóricamente, la señal en el dominio del tiempo puede modelarse como la suma de sinusoides complejas exponencialmente amortiguadas, donde K , que es el número esperado de sinusoides que forman la muestra de N datos exponenciales, se supone conocida:

$$y(t_n)_{n=0,1,\dots,N-1} = \sum_{k=1}^K A_k e^{i\phi_k} d(t_n) e^{i2\pi\nu_k t_n} + e(t_n) \quad (1)$$

donde $d(t_n)$ es la función de decaimiento exponencial que gobierna la FID, que normalemente viene descrita por una de estas tres funciones:

$$d(t_n) = e^{-\alpha_k t_n} \quad \text{Perfil Lorentziano}$$

$$d(t_n) = e^{-\beta_k t_n^2} \quad \text{Perfil Gaussiano}$$

$$d(t_n) = e^{-\alpha_k t_n - \beta_k t_n^2} \quad \text{Perfil de Voigt}$$

Los parámetros de la función de modelado son la amplitud, A_k , que es proporcional al número de núcleos que contribuyen a la componente espectral de frecuencia ν_k ; $i = \sqrt{-1}$, $t_n = t_0 + n\Delta t$, donde Δt es el intervalo de muestreo. Los factores de amortiguación, α_k para el perfil Lorentziano y β_k

para el Gaussiano, proporcionan información sobre la movilidad y el ambiente macromolecular de los núcleos. El término $2\pi\nu_k t_n$ corresponde al primer orden de la fase y ϕ_k representa el orden cero de la misma. El término de ruido, $e(t_n)$, se supone ruido blanco Gaussiano complejo y circular [5].

El análisis de los datos comprende el pre-procesado de los datos, seguido de una cuantificación de la FID. La cuantificación consiste en determinar que metabolitos hay presentes en la FID de una región determinada de un tejido blando y en que cantidad, es decir, cuál es la concentración de los mismos. Ésto se consigue mediante la caracterización de la señal a través de una función de modelado definida en el dominio del tiempo. Los métodos para realizar la cuantificación se dividen en dos grupos: métodos iterativos y métodos no-iterativos o de *caja negra*[6].

Los métodos iterativos, típicamente basados en mínimos cuadrados no-lineales, se basan en la minimización de las diferencias entre los datos experimentales y el modelo, empleando una optimización local o global. Permiten la inclusión de información previa sobre el experimento que se esté llevando a cabo, como relaciones entre las componentes espectrales. Este conocimiento previo juega un papel muy importante, a la hora de resolver solapamientos de picos o imponer anchuras en un espectro ruidoso, para mejorar la precisión de las estimaciones. Uno de los primeros métodos de este tipo fue VARPRO (VARiable PROjection)[8], donde se asume un perfil Lorentziano para las componentes individuales de las resonancias y se ajustan las exponenciales de decaimiento en el dominio del tiempo usando el algoritmo de optimización local Osborne's Levenberg-Marquardt. Otro método es el AMARES (Advanced Method for Accurate, Robust and Efficient Spectral)[9], que permite añadir al ajuste información sobre los parámetros espectrales y así mejorar su eficiencia, las tasas de convergencia y la precisión total. Además, está pre-programado para cambiar entre los perfiles Lorentziano, Gaussiano y de Voigt, y permite ajustar *echoes* en lugar de FID's. También existe QUEST (Quantitation based on Quantum ESTimation)[10], que usa una base de metabolitos generada a partir de espectros simulados e *in vitro*, para realizar la cuantificación.

Los métodos no-interactivos se basan, bien en el principio de predicciones lineales (LP) o en la teoría del espacio de estados. Son menos flexibles que los iterativos y menos útiles en espectros complicados. La función de modelado

sólo puede presentar un perfil Lorentziano y los parámetros del modelo se estiman en una única iteración. Algunos de estos métodos son, el HSVD (Hankel Singular Value Decomposition)[11], HTLS (Hankel lanczos Total Least Squares)[12] o LPSVD (Linear Prediction Singular Value Decomposition)[13,14]. Mediante el modelado de la señal se obtiene la contribución de cada una de las señales independientes que la forman. Normalmente, el límite inferior de Cramèr-Rao (CRLB)[15] proporciona una estimación del error del ajuste o de la incertidumbre estadística de la estimación de la concentración del metabolito.

1.2.2. Análisis en el dominio de la Frecuencia

La inspección visual de una FID multicomponente en el dominio temporal no es propicio. La solución es mostrar las componentes individuales de la señal en función de la frecuencia. Para poder trabajar con los datos en el dominio de la frecuencia se ha de transformar la FID obtenida experimentalmente en un espectro de potencias. Matemáticamente, ésto se consigue mediante la FT, que es una transformación lineal que descompone la señal en componentes individuales oscilantes, mostrando cada una de ellas como un pico de una determinada amplitud y posición en el eje frecuencial. Al igual que la FID, el espectro también posee parte real e imaginaria, que se corresponden con las líneas de absorción y dispersión respectivamente, y de nuevo se puede realizar un estudio de la señal con ambas partes independientemente. Ya que la parte real, o líneas de absorción, presenta intensidad únicamente positiva y colas más estrechas, el análisis suele llevarse a cabo con estos datos. En los espectros, el área bajo el pico es proporcional a la amplitud de la componente correspondiente, la posición en el eje de frecuencias se corresponde con la frecuencia de resonancia de la sustancia, y la anchura a media altura es inversamente proporcional a T_2^* . Puede demostrarse que el perfil del pico depende de la función de decaimiento, o envolvente, de la FID. Una función de decaimiento exponencial lleva a picos Lorentzianos. Sin embargo, si el campo magnético estático presenta inhomogeneidades, el perfil sufre alteraciones. Tan sólo en el caso especial de que dichas inhomogeneidades se distribuyan de manera Lorentziana, los picos permanecen con este perfil, aunque con anchuras mayores. La altura del pico depende tanto de A como de T_2^* , pero el área bajo éste sólo depende de A [3,4,5].

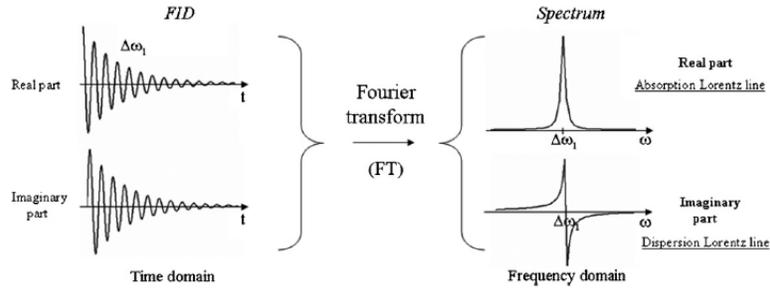


Figura 4: Procedimiento para transformar la señal del dominio del tiempo al de la frecuencia [7].

Al realizar esta transformación pueden aparecer distorsiones debido al efecto de inducción en el receptor en el punto inicial o por el truncamiento de la FID. Esto genera una base móvil. La contribución de componentes inmóviles que presentan un T_2^* más corto, genera un ensanchamiento de los picos. Esto da lugar a una línea de base macromolecular que hace necesaria su eliminación o mitigación mediante técnicas de pre-procesado.

La cuantificación se lleva a cabo típicamente de dos formas, bien integrando el área bajo los picos o bien realizando ajustes no-lineales por mínimos cuadrados usando funciones de modelado. En estudios *in vivo*, debido a la complejidad del espectro y el solapamiento de componentes, la cuantificación se realiza mediante algoritmos de optimización basados en modelos. Éstos se basan en el método no-lineal de mínimos cuadrados regularizado. Idealmente, las líneas espectrales son Lorentzianas, pero en los datos experimentales reales las líneas presentan una mezcla de perfiles, Lorentzianos, Gaussianos y otros perfiles adicionales. Una ecuación general sería del tipo:

$$Y(\nu) = \sum_{n=1}^N A_{1n} L_n(\theta_n, \nu) + \sum_{n=1}^N A_{2n} G_n(\theta_n, \nu) + \sum_{n=1}^N B_n \nu^n + W(\nu) \quad (2)$$

donde A_{1n} y A_{2n} son los factores de peso de la Lorentziana, L_n , y la Gaussiana, G_n , respectivamente. $W(\nu)$ es el ruido blanco Gaussiano. A_1 y A_2 pueden estimarse o fijarse a través de información previa. El tercer término da cuenta de posibles distorsiones de base, usando un polinomio de orden N con coeficiente B_n para modelar la base. La estimación de los parámetros del modelo, θ_n , se obtiene resolviendo el problema clásico de los mínimos cuadrados no-lineales[5]. Existen varios métodos no-iterativos como el FDM

(Filter Diagonalization Method)[16], una versión modificada de MODE [17], FIDO (Filtering and DOWsampling)[18] o el método SELF-SVD (SElective Frequency band Singular Value Decomposition)[19]. Realizar métodos iterativos en este dominio es más complicado debido a que no existen expresiones analíticas en este dominio de las funciones que se emplean. El más conocido es el LCMModel [20] que hace uso únicamente de la parte real de los datos.

1.2.3. Otros métodos

El análisis de datos MRS en el dominio del tiempo presenta la ventaja de manejar fácilmente la falta de algunos datos y/o eliminar datos iniciales debidos a macromoléculas. El análisis en frecuencia presenta la ventaja de poder seleccionar únicamente ciertas frecuencias para su análisis y así reducir el número de parámetros del modelo. La combinación de ambos dominios permite usar modelos en el dominio del tiempo, cuyos parámetros son determinados ajustando una parte concreta de la FT de los datos. Estos métodos permiten manejar fácilmente modelos complejos, casos de desaparición de datos y llevar a cabo análisis selectivos en frecuencia [21].

En los últimos años se han desarrollado métodos de análisis basados en wavelets [22,23,24]. Esta forma de análisis parece adecuada para este tipo de datos, ya que permite realizar el mismo simultáneamente en los dominios de tiempo y escala, pudiendo proporcionar información de interés. Primero se desarrollaron métodos basados tanto en transformadas de wavelet discretas como continuas para obtener estimaciones de los parámetros de las diferentes componentes que forman la señal FID. La transformada de Wavelet discreta se presenta más adecuada para compresión de datos y síntesis de señales, mientras que la transformada continua, debido a su redundancia, parece más adecuada para el análisis y determinación de características. Básicamente, cada señal se obtiene a partir de dilataciones y traslaciones de una wavelet madre. En la mayoría de estudios al respecto, se utilizaba como wavelet madre la wavelet de Morlet. Se ha probado la efectividad de dichos métodos, aunque siguen sufriendo las mismas limitaciones que los métodos temporales o frecuenciales descritos anteriormente.

Sin embargo, ha aparecido desde hace un par de años un nuevo método basado en wavelets que parece muy prometedor [25]. Consiste en crear wavelets específicas, denominadas wavelets de autocorrelación, para el estudio individual de metabolitos usando la transformada de wavelet continua.

Reciben este nombre debido a que se construyen a partir de funciones normalizadas de autocorrelación de datos de metabolitos almacenados en bases de datos. Según las primeras pruebas llevadas a cabo, éstas son capaces de determinar sin ambigüedad la presencia de un metabolito específico en una señal con presencia de otras componentes, incluso en superposición, y determinar algunos de sus parámetros. Ésto permitiría determinar si el metabolito deseado está presente o no en el espectro bajo inspección sin necesidad de realizar un estudio completo del mismo, lo que ahorraría mucho tiempo en la cuantificación de biomarcadores para su uso clínico. Desafortunadamente, se necesita poseer una base de datos de los parámetros de la señal asociada al metabolito, con lo que este método no evita el tener que usar previamente alguno de los métodos descritos anteriormente cuando el objetivo es encontrar una señal asociada a un nuevo proceso.

1.2.4. Paquetes de Software

Cada tipo de máquina de resonancia magnética proporciona un software básico para el procesamiento de datos MRS y, en la mayoría de los casos, los investigadores crean sus propios programas o compran software más sofisticado para obtener mayor flexibilidad y profundidad de análisis. Los dos softwares más usados son el LCMoDel[20,26] y jMRUI[27,28]. En el Laboratorio de Neuroimagen del Hospital Universitario Marques de Valdecilla se dispone de jMRUI.

jMRUI es un software basado en Java a través de un interfaz gráfico, que analiza datos MRS en el dominio del tiempo y funciona en los principales sistemas operativos (Windows, Linux, MacOS). Funciona a través del Java Development Kit (JDK), que se usa como interfaz entre el código nativo, escrito en FORTRAN y ANSI C, y el Graphic User Interface (GUI). Este software permite trabajar con uno o varios voxel y manejar gran cantidad de datos en el dominio del tiempo para analizar datos MRS tanto en forma de espectro como de imagen. Además, las señales de los metabolitos se pueden computar mediante el formalismo productor-operador, inspirado en la mecánica cuántica, aplicado a la información NMR. Incorpora la mayoría de técnicas de pre-procesado y cuantificación existentes, algunas de las cuales se han comentado en los apartados anteriores [27,28,29].

1.3. Caso Práctico: Neurogénesis

La neurogénesis es el proceso mediante el cual células progenitoras (Neural Progenitor Cells) (NPC's) se especializan para generar nuevas neuronas. Se ha descubierto que el cerebro de los mamíferos adultos aún retiene la habilidad para generar nuevas neuronas mediante tallos neuronales y NPC's que residen en la zona del hipocampo y la zona subventricular. Las NPC's poseen la habilidad de autorenovarse y generar descendencia que da lugar a células madres de diferentes tipos. Esta habilidad de las NPC's de producir neuronas, astrocitos y oligodendrocitos, abre la posibilidad de su uso para reparar tejido nervioso dañado o perdido, y tratar enfermedades o traumas neurológicos. Por tanto, la identificación de un biomarcador que permita determinar y cuantificar la presencia de este tipo de células *in vivo* en el cerebro, puede tener implicaciones muy importantes en el diagnóstico, pronóstico y terapias para un amplio rango de enfermedades cerebrales en humanos.

Hasta hace unos años se han usado técnicas como la tomografía de emisión positrónica o el escaneado tomográfico mediante la computación de fotón único, para intentar proceder a la identificación y seguimiento *in vivo* de NPC's. Pero dichas técnicas requieren que las NPC's sean etiquetadas con agentes radiactivos o derivados óxido-base de hierro supramagnético *ex vivo*, por lo que no permiten detectar NPC's propias del cerebro. La $^1H - NMR$ ha sido usada ampliamente para la detección *in vitro* de pequeñas cantidades de metabolitos conocidos y la identificación de compuestos desconocidos presentes en fluidos y tejidos. También ha sido capaz de detectar e identificar metabolitos específicos para neuronas, como el N-Acetil Aspartato (NAA), o neuroglías, como la colina (Cho) o el Myoinositol (mI), que han sido empleadas como biomarcadores útiles para los correspondientes tipos celulares en muestras aisladas de tejidos. Este tipo de espectroscopia no puede usarse en organismos vivos. Para este menester se usa su correlacionado $^1H - MRS$ que proporciona información sobre el estado de los metabolitos en tejido vivo. Ambas técnicas se complementan, ya que la $^1H - NMR$ se usa para buscar metabolitos específicos *in vitro* y $^1H - MRS$ para detectarlos *in vivo*.

En [30], se aseguraba haber descubierto un biomarcador metabólico para la neurogénesis que, mediante la identificación y caracterización de su señal en espectros $^1H - MRS$, permite detectar y cuantificar NPC's *in vivo* en el cerebro humano y, por tanto, monitorizar la neurogénesis. A pesar del amplio estudio llevado a cabo, la publicación tuvo varios artículos de crítica

[31-34], cuestionando la fiabilidad, reproducibilidad y viabilidad del mismo. Se cuestionaba la fiabilidad del análisis espectral realizado, así como la interpretación de los resultados.

Se exponía que el método SVD selectivo en frecuencia usado para la cuantificación no era adecuado al no ser capaz de resolver superposición de componentes. Si el decaimiento de las señales no es exponencial, las componentes producidas por el método SVD pueden no tener correspondencia en el mundo real, y no ser aceptadas o interpretadas como válidas. Si un método SVD trata de ajustar una señal no ideal mediante un número de sinusoides exponencialmente amortiguadas, éste modifica sus parámetros de cualquier forma necesaria para minimizar las diferencias entre la señal y el ajuste. Por otro lado, la presencia de acoplamientos J , solapamientos o ruido, también pueden afectar la interpretación de los resultados, y en este caso la región frecuencial en la que se ubica el biomarcador es una zona en la que aparecen picos en el espectro debidos a lípidos. En estos casos, los investigadores críticos con el estudio aseguran que la información sobre la fase es de vital importancia para evaluar si una componente es válida. Dos componentes espectrales adyacentes no deberían diferir significativamente en sus fases, debido a los errores de orden cero y de primer orden de la fase, y a que la evolución de la fase por el acoplamiento J para tiempos cortos es pequeña. Si es muy grande, se producen errores a la hora de estimar los parámetros de las señales. En el estudio, se establece la detección del biomarcador en base únicamente a la frecuencia espectral de la componente, desestimando su fase y la de las componentes adyacentes. Por otro lado, también se discutía que otros grupos de investigación habían llevado a cabo similares estudios de cultivos de NPC's *in vitro* sin detectar la presencia de la componente resonante a 1.28 ppm.

Hasta ahora, nadie ha conseguido reproducir los resultados expuestos en [30]. Si esto fuera posible, las implicaciones médicas serían importantísimas. Gracias a la colaboración del IFIMAV, se dispone de una muestra de datos de pacientes reales (anónimos) sobre la que intentar repetir el experimento. En ellos, se va a buscar de forma explícita complicaciones de tipo sistemático o metodológico en la implementación de los métodos SVD que pudieran afectar a la calidad de los resultados.

El objetivo de este trabajo es reproducir la última parte del estudio descrito anteriormente, para intentar confirmar o refutar su veracidad. Con la

ayuda de MATLAB, se van a implementar tres métodos SVD, dos en el dominio del tiempo y uno selectivo en frecuencia, para analizar, por un lado espectros simulados, para evaluar el funcionamiento de los mismos, y por otro espectros reales de hipocampos de cerebros humanos para intentar hallar la presencia de la señal resonante asociada a la neurogénesis. Por último, si se consigue hallar la evidencia de la presencia de esta señal en los espectros, el objetivo es crear una wavelet de autocorrelación específica para dicho metabolito y testar su viabilidad.

2. Método Experimental

Como se ha comentado al final de la introducción, el procedimiento experimental que se va a seguir consta de dos partes bien diferenciadas. Primero, se va a proceder a implementar y testar el funcionamiento de tres métodos SVD, dos en el dominio del tiempo y otro en el de la frecuencia, para determinar las habilidades de éstos a la hora de analizar diferentes situaciones que se pueden dar en los espectros MRS *in vivo* y usarlos posteriormente como métodos de cuantificación de los espectros de hipocampos. Después, si se consigue hallar la presencia del biomarcador y se puede realizar una caracterización adecuada del mismo, se procederá a construir una wavelet de autocorrelación específica para la búsqueda directa del mismo en los espectros.

Los espectros del hipocampo de cerebros humanos, en concreto del hipocampo izquierdo, han sido proporcionados por el Laboratorio de Neuroimagen de Valdecilla. En este laboratorio se dispone de un aparato de resonancia magnética del tipo *Philips Achieva 3T*. Como su propio nombre indica, se trata de un aparato de resonancia magnética cuyo campo externo fijo es de 3T.



Figura 5: Equipo de resonancia magnética utilizado para adquirir los espectros. Imagen tomada de www.healthcare.phillips.com

Presenta una frecuencia de transmisor de $f_{trans} = 1,2778 \cdot 10^8 \text{Hz}$, dato necesario para poder transformar los Hertzios del eje frecuencial a una escala de partes por millón (ppm), la cual permite que la posición de los picos del espectro se mantenga invariante aunque los espectros se realicen con diferentes aparatos de resonancia. Esta transformación se consigue mediante la relación:

$$f \text{ (ppm)} = \frac{f \text{ (Hz)} - f_{ref} \text{ (Hz)}}{f_{trans} \text{ (Hz)}} 10^6 \quad (3)$$

donde en el numerador aparece la diferencia de frecuencias entre la frecuencia de la componente, f , y la frecuencia de la componente de referencia con la que se ha adquirido el espectro, f_{ref} . Ésto, se divide entre la frecuencia de operación de transmisión del aparato, f_{trans} . De esta manera se consigue que, por ejemplo, si una componente presenta un pico de resonancia tomada con un 3T en 1,10ppm, ésta seguirá apareciendo ahí al realizar esta transformación aunque la señal se adquiera con una resonancia de 1T, 1,5T, 2T ó 4T, mientras se mantenga la misma frecuencia de referencia. Para adquirir los espectros se ha seguido el mismo procedimiento que el realizado en [30]. Se ha empleado un único voxel centrado en el hipocampo izquierdo de sujetos de prueba, usándose una secuencia de pulsos PRESS.

Para realizar estudios del comportamiento de los métodos a la hora de analizar los espectros, se han empleado espectros simulados. Los datos simulados se han realizado siguiendo la estructura de espectros reales del hipocampo. Debido a ello, la componente que se simula como biomarcador de la neurogénesis se genera con una amplitud pequeña, escasamente superior al nivel del ruido, para mantener las simulaciones lo más fieles posibles a la realidad. El ruido se simula, bien a partir de espectros a los que se les han eliminado las componentes de metabolitos, o bien aleatoriamente, definiendo y controlando su máxima amplitud.

Los diferentes métodos de cuantificación que se han desarrollado para analizar los espectros se han implementado usando MATLAB. Esta herramienta trae ya implementados muchos métodos algebraicos necesarios para que las rutinas de cuantificación funcionen correctamente y de esta forma se ha podido hacer uso de ellas sin necesidad de implementarlas de nuevo.

A continuación se va a detallar cada uno de los métodos de cuantificación que se ha desarrollado y empleado en la caracterización de los espectros para buscar el biomarcador asociado a la neurogénesis.

2.1. Métodos SVD

Los métodos que se van a describir a continuación, emplean la descomposición en valores singulares (Singular Value Decomposition) (SVD) de la

matriz de datos para analizar la señal trabajando en el espacio de estados. Dicha descomposición genera un subespacio *señal* y un subespacio *ruido* con los que se trabaja para recuperar la información física que proporcionan las diferentes componentes constituyentes de la señal. En general, el método SVD puede definirse como se detalla a continuación.

Para cualquier matriz, \mathbf{A} de dimensión $(m \times n)$ existen las matrices unitarias $\mathbf{U} \in \mathbf{C}^{m \times m}$, $\mathbf{V} \in \mathbf{C}^{n \times n}$ y una matriz diagonal $\Sigma \in \mathfrak{R}^{m \times m}$ con elementos no negativos en la diagonal, tales que:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^* \quad (4)$$

Mediante las permutaciones adecuadas, los elementos diagonales de Σ se pueden reordenar de forma decreciente:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)}$$

Esta factorización se denomina factorización en valores singulares de \mathbf{A} y su existencia es un resultado significativo, tanto a nivel teórico como práctico. Las matrices \mathbf{U} , \mathbf{V} y Σ , satisfacen:

$$\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I} \quad (m \times m)$$

$$\mathbf{V}^*\mathbf{V} = \mathbf{V}\mathbf{V}^* = \mathbf{I} \quad (n \times n)$$

$$\Sigma_{ij} = \begin{cases} \sigma_i \geq 0 & \text{para } i = j \\ 0 & \text{para } i \neq j \end{cases}$$

La siguiente terminología va asociada con el método SVD:

- Los vectores singulares izquierdos de \mathbf{A} son las columnas de \mathbf{U} . Asimismo, son los autovectores de $\mathbf{A}\mathbf{A}^*$
- Los vectores singulares derechos de \mathbf{A} son las columnas de \mathbf{V} . Asimismo, son los autovectores de $\mathbf{A}^*\mathbf{A}$
- Los valores singulares de \mathbf{A} son los elementos diagonales $\{\sigma_i\}$ de Σ . σ_i son las raíces cuadradas de los $\min\{m, n\}$ autovalores mayores de $\mathbf{A}^*\mathbf{A}$ o $\mathbf{A}\mathbf{A}^*$

- El triple singular de \mathbf{A} es el triplete (σ_k, u_k, v_k) donde u_k (v_k) es la columna k -ésima de \mathbf{U} (\mathbf{V})

Otra propiedad interesante es que si se tiene que $\text{rango}(\mathbf{A}) = r \leq \min(m, n)$, puede demostrarse que:

$$\begin{cases} \sigma_k > 0 & \text{si } k = 1, \dots, r \\ \sigma_k = 0 & \text{si } k = r + 1, \dots, \min(m, n) \end{cases}$$

Entonces, para una matriz de rango r , el método SVD puede escribirse como:

$$\mathbf{A} = [\mathbf{U}_1 \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^* \\ \mathbf{V}_2^* \end{bmatrix} = \mathbf{U}_1 \Sigma_1 \mathbf{V}_1^*$$

donde $\Sigma_1 \in \mathfrak{R}^{r \times r}$ es no singular. Esto lleva asociadas una serie de importantes consecuencias:

- \mathbf{U}_1 es una base ortonormal de $\mathfrak{R}(\mathbf{A})$
- \mathbf{U}_2 es una base ortonormal de $\mathbf{N}(\mathbf{A}^*)$
- \mathbf{V}_1 es una base ortonormal de $\mathfrak{R}(\mathbf{A}^*)$
- \mathbf{V}_2 es una base ortonormal de $\mathbf{N}(\mathbf{A})$

Un corolario de lo anterior es que para cualquier matriz $\mathbf{A} \in \mathbf{C}^{m \times n}$, los subespacios $\mathfrak{R}(\mathbf{A})$ y $\mathbf{N}(\mathbf{A}^*)$ son ortonormales entre sí y juntos forman todo \mathbf{C}^m . Consecuentemente, se dice que $\mathbf{N}(\mathbf{A}^*)$ es el complemento ortogonal de $\mathfrak{R}(\mathbf{A})$ en \mathbf{C}^m y vice versa.

Por último, el método SVD de una matriz \mathbf{A} también proporciona una conveniente representación de los proyectores en los subespacios rango y nulo de \mathbf{A} y \mathbf{A}^* . Sea $y \in \mathbf{C}^{m \times 1}$ un vector arbitrario. Por definición, el proyector ortogonal en $\mathfrak{R}(\mathbf{A})$ es la matriz Π que presenta las propiedades:

$$\mathfrak{R}(\Pi) = \mathfrak{R}(\mathbf{A})$$

$$\|y - \Pi y\|^2 = \min \text{ sobre } \mathfrak{R}(\mathbf{A})$$

donde $\|\mathbf{X}\|^2 = \mathbf{X}^* \mathbf{X}$ es la norma euclidiana, lo que significa que la segunda propiedad muestra que la distancia euclidiana entre y y $\Pi y \in \mathfrak{R}(\mathbf{A})$

es mínima. Entonces, si se dispone de una matriz $\mathbf{A} \in \mathbf{C}^{m \times n}$, el proyector ortogonal sobre $\mathfrak{R}(\mathbf{A})$ viene dado por:

$$\Pi = \mathbf{U}_1 \mathbf{U}_1^*$$

mientras que el proyector ortogonal sobre $\mathbf{N}(\mathbf{A}^*)$ es:

$$\Pi_{\perp} = \mathbf{I} - \mathbf{U}_1 \mathbf{U}_1^* = \mathbf{U}_2 \mathbf{U}_2^*$$

Muchas de estas propiedades son explotadas por los métodos SVD que se van a emplear para analizar los conjuntos de datos MRS. Existen diferentes métodos basados en SVD, en función de como se genera la matriz de datos a partir de la FID, y en función de como se procede a la recuperación de la información física de la señal. A continuación se van a describir los tres métodos empleados en este trabajo que, aunque se basan en el formalismo del espacio de estados, no es necesario recurrir al mismo para explicar las bases de su funcionamiento y basta con recurrir al álgebra elemental con matrices.

2.1.1. LPSVD

El método de descomposición en valores singulares mediante predicción lineal (Linear Predictive Singular Value Decomposition) (LPSVD) es uno de los métodos más utilizados en el procesamiento de señales biomédicas para estimar los parámetros de sinusoides exponencialmente amortiguadas. Se basa en dos principios, el SVD y la estimación de coeficientes de predicción lineal.

El algoritmo [13], usa un modelo para ajustar los datos minimizando la diferencia entre ambos en el sentido de mínimos-cuadrados. Dicho modelo es igual al de la *Ec.(1)* de la introducción con un perfil Lorentziano. La señal, $y(n)$, se supone que es una secuencia de datos de N muestras constituida por M resonancias sumadas en un ruido blanco Gussiano complejo:

$$y(n)_{n=0,1,\dots,N-1} = \sum_{k=1}^M a_k e^{s_k n} + w(t_n) \quad (5)$$

con $k = 1, \dots, M$, donde w es el ruido, $a_k = A_k e^{i\phi_k}$ son las amplitudes complejas, conteniendo la fase inicial de la señal, ϕ_k , y $s_k = -\alpha_k + i2\pi f_k$ son los polos de la señal, que contienen el factor de amortiguación y la frecuencia de cada una de las resonancias, los cuales se van a tratar de identificar a

partir de este método de análisis. Para ello, se construye el siguiente sistema de ecuaciones de predicción lineal, utilizando los datos complejos conjugados de la señal y utilizando la predicción lineal inversa:

$$\begin{bmatrix} y^*(1) & y^*(2) & \cdots & y^*(L) \\ y^*(2) & y^*(3) & \cdots & y^*(L+1) \\ \vdots & \vdots & & \vdots \\ y^*(N-L) & y^*(N-L+1) & \cdots & y^*(N-1) \end{bmatrix} \cdot \begin{bmatrix} b(1) \\ b(2) \\ \vdots \\ b(L) \end{bmatrix} = - \begin{bmatrix} y^*(0) \\ y^*(1) \\ \vdots \\ y^*(N-L-1) \end{bmatrix}$$

$$\mathbf{A}\mathbf{b} = -\mathbf{h} \tag{6}$$

donde \mathbf{A} posee estructura de matriz de Hankel, es decir, una matriz cuadrada con todas sus diagonales de derecha a izquierda paralelas numéricamente, y \mathbf{b} es el vector de los coeficientes de predicción inversa. * denota el complejo conjugado. La ecuación anterior se puede escribir de forma aumentada como $\mathbf{A}'\mathbf{b}' = 0$, donde $\mathbf{A}' = (\mathbf{h}|\mathbf{A})$ y $\mathbf{b}' = (1, \mathbf{b}^T)^T$, denotando T la traspuesta de la matriz. \mathbf{b}' denota el vector de coeficientes de predicción del filtro de error. Si se está trabajando con datos en ausencia de ruido, dado que \mathbf{b}' se encuentra en el espacio nulo de \mathbf{A} , el polinomio de filtro predicción-error sería:

$$B(z) = 1 + b(1)z^{-1} + b(2)z^{-2} + \cdots + b(L)z^{-L} \tag{7}$$

La expresión anterior poseerá ceros en $e^{-s_k^*}$, $k = 1, \dots, M$ si L se elige tal que satisfaga la desigualdad $M \leq L \leq N - M$. De estos ceros se obtienen los polos buscados de la señal, e^{-s_k} , reflectándolos en un círculo unidad. Estos ceros, denominados ceros de la señal, caerán fuera de dicho círculo unidad dada la forma de definir el sistema de ecuaciones, *Ec.(5)*, en el sentido inverso de la dirección del tiempo. Si $L > M$, $B(z)$ posee $L - M$ ceros adicionales, denominados superfluos. Si $L < M$, el sistema de ecuaciones posee más de una solución cuando los datos están libres de ruido, ya que el rango de \mathbf{A} , o \mathbf{A}' , es $M(< L)$. Sin embargo, si se encuentra la solución única de mínima norma del sistema, que minimiza $\|\mathbf{b}\|^2 = |b(1)|^2 + |b(2)|^2 + \cdots + |b(L)|^2$, puede demostrarse que los $L - M$ ceros superfluos caerán siempre dentro del círculo unidad.

Se puede factorizar $B(z)$ como $B(z) \approx B_1(z)B_2(z)$, donde $B_1(z)$ es un polinomio de grado M con M ceros. Tal factor existe debido a que $B(z)$

posee ceros en $e^{-s_k^*}$, $k = 1, \dots, M$. Por otro lado, $B_2(z)$ posee los $M - L$ ceros superfluos, en cuya localización se centra el interés. Minimizar la norma de \mathbf{b} es igual a minimizar $\int_{-\pi}^{\pi} |B(e^{iw})|^2 dw$, dado que el primer coeficiente de $B(z)$ es la unidad. Como $B_1(z)$ es fija, minimizar lo anterior es equivalente a resolver el problema estándar de predicción lineal, método de autocorrelación, donde $B_2(z)$ es el filtro predicción-error, $B_1(z)$ corresponde a la transformación Z de la *secuencia de datos* y $B(z)$ es la transformación Z de la *secuencia error*. Se sabe que $B_2(z)$ posee fase mínima. En resumen, los M ceros de señal de $B(z)$ caen fuera del círculo unidad en las posiciones $e^{-s_k^*}$, $k = 1, \dots, M$, mientras que los $L - M$ ceros superfluos caen dentro del mismo, facilitando así la identificación de los ceros buscados. Este es el principal motivo de utilizar los datos en sentido del tiempo inverso en *Ec.(6)*. La razón de usar un polinomio de grado $L > M$ es incrementar la precisión de la localización de los ceros de la señal, es decir, de la estimación de los polos.

Aunque todos estos resultados son totalmente ciertos únicamente en el caso ideal de datos sin ruido, se demuestra experimentalmente que siguen siendo válidos en los casos de datos que presentan un nivel de ruido moderado. Si se intenta resolver en el sentido de mínimos cuadrados el sistema de ecuaciones lineales para datos con ruido, aparecen perturbaciones en el vector \mathbf{b} . La razón de esto, es que para $L > M$, $L - M$ columnas de \mathbf{A} tienden a ser dependientes, dando lugar a un problema de mínimos cuadrados muy inestable, y es esencial un valor de L moderadamente grande para incrementar la exactitud de la localización de la estimación de los polos. Para mejorar el condicionamiento enfermo de \mathbf{A} se emplea SVD en lugar de mínimos-cuadrados. Computando esta descomposición en valores singulares se obtiene:

$$\mathbf{A}_{L,M} = \mathbf{U}_{L,L} \cdot \mathbf{S}_{L,M} \cdot \mathbf{V}_{M,M}^* \quad (8)$$

Se crea una matriz cuadrada a partir de \mathbf{S} , truncando la misma de tal manera que se retengan, en lo posible, únicamente los valores singulares relativos a las componentes resonantes, en este caso M , y descartando los relativos al ruido. De igual manera hay que eliminar las columnas de \mathbf{U} y \mathbf{V} necesarias para tal propósito. Esto permite realizar una inversión de la matriz \mathbf{A} :

$$inv(\mathbf{A}) = \mathbf{V}_{M,M} \cdot \mathbf{S}_{M,M} \cdot \mathbf{U}_{L,M}^* \quad (9)$$

Haciendo uso de ésto y de la *Ec.(6)*, se obtiene la siguiente expresión para \mathbf{b} :

$$\mathbf{b} = \sum_{k=1}^M \sigma_k^{-1} [\mathbf{u}_k^\perp \mathbf{h}] \mathbf{v}_k \quad (10)$$

donde $\sigma_k, k = 1, 2, \dots, L$ ó $N - L$, dependiendo del rango de \mathbf{A} , son los valores singulares de \mathbf{A} , y $\mathbf{v}_k, k = 1, 2, \dots, L$ y $\mathbf{u}_k, k = 1, 2, \dots, N - L$, son los autovectores de $\mathbf{A}^\perp \mathbf{A}$ y $\mathbf{A} \mathbf{A}^\perp$ respectivamente. $^\perp$ denota la matriz compleja conjugada traspuesta. El efecto de usar el truncamiento del SVD es incrementar la tasa señal-ruido, (Signal to Noise Ratio) (SNR), en los datos antes de obtener la solución del vector \mathbf{b} . Hasta ahora se ha supuesto que M era conocida. Si no es así, puede estimarse a partir del valor de los valores singulares de A . En el caso de datos sin ruido, dado que el rango de \mathbf{A} es M , sólo los valores singulares $\sigma_1, \sigma_2, \dots, \sigma_M$ serán distintos de cero, y b será la solución de norma mínima deseada, ya que sólo se usan los M principales autovectores de $\mathbf{A}^\perp \mathbf{A}$ y $\mathbf{A} \mathbf{A}^\perp$ que son más robustos a las perturbaciones del ruido. El resultado es que los $L - M$ ceros externos de B_z tienden a estar menos perturbados. Por tanto, la estabilidad del vector de coeficientes \mathbf{b} se consigue eliminando los valores singulares menos robustos. La solución mediante mínimos-cuadrados puede escribirse de esta misma forma, pero incluyendo todos los valores singulares y los vectores singulares completos, lo que hace que la estabilidad de \mathbf{b} empeore.

Una vez estimados los polos de la señal, pueden usarse para obtener los factores de amortiguación y las frecuencias de cada una de las componentes resonantes que forman la señal como:

$$\alpha_k = \frac{1}{\text{abs}(z_k) \Delta t} \quad (11)$$

$$\omega_k = \frac{-\text{arg}(z_k)}{\Delta t} \quad (12)$$

donde Δt es el intervalo de muestreo de la señal, y $\text{abs}(z_k)$ y $\text{arg}(z_k)$ denotan el módulo y la fase del polo z_k respectivamente. A partir de los polos también se puede recuperar la estimación de cada una de las componentes sin ruido como:

$$x_i^{sig} = \sum_{k=1}^M z_k^0 \left(\frac{1}{z_k} \right)^{i-1} \quad (13)$$

con $i = 1, 2, \dots, N - 1$, donde z_k^0 indica el valor inicial del polo y contiene la amplitud y fase inicial de cada una de ellas. Ambas se pueden recuperar mediante las expresiones:

$$A_k = abs(z_k^0) \quad (14)$$

$$\phi_k = arg(z_k^0) \quad (15)$$

Un sumario de todo el proceso que se ha seguir con este método para realizar la estimación de parámetros, sería:

1. Elegir un modelo para llevar a cabo el ajuste de los datos. En este caso se opta por modelar la señal como la suma de diferentes componentes que siguen un perfil Lorentziano.
2. Formar una matriz de datos \mathbf{A} con estructura de matriz de Hankel, a partir del vector de datos experimentales $y(n)$
3. Aplicar el método SVD sobre la matriz de datos \mathbf{A} para estimar los coeficientes de predicción \mathbf{b}
4. Construir el polinomio característico a partir de los coeficientes en \mathbf{b} para hallar los polos de la señal calculando las raíces del mismo
5. Obtener las estimaciones de los parámetros que definen cada componente resonante: Factor de amortiguación, α_k , frecuencia angular, ω_k , amplitud, A_k y fase inicial, ϕ_k .

Siguiendo estos pasos, se ha creado una función en MATLAB, denominada *LPSVD*, que implementa este método para estimar los parámetros de las componentes resonantes que forman una señal MRS. El código de dicho método, así como toda la documentación del mismo, aparece en el recurso electrónico que se adjunta con esta memoria.

2.1.2. HSVD

La descomposición en valores singulares de Hankel, (Hankel Singular Value Decomposition) (HSVD), es un método de estimación de parámetros que se basa en el formalismo del espacio de estados, en el que la matriz de datos se ordena en forma de matriz de Hankel. Al igual que en el método anterior, el ajuste de la señal se lleva a cabo con los datos en el dominio del tiempo, lo que permite que el proceso de ajuste pueda realizarse de forma no-iterativa [11]. A priori presenta una ventaja sobre el método LPSVD anterior, y es que puede manejar una cantidad mayor de datos al evitar el cálculo de raíces de polinomios.

De nuevo se necesita una función de modelado para realizar el ajuste. En este caso, de nuevo se hace uso de un perfil Lorentziano:

$$x_n = \sum_{k=1}^K c_k z_k^n \quad \text{con } n = 0, 1, \dots, N-1 \quad (16)$$

donde las cantidades $z_k, k = 1, \dots, K$ son los polos de la señal, definidos como:

$$z_k = e^{(-b_k + i\omega_k)\Delta t} \quad (17)$$

donde b_k y ω_k son el factor de amortiguación y la frecuencia angular de la sinuside amortiguada k , y c_k es la amplitud compleja asociada a la misma, donde se ha absorbido la fase inicial, $c_k = A_k e^{i\phi_k}$. Δt es el intervalo de muestreo e i es la unidad imaginaria, $i = \sqrt{-1}$. Se puede ver como esta función depende linealmente de la amplitud y no-linealmente de las frecuencias angulares y factores de amortiguación. Generalmente, dicha dependencia no-lineal acarrea un proceso de ajuste iterativo no-lineal por mínimos-cuadrados. Pero ésto puede solventarse manipulando adecuadamente los datos experimentales, construyendo una matriz de datos, \mathbf{X} , a partir de los datos de $x(n)$, en forma de matriz de Hankel:

$$\mathbf{X} = \begin{bmatrix} x_0 & x_1 & x_2 & x_3 & \cdots & x_M \\ x_1 & x_2 & x_3 & x_4 & \cdots & x_{M+1} \\ x_2 & x_3 & x_4 & x_5 & \cdots & x_{M+2} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ x_L & x_{L+1} & x_{L+2} & x_{L+3} & \cdots & x_{N-1} \end{bmatrix} \quad (18)$$

donde $L = N - M - 1$. Todos los elementos de las diagonales de derecha a izquierda poseen lo mismos valores, lo que es característico de la matriz de Hankel. En una señal sin ruido, el rango de esta matriz de datos es igual al número de sinusoides, K . Como regla, K ha de ser más pequeño que $\min(N - M, M + 1)$, por lo que hay que definir M en concordancia. Ésto implica que \mathbf{X} es no completa. Por tanto, lo que hay que buscar es una factorización de \mathbf{X} en dos matrices de rango completo K . Para ello, el modelo se reescribe como:

$$x_n = \tilde{\mathbf{e}}\mathbf{Z}^n\mathbf{c} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \cdot \begin{bmatrix} z_1^n & & & \\ & z_2^n & & 0 \\ & & \ddots & \\ & 0 & & z_K^n \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_K \end{bmatrix} \quad (19)$$

donde $\tilde{\mathbf{e}}$ es un vector fila con entradas igual a la unidad ($\tilde{}$ denota trasposición), \mathbf{Z} es una matriz diagonal $K \times K$ cuyas entradas en la diagonal son iguales a $z_k, k = 1, \dots, K$, y \mathbf{c} es un vector columna cuyas entradas son las amplitudes complejas. Entonces, usando $X_{ij} = x_{i+j-2} = \tilde{\mathbf{e}}\mathbf{Z}^{i+j-2}\mathbf{c}$, la factorización buscada es:

$$\mathbf{X} = \mathbf{X}_l\mathbf{X}_r = \begin{bmatrix} \tilde{\mathbf{e}} & & & \\ \tilde{\mathbf{e}} & & \mathbf{Z} & \\ \vdots & & & \\ \tilde{\mathbf{e}} & \mathbf{Z}^{N-M-1} & & \end{bmatrix} \cdot \begin{bmatrix} \mathbf{c} & \mathbf{c}\mathbf{Z} & \cdots & \mathbf{Z}^M\mathbf{c} \end{bmatrix} \quad (20)$$

donde la matriz izquierda, \mathbf{X}_l , de rango $(N - M) \times K$ y la matriz derecha, \mathbf{X}_r , de rango $K \times (M + 1)$, tienen rango completo K . \mathbf{X}_l posee una estructura de matriz de Vandermonde, ya que cada fila es igual a la anterior multiplicada por \mathbf{Z} . La misma relación existe entre las sucesivas columnas de \mathbf{X}_r . Es importante destacar que la transformación:

$$x_n = \tilde{\mathbf{e}}\mathbf{Z}^n\mathbf{c} = (\tilde{\mathbf{e}}\mathbf{Q}) (\mathbf{Q}^{-1}\mathbf{Z}^n\mathbf{Q}) (\mathbf{Q}^{-1}\mathbf{c}) = \tilde{\mathbf{e}}'\mathbf{Z}'^n\mathbf{c}' \quad (21)$$

donde \mathbf{Q} es una matriz arbitraria de rango completo y dimensión $K \times K$, deja inalterada la descomposición llevada a cabo en la *Ec.(19)*, convirtiendo simplemente el conjunto $\tilde{\mathbf{e}}\mathbf{Z}^n\mathbf{c}$ en $\tilde{\mathbf{e}}'\mathbf{Z}'^n\mathbf{c}'$. Gracias a esto, la relación especial entre las sucesivas filas de \mathbf{X}_l se retiene tras la transformación, pero reemplazando la matriz diagonal \mathbf{Z} por la no-diagonal \mathbf{Z}' . Dado que \mathbf{Q} es una

matriz no-singular generalizada, cualquier factorización de \mathbf{X} en dos matrices de rango completo, de dimensioness $(N - M) \times K$ y $K \times (M + 1)$, puede expresarse en términos de un conjunto transformado $\tilde{\mathbf{e}}' \mathbf{Z}'^n \mathbf{c}'$. Ahora, tan sólo habría que diagonalizar la matriz resultante, \mathbf{Z}' , para llegar a los polos de la señal, z_k .

Para determinar la esencia del método, hay que explicar como conseguir la factorización de una matriz de datos experimentales y la posterior recuperación de la matriz \mathbf{Z}' . Aplicando SVD, se puede factorizar una matriz de rango K y dimensión $(N - M) \times (M + 1)$ en un producto de tres matrices de rango completo:

$$\mathbf{X} = \mathbf{U}_K \Lambda_K \mathbf{V}_K^\perp \quad (22)$$

donde Λ_K es la matriz truncada de valores singulares de dimensión $K \times K$, \mathbf{U}_K es la matriz de truncada de vectores singulares izquierdos de dimensión $(N - M) \times K$ y \mathbf{V}_K es la matriz de truncada de vectores singulares derechos de dimensión $K \times (M + 1)$, denotando $^\perp$ el complejo conjugado traspuesto o conjugado hermítico. Pasar de esta factorización en tres matrices a una factorización de dos es relativamente sencillo si se absorbe Λ_K en \mathbf{U}_K , \mathbf{V}_K o parcialmente en ambas. Aquí se va a realizar la siguiente identificación:

$$\mathbf{X}_l = \mathbf{U}_K \quad y \quad \mathbf{X}_r = \Lambda_K \mathbf{V}_K^\perp \quad (23)$$

En el caso de señales que contienen contribución del ruido, las expresiones anteriores serían tan sólo aproximaciones.

A continuación se va a proceder a calcular \mathbf{Z}' haciendo uso de la estructura de Vandermonde de la matriz \mathbf{U}_K :

$$\mathbf{U}_b \mathbf{Z}' = \mathbf{U}_t \quad (24)$$

donde \mathbf{U}_b y \mathbf{U}_t son idénticas a la matriz \mathbf{U}_K a las que se las ha eliminado la última y primera fila respectivamente. En ausencia de ruido esta ecuación es exacta y se necesitarían solamente K filas sucesivas de \mathbf{U}_b y \mathbf{U}_t para resolver \mathbf{Z}' . Sin embargo, en presencia de ruido, lo anterior es sólo una aproximación y todas las filas disponibles se usan para resolverlo en el sentido de mínimos-cuadrados. La solución estándar lineal, no-iterativa, de mínimos-cuadrados es:

$$\mathbf{Z}' = (\mathbf{U}_b^\perp \mathbf{U}_b)^{-1} \mathbf{U}_b^\perp \mathbf{U}_t \quad (25)$$

Finalmente, tan sólo restaría diagonalizar esta matriz \mathbf{Z}' para obtener los polos de la señal, y a partir de éstos, obtener las estimaciones de los parámetros de cada componente a partir de las mismas expresiones que se relataban en el método anterior en las ecuaciones (11),(12),(14) y (15).

Un resumen de todo el proceso que se ha de seguir con este método para realizar la estimación de parámetros, sería:

1. Formar una matriz de datos \mathbf{X} con estructura de matriz de Hankel, a partir del vector de datos experimentales x_n
2. Aplicar el método SVD sobre \mathbf{X} , truncándola apropiadamente
3. Calcular la matriz \mathbf{Z}'
4. Diagonalizar la matriz \mathbf{Z}' para obtener los polos de la señal.
5. Obtener las estimaciones de los parámetros que definen cada componente resonante: Factor de amortiguación, α_k , frecuencia angular, ω_k , amplitud, A_k y fase inicial, ϕ_k .

Siguiendo estos pasos, se ha creado una función en MATLAB, denominada *HSVD*, que implementa este método para estimar los parámetros de las componentes resonantes que forman una señal MRS. El código de dicho método, así como toda la documentación del mismo, aparece en el recurso electrónico que se adjunta con esta memoria.

2.1.3. SELF-SVD

La descomposición en valores singulares de una banda de frecuencia seleccionada, (SElected Frequency band Singular Value Decomposition) (SELF-SVD), es un método de ajuste para modelos de sinusoides amortiguadas en el dominio de la frecuencia, usando una implementación en frecuencia del espacio de estados y matrices empleadas en el análisis espectral. Este tipo de métodos surge del hecho de que a la hora de analizar datos MRS el usuario puede estar interesado únicamente en las componentes presentes en una pequeña banda de frecuencias del espectro. Con este método se pueden estimar sólo los parámetros de las componentes contenidas en esta banda,

con la ventaja de contener la menor interferencia posible de las componentes externas a la misma. Desde el punto de vista práctico, ésto puede suponer una ventaja, ya que las componentes externas pueden no poseer un perfil Lorentziano puro, y tratarlas como interferencias no deseadas y eliminarlas del modelo puede mejorar la estimación de las componentes de interés.

Al igual que se ha hecho con los otros dos métodos, se va a describir el método y su funcionamiento, y se van a describir los pasos para realizar una implementación computacionalmente simple del mismo [19].

De nuevo, el modelo considerado para describir la señal MRS viene dado por un conjunto de sinusoides exponencialmente amortiguadas que siguen un perfil Lorentziano:

$$y(t)_{t=0,1,\dots,N-1} = \sum_{k=1}^m \rho_k \lambda_k^t + \epsilon(t) \quad \text{con} \quad \lambda_k = e^{(-\alpha_k + i\omega_k)\Delta t} \quad (26)$$

donde m denota el número de componentes, ρ_k es la amplitud compleja en la que se ha absorbido la fase inicial, $\rho_k = A_k e^{i\phi_k}$; α_k y ω_k son el coeficiente de amortiguación y la frecuencia angular de la componente k respectivamente, contenidos en los modos de la señal λ_k ; Δt es el intervalo de muestreo, N es el número de puntos de muestreo y $\epsilon(t)$ es el término de ruido, que consiste en ruido complejo blanco Gaussiano circular. Si se denotan estos datos como un vector, se tendría:

$$\mathbf{y} = [y(0), \dots, y(N-1)]^T \quad (27)$$

donde T denota trasposición. Con los modos de la señal pueden definirse dos cantidades asociadas con la estructura de Vandermonde que permitirán posteriormente describir en profundidad el desarrollo de este método:

$$\mathbf{a}(\lambda_k) = [\lambda_k, \dots, \lambda_k^S]^T \quad (28)$$

$$\mathbf{b}(\lambda_k) = [1, \lambda_k, \dots, \lambda_k^{N-1}]^T \quad (29)$$

que permiten reescribir la expresión del modelo de la señal como:

$$\mathbf{y} = \sum_{l=1}^m \rho_l \mathbf{b}(\lambda_l) + \epsilon(t) \quad (30)$$

S es un parámetro a definir por el usuario que se detallará más adelante. Este método puede trabajar con todo el espectro, pero su peculiaridad es que se puede seleccionar una banda frecuencial de interés, cuyas frecuencias de Fourier se podrían describir como:

$$\left\{ \frac{2\pi k_1}{N}, \frac{2\pi k_2}{N}, \dots, \frac{2\pi k_M}{N} \right\} \quad (31)$$

donde k_1, \dots, k_M son M enteros dados. Habitualmente, estos enteros son consecutivos, pero no es una condición obligatoria para este método y se puede trabajar con bandas de frecuencia no adyacentes. Lo que si se ha de asumir es que se conoce el número de componentes que están presentes en esta banda, n y se ha de cumplir que $n \leq m$. Si n no es conocida en un principio, ésta puede estimarse, bien contando el número de picos en la banda, o bien a través del rango de la matriz de datos que se va a construir a partir de los datos experimentales, como se explicará más adelante.

Estas frecuencias de Fourier permiten definir w_k como $w_k = e^{i2\pi k/N}$ para $k = 0, \dots, N - 1$. Ésto permite definir a su vez, cantidades asociadas al vector de Fourier k -ésimo siguiendo también una estructura de Vandermonde:

$$\mathbf{u}_k = [w_k, \dots, w_k^S]^T \quad (32)$$

$$\mathbf{v}_k = [1, w_k, \dots, w_k^{N-1}]^T \quad (33)$$

que permitirían calcular la FT de los datos como:

$$Y_k = \mathbf{v}_k^* \mathbf{y} \quad \text{con } k = 0, \dots, N - 1 \quad (34)$$

donde $*$ denota la traspuesta conjugada. Con todas estas cantidades que se han definido se puede calcular lo siguiente, definiendo p como $p = 1, \dots, S$:

$$\begin{aligned}
w_k^p [\mathbf{v}_k^* \mathbf{b}(\lambda)] &= \sum_{t=0}^{N-1} e^{[-\alpha+i(\omega-(2\pi/N)k)]t} e^{i2\pi kp/N} \\
&= \lambda^p \sum_{t=0}^{N-1} e^{[-\alpha+i(\omega-(2\pi/N)k)](t-p)} \\
&= \lambda^p [\mathbf{v}_k^* \mathbf{b}(\lambda)] + \lambda^p \left[\sum_{t=0}^{p-1} \lambda^{t-p} e^{-i2\pi k(t-p)/N} - \sum_{t=N}^{N+p-1} \lambda^{t-p} e^{-i2\pi k(t-p)/N} \right] \\
&= \lambda^p [\mathbf{v}_k^* \mathbf{b}(\lambda)] + \lambda^p \sum_{l=1}^p [\lambda^{-l} e^{i2\pi kl/N} - \lambda^{N-l} e^{i2\pi kl/N}] \\
&= \lambda^p [\mathbf{v}_k^* \mathbf{b}(\lambda)] + \sum_{l=1}^p \lambda^{p-l} (1 - \lambda^N) w_k^l
\end{aligned} \tag{35}$$

Definiendo γ_p como:

$$\gamma_p^*(\lambda) = (1 - \lambda^N) [\lambda^N, \dots, \lambda, 1, 0, \dots, 0] \tag{36}$$

se puede escribir de forma más compacta lo obtenido en la Ec.(35) como:

$$w_k^p [\mathbf{v}_k^* \mathbf{b}(\lambda)] = \lambda^p [\mathbf{v}_k^* \mathbf{b}(\lambda)] + \gamma_p^*(\lambda) \mathbf{u}_k \tag{37}$$

o, equivalentemente:

$$\mathbf{u}_k [\mathbf{v}_k^* \mathbf{b}(\lambda)] = \mathbf{a}(\lambda) [\mathbf{v}_k^* \mathbf{b}(\lambda)] + \begin{bmatrix} \gamma_1^*(\lambda) \\ \vdots \\ \gamma_S^*(\lambda) \end{bmatrix} \mathbf{u}_k \tag{38}$$

Combinando las ecuaciones (27) y (38) se tiene:

$$\begin{aligned}
\mathbf{u}_k Y_k &= \sum_{l=1}^m \rho_l \mathbf{u}_k [\mathbf{v}_k^* \mathbf{b}(\lambda_l)] + \mathbf{u}_k \epsilon_k \\
&= [\mathbf{a}(\lambda_1, \dots, \lambda_m)] \begin{bmatrix} \rho_1 \mathbf{v}_k^* \mathbf{b}(\lambda_1) \\ \vdots \\ \rho_m \mathbf{v}_k^* \mathbf{b}(\lambda_m) \end{bmatrix} + \left\{ \sum_{l=1}^m \rho_l \begin{bmatrix} \gamma_1^*(\lambda_l) \\ \vdots \\ \gamma_S^*(\lambda_l) \end{bmatrix} \right\} \mathbf{u}_k + \mathbf{u}_k \epsilon_k \\
&= [\mathbf{a}(\lambda_1, \dots, \lambda_m)] \begin{bmatrix} \rho_1 \mathbf{v}_k^* \mathbf{b}(\lambda_1) \\ \vdots \\ \rho_m \mathbf{v}_k^* \mathbf{b}(\lambda_m) \end{bmatrix} + \Gamma \mathbf{u}_k + \mathbf{u}_k \epsilon_k
\end{aligned} \tag{39}$$

donde $\{\epsilon_k\}$ se define de forma similar a $\{Y_k\}$, ya que $\{\epsilon_k\}$ es la transformada de Fourier del ruido $\{\epsilon(t)\}_{t=0}^{N-1}$. A partir de ahora, $\{\lambda(t)\}_{k=1}^n$ van a denotar los modos de la Ec.(26) que se encuentran en la banda definida por la Ec.(31). Para separar los modos de las componentes de interés de los modos despreciables, se define:

$$\mathbf{A} = [\mathbf{a}(\lambda_1), \dots, \mathbf{a}(\lambda_n)] \tag{40}$$

$$\mathbf{x}_k = \begin{bmatrix} \rho_1 \mathbf{v}_k^* \mathbf{b}(\lambda_1) \\ \vdots \\ \rho_n \mathbf{v}_k^* \mathbf{b}(\lambda_n) \end{bmatrix} \tag{41}$$

para los modos de interés y de la misma forma se definen $\tilde{\mathbf{A}}$ y $\tilde{\mathbf{x}}_k$ para el resto de modos. Finalmente, para escribir la Ec.(39) de forma totalmente compacta mediante matrices para $k = k_1, \dots, k_M$, se definen:

$$\mathbf{Y} = [\mathbf{u}_{k_1} Y_{k_1}, \dots, \mathbf{u}_{k_M} Y_{k_M}] \tag{42}$$

$$\mathbf{U} = [\mathbf{u}_{k_1}, \dots, \mathbf{u}_{k_M}] \tag{43}$$

$$\mathbf{X} = [\mathbf{x}_{k_1}, \dots, \mathbf{x}_{k_M}] \tag{44}$$

Similarmente a \mathbf{X} se obtendría $\tilde{\mathbf{X}}$. Sustituyendo estas expresiones en la Ec.(39) se obtiene:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \Gamma\mathbf{U} + \tilde{\mathbf{A}}\tilde{\mathbf{X}} + \epsilon \quad (45)$$

donde la matriz ϵ relativa al ruido, se define de forma similar a \mathbf{Y} . Se pueden realizar una serie de observaciones sobre los dos últimos términos de la expresión anterior:

- La contribución de los elementos del término de ruido son mucho menores que los de la componente $\mathbf{A}\mathbf{X}$. De hecho, se tiene que los elementos del ruido siguen un patrón $O(N^{1/2})$ (estocásticamente), mientras que los elementos de \mathbf{X} siguen un patrón $O(N)$, siempre y cuando la SNR sea suficientemente grande.
- Si se supone que las componentes externas a la banda no son mucho más grandes que las buscadas y que sus frecuencias no están muy cercanas al intervalo de interés, los elementos de $\tilde{\mathbf{X}}$ serán mucho más pequeños que los de \mathbf{X} .

A partir de aquí se va a asumir que $M \geq n + S$, la cuál no es una condición muy exigente ya que el parámetro S es definido por el usuario y siempre puede elegirse de tal forma que cumpla esta condición. Usualmente, la forma de definir el valor de S viene dado por, $S \in \left[\left\lceil \frac{M}{3} \right\rceil, \left\lceil \frac{M}{2} \right\rceil\right]$, donde $\lceil x \rceil$ denota la parte entera de x . Teniendo esto en mente, se puede definir la matriz de proyección ortogonal en el espacio nulo de \mathbf{U} como:

$$\Pi_{\mathbf{U}}^{\perp} = \mathbf{I} - \mathbf{U}^* (\mathbf{U}\mathbf{U}^*)^{-1} \mathbf{U} \quad (46)$$

Si se postmultiplica con esta expresión la Ec.(45), por un lado se elimina el segundo término, y por otro, si no se cumplen las observaciones sobre la Ec.(45), la postmultiplicación por esta matriz de proyección ortogonal atenuará tanto las contribuciones del ruido como las contribuciones de las componentes externas a la banda de interés. Por tanto, llevando a cabo esta postmultiplicación, se obtiene:

$$\mathbf{Z} = \mathbf{Y}\Pi_{\mathbf{U}}^{\perp} \approx \mathbf{A}\mathbf{X}\Pi_{\mathbf{U}}^{\perp} \quad (47)$$

El rango de esta matriz \mathbf{Z} puede utilizarse para estimar, si no se conocía a priori, o verificar, si ya se conocía, el número de señales presentes en la banda de frecuencias de interés. Para el caso de datos sin ruido se tiene que:

$$\text{rango}(\mathbf{Z}) = \text{rango}(\mathbf{Y}\Pi_{\bar{U}}^\perp) = n \quad (48)$$

Para el caso de señales que contienen contribuciones del ruido, esto puede servir de estimación, pero una mejor forma de estimar el número de señales presentes, es a través de la inspección de los valores singulares de la matriz \mathbf{Z} tras someterla a una descomposición mediante el método SVD. Una vez definido el número de señales presentes en el intervalo frecuencial, a partir del truncamiento de la matriz de vectores singulares izquierdos de la descomposición SVD de \mathbf{Z} en función de éste se construye la matriz \mathbf{W} , de dimensión $S \times n$, y cuyas columnas son los vectores singulares izquierdos asociados a los n valores singulares mayores. A partir de esta definición de \mathbf{W} y la Ec.(47) se tiene que:

$$\mathbf{A} \approx \mathbf{W}\mathbf{P} \quad (49)$$

donde \mathbf{P} es una matriz de transformación no-singular. A partir de esta matriz \mathbf{W} se pueden estimar los modos de las componentes presentes en el rango frecuencial estudiado, $\{\lambda(t)\}_{k=1}^n$, de los cuales se pueden obtener las estimaciones de los coeficientes de amortiguación y frecuencias angulares de cada una de ellas. Para ello, se construyen las matrices \mathbf{W}_u y \mathbf{W}_l a partir de \mathbf{W} , eliminando la primera fila y la última de \mathbf{W} respectivamente. Similarmente se definen \mathbf{A}_u y \mathbf{A}_l a partir de \mathbf{A} . Aplicando estas expresiones en la Ec.(49) se tiene que:

$$\mathbf{W}_l\mathbf{P} \approx \mathbf{A}_l = \mathbf{A}_u \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \approx \mathbf{W}_u\mathbf{P} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \equiv \mathbf{W}_l \approx \mathbf{W}_u\Phi \quad (50)$$

donde

$$\Phi = \mathbf{P} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \mathbf{P}^{-1} \quad (51)$$

Por tanto, se podría estimar Φ como

$$\hat{\Phi} = (\mathbf{W}_u^*\mathbf{W}_u)^{-1} \mathbf{W}_u^*\mathbf{W}_l \quad (52)$$

Los autovalores de esta matriz proporcionan los modos asociados a las diferentes componentes presentes en la banda de frecuencias bajo estudio. A partir de ellos y mediante las ecuaciones (11), (12), (14) y (15) detalladas en el método LPSVD, se pueden obtener las estimaciones de los parámetros que proporcionan la información física de cada una de estas componentes.

Un resumen de todo el proceso que se ha de seguir con este método para realizar la estimación de parámetros, sería:

1. Realizar la transformada de Fourier sobre el vector de datos experimentales $\{y(t)\}_{t=0}^{N-1}$ para obtener $\{Y_k\}_{k=0}^{N-1}$. También se ha de elegir la banda de frecuencias sobre la que se quiere llevar a cabo el análisis y definir el valor del parámetro S .
2. Construir la matriz Π_U^\perp
3. Calcular la matriz \mathbf{Z}
4. Aplicar el método SVD a la matriz \mathbf{Z} para determinar el número de componentes, n , presentes en la banda bajo estudio y construir la matriz \mathbf{W} .
5. Calcular la matriz $\hat{\Phi}$ y obtener sus autovalores
6. Obtener las estimaciones de los parámetros que definen cada componente resonante: Factor de amortiguación, α_k , frecuencia angular, ω_k , amplitud, A_k y fase inicial, ϕ_k .

Siguiendo estos pasos, se ha creado una función en MATLAB, denominada *SELF-SVD*, que implementa este método para estimar los parámetros de las componentes resonantes que forman una señal MRS. El código de dicho método, así como toda la documentación del mismo, aparece en el recurso electrónico que se adjunta con esta memoria.

3. Discusión Y Resultados

Se va a proceder al análisis de los resultados obtenidos tras someter, tanto espectros simulados como reales, a los diferentes métodos de cuantificación que se han relatado en la sección anterior. Los datos simulados se han empleado para testar el funcionamiento, comportamiento y fidelidad de las diferentes técnicas que se han implementado, mientras que los datos reales se han sometido a estudio con éstas para caracterizarlos y determinar la presencia o ausencia de la componente relativa al biomarcador asociado al proceo de neurogénesis.

3.1. Métodos SVD

Primero, se implementaron los métodos SVD descritos en el Método Experimental como funciones de MATLAB. Como ya se ha comentado en la Introducción, los dos primeros métodos, el LPSVD y el HSVD, aparecen implementados en el software ya existente para el análisis de datos MRS. Presentan la desventaja de que en dichas implementaciones funcionan como métodos estilo *caja negra*, en los que la única interacción que permiten con el usuario es determinar el número de señales presentes en el espectro bajo estudio. Precisamente ésto no es lo más adecuado cuando se estudia un conjunto de datos sobre el que no se posee información previa o, como en este caso, cuando se busca la presencia de picos de baja intensidad de los que se dispone de información. Por otro lado, el método SELF-SVD no aparece implementado en el software de uso común en este ámbito. En las implementaciones se ha intentado mejorar la interactividad de estos métodos con el usuario, proporcionando un mayor control sobre los parámetros del análisis. Además, se han incluido en un interfaz gráfico (GUI) de MATLAB para facilitar su empleo por parte de los usuarios. Una mejor descripción tanto de los métodos como del GUI aparece en el Anexo I de esta memoria y los códigos de todo ello se incluye en el recurso electrónico que se ha adjuntado con la misma.

3.1.1. Espectros simulados

Lo primero que se ha hecho es someter los métodos de análisis a datos MRS simulados para evaluar el funcionamiento de los mismos. Se han simulado datos que contienen desde una hasta cuatro componente resonantes,

tanto en ausencia como en presencia de ruido. A continuación se detallan los parámetros de cada una de las componentes resonantes simuladas:

Señal	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
Señal 1	10.00	163.56	1.280	40.00	$\frac{\pi}{6}$
Señal 2	20.00	10.00	0.078	70.00	$\frac{\pi}{4}$
Señal 3	14.30	500.00	3.913	100.00	$\frac{\pi}{9}$
Señal 4	33.30	700.00	5.478	120.03	$\frac{\pi}{3}$

Cuadro 1: Valores de las componentes resonantes que se van a ajustar y caracterizar mediante los métodos SVD por separado o en combinación.

En los estudios con estas señales se pretendía ver el comportamiento de los métodos SVD a la hora de caracterizar señales aisladas y conjuntos de señales, en ausencia de ruido y con ruido presente en las mismas. Los resultados que se han obtenido se muestran en las siguientes tablas:

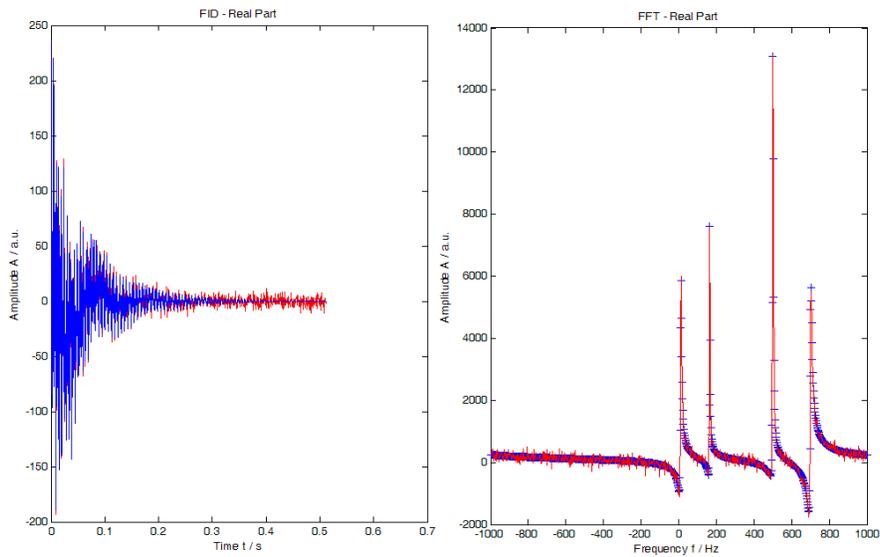


Figura 6: Ejemplo de los ajustes (azul) en el dominio del tiempo y de la frecuencia, obtenidos mediante el análisis con los métodos SVD de una señal simulada con 4 componentes con ruido añadido.

Señal	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
1	10.00	163.56	1.280	40.00	0.5236
1 + r	10.10	163.58	1.2802	40.64	0.5215
1,2	10.00	163.56	1.280	40.00	0.5236
	20.00	10.00	0.078	70.00	0.7854
1,2 + r	9.75	163.58	1.280	39.90	0.5331
	19.58	10.01	0.078	69.01	0.7819
1,2,3	10.00	163.56	1.280	40.00	0.5236
	20.00	10.00	0.078	70.00	0.7854
	14.30	500.00	3.913	100.00	0.3491
1,2,3 + r	10.34	163.51	1.279	41.15	0.5459
	20.58	10.00	0.078	71.90	0.7734
	14.29	500.01	3.913	100.01	0.3426
1,2,3,4	10.00	163.56	1.280	40.00	0.5236
	20.00	10.00	0.078	70.00	0.7854
	14.30	500.00	3.913	100.00	0.3491
	33.30	700.00	5.478	120.00	1.0472
1,2,3,4 + r	9.77	163.54	1.279	39.33	0.5261
	20.52	10.11	0.079	70.11	0.7799
	14.20	499.97	3.913	98.75	0.3573
	33.05	700.06	5.478	119.42	1.0434

Cuadro 2: Resultados de la estimación de parámetros de las diferentes señales simuladas usando el método HSVD. Los números indican la componente o componentes presentes en la simulación y r indica que se ha añadido ruido.

Señal	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
1	10.00	163.56	1.280	40.00	0.5236
1 + r	10.52	163.54	1.279	41.30	0.5212
1,2	10.00	163.56	1.280	40.00	0.5236
	20.00	10.00	0.08	70.00	0.7854
1,2 + r	10.62	163.49	1.279	40.89	0.5407
	21.09	9.99	0.078	71.88	0.7939
1,2,3	10.00	163.56	1.280	40.00	0.5236
	20.00	10.00	0.078	70.00	0.7854
	14.30	500.00	3.91	100.00	0.3491
1,2,3 + r	10.52	163.57	1.280	41.34	0.5316
	21.20	9.93	0.077	72.03	0.7872
	14.90	499.99	3.913	102.46	0.3482
1,2,3,4	10.00	163.56	1.280	40.00	0.5236
	20.00	10.00	0.078	70.00	0.7854
	14.30	500.00	3.913	100.00	0.3491
	33.30	700.00	5.478	120.00	1.0472
1,2,3,4 + r	10.54	163.57	1.280	41.08	0.5202
	21.18	9.95	0.077	71.44	0.7981
	14.88	499.98	3.913	101.80	0.3503
	35.63	700.07	5.478	124.44	1.0435

Cuadro 3: Resultados de la estimación de parámetros de las diferentes señales simuladas usando el método LPSVD. Los números indican la componente o componentes presentes en la simulación y r indica que se ha añadido ruido.

Señal	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
1	10.00	165.51	1.290	34.09	-0.0237
1 + r	10.20	165.48	1.295	34.87	0.0002
1,2	10.00	165.51	1.290	34.38	-0.0151
	20.00	11.95	0.090	66.31	0.5001
1,2 + r	10.15	165.54	1.295	34.86	-0.0115
	20.62	12.00	0.094	67.00	0.5083
1,2,3	10.00	165.51	1.290	34.05	-0.0091
	20.00	11.95	0.090	66.19	0.5084
	14.30	501.95	3.930	92.19	-0.0494
1,2,3 + r	10.21	165.48	1.295	34.85	0.0161
	20.29	11.94	0.0934	67.54	0.5037
	14.34	501.95	3.928	92.50	-0.0527
1,2,3,4	10.00	165.51	1.290	33.90	-0.0097
	20.00	11.95	0.090	65.99	0.5109
	14.30	501.95	3.930	91.76	-0.0529
	33.30	701.95	5.490	120.53	0.8606
1,2,3,4 + r	9.96	165.51	1.295	33.90	-0.0124
	19.74	11.84	0.093	65.93	0.5282
	14.34	501.96	3.928	92.75	-0.0585
	33.68	701.95	5.493	122.11	0.8597

Cuadro 4: Resultados de la estimación de parámetros de las diferentes señales simuladas usando el método SELF-SVD. Los números indican la componente o componentes presentes en la simulación y r indica que se ha añadido ruido.

Se puede ver como el método más preciso de los tres es el HSVD. El SELF-SVD proporciona los peores resultados de ajuste a la hora de utilizarlo para analizar el espectro al completo. Esto se debe a que es el método que más transformaciones de los datos necesita y en cada una de ellas se van acumulando los errores que se producen debido al ruido presente en el espectro y las desviaciones que puedan tener las componentes resonantes respecto al perfil Lorentziano que se supone idealmente para ellas. Se puede ver a tenor de los resultados que la presencia de un ruido leve no afecta demasiado a la estimación de parámetros y ésta va empeorando a medida que aumenta el número de señales presentes en el espectro.

En este sentido, se van a mostrar los resultados de las estimaciones que se obtienen cuando en el análisis no se especifica el número de señales previamente, y cuando el número de señales es menor o mayor que el auténtico. Ésto sólo se va a realizar con las simulaciones con ruido añadido, ya que en las simulaciones ideales no tiene ningún efecto, siempre se recupera el número de señales presentes. Además, el método LPSVD no se ha podido someter a este estudio ya que estima automáticamente el número de señales presentes a través de las raíces del polinomio característico que usa para el cálculo de los polos de la señal, y únicamente se puede influir en el orden de éste polinomio. Los resultados obtenidos fueron:

Señales	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
1 + r	10.3	163.59	1.280	41.34	0.5167
2 + r	9.76	165.58	1.280	39.98	0.5343
	19.50	10.01	0.078	69.34	0.7780
3 + r	10.04	163.48	1.279	39.99	0.5628
	20.53	9.88	0.077	71.94	0.8108
	14.28	500.00	3.913	99.88	0.3448
4 + r	9.71	163.57	1.280	39.09	0.5067
	20.60	10.10	0.079	70.76	0.7776
	14.13	499.95	3.912	98.13	0.3657
	32.83	700.03	5.48	117.45	1.0446
4 + r (3)	14.43	500.00	3.913	101.54	0.3454
	55.29	30.45	0.238	49.57	-0.1349
	38.03	699.52	5.474	129.27	1.0816
4 + r (5)	10.10	163.56	1.280	40.38	0.5257
	19.54	9.90	0.077	69.58	0.8069
	14.36	500.03	3.913	101.03	0.3363
	33.42	700.00	5.478	121.11	1.0447
	23.84	200.30	1.567	0.5613	-2.1601

Cuadro 5: Resultados de la estimación de parámetros de las diferentes señales simuladas en presencia de ruido (r) usando el método HSVD cuando no se especifica de antemano el número de componentes. Para el último caso, en el que hay presentes 4 componentes, se muestran los resultados de introducir en el método como número de componentes (3) ó (5).

Señales	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
4 + r (3)	14.38	501.98	3.929	93.83	-0.0626
	11.54	165.30	1.294	39.290	0.0364
	40.77	700.89	5.485	136.39	0.0953
4 + r (5)	9.97	165.51	1.295	33.99	-0.0128
	19.75	11.84	0.093	65.94	0.5280
	14.35	501.96	3.9284	92.73	-0.0585
	33.67	701.95	5.493	122.09	0.8599
	2.66	201.94	1.580	0.8189	0.9639

Cuadro 6: Resultados de la estimación de parámetros de las diferentes señales simuladas en presencia de ruido (r) usando el método SELF-SVD, para el caso en el que hay presentes 4 componentes y se introducen en el método como número de componentes (3) ó (5).

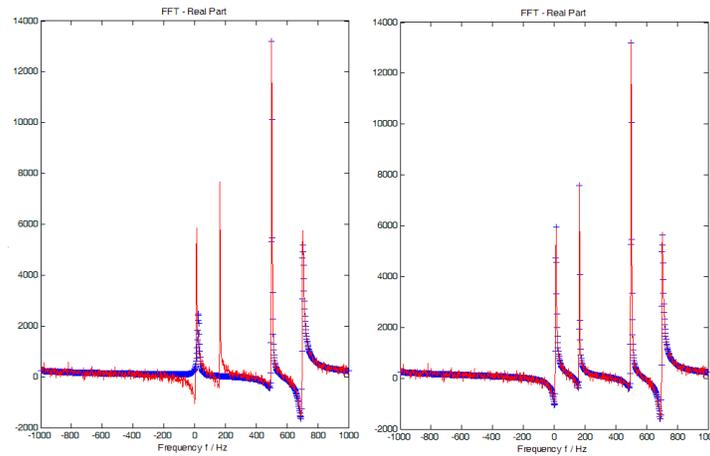


Figura 7: Ejemplo de los ajustes (azul) que se obtienen al suponer tres (izq.) y cinco (dcha.) componentes al analizar un espectro simulado con cuatro componentes más ruido.

Aquí sólo se muestran las estimaciones obtenidas para las componentes que se buscaban, pero en los datos de salida aparece un número de componentes resonantes que forman la señal igual a la dimensión de la matriz de datos usada por el método para los casos en los que no se especifica el número de componentes presentes. Esto supone una desventaja ya que usar un número de componentes para simular la señal igual a la dimensión de la matriz de datos, hace que la contribución del ruido sea mucho mayor y se pierda precisión. Para SELF-SVD no se muestra esta parte del estudio, ya

que las estimaciones que se obtenían carecían de todo sentido. Respecto al hecho de especificar un número de señales distinto al real, se consigue una mejor estimación cuando se introduce un valor de número de señales superior al real frente a introducir uno inferior, lo cuál es lógico ya que si se hace lo último se pierde información al eliminar del ajuste las componentes con valores singulares más pequeños. Al añadir un número mayor, el modelo genera componentes extra que afectan menos a la precisión del ajuste que perder información. Por su parte, al ejecutar el método LPSVD sobre el espectro simulado con cuatro componentes más ruido, predice la existencia de cinco componentes, lo que es una estimación bastante precisa.

A pesar de que el método LPSVD realice una estimación del número de señales presentes en un espectro, siempre añade más componentes de las que de verdad hay presentes. El método más fiable a la hora de estimar el número de señales presentes es la observación de los valores singulares obtenidos mediante la descomposición SVD. Un ejemplo de representación de valores singulares se muestra a continuación:

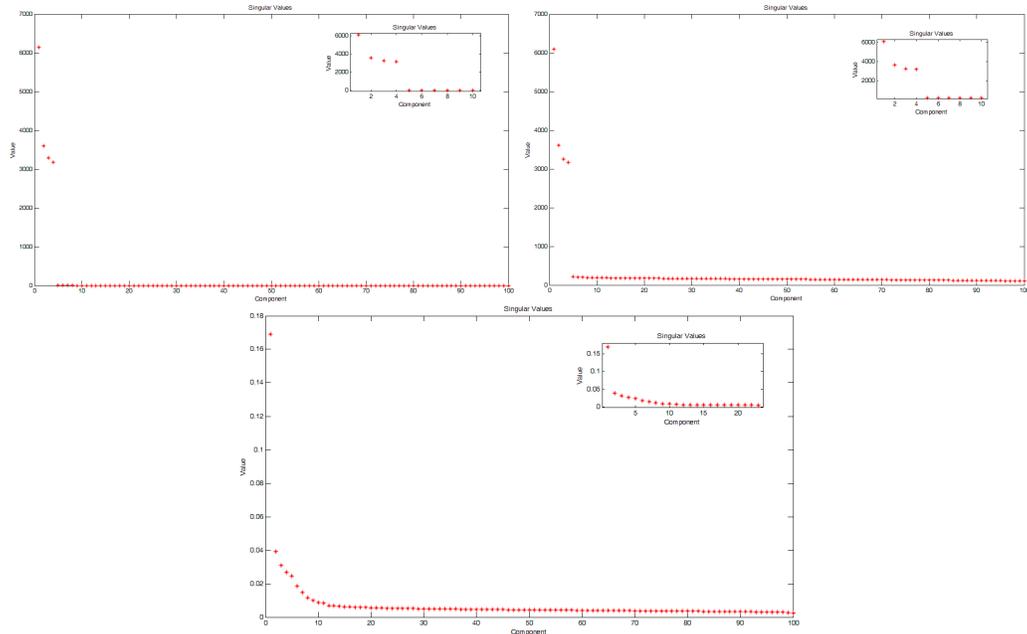


Figura 8: Representación gráfica de los valores singulares de un espectro simulado con 4 componentes (arriba izq.), el mismo espectro con ruido añadido (arriba dcha.) y un espectro real *in vivo* (abajo).

En el caso de la señal sin ruido y con presencia leve del mismo, se puede ver de forma clara como hay una discontinuidad que separa los valores singulares referentes a las componentes, del resto. Sin embargo, en el caso de espectros reales, en los que pueden existir componentes con una intensidad similar a la del ruido, se puede ver como esta discontinuidad ya no está tan clara. Aún así, sirve para poder estimar el número de componentes y proceder así a realizar un nuevo análisis con mayor precisión. Por tanto, la forma más segura para determinar una estimación del número de componentes presentes en un espectro del que no se dispone información previa, sería comparar el número de componentes que designa el método LPSVD junto con una observación de la representación de los valores singulares de la señal.

A continuación se va a estudiar la habilidad de los métodos para resolver superposición de componentes. Se simuló la superposición de dos y tres de las componentes de un espectro compuesto por cuatro de ellas. Se definieron como se muestra en las siguientes tablas:

Señal	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
Señal 1	10.00	163.56	1.280	40.00	$\frac{\pi}{6}$
Señal 2	20.00	164.00	1.283	70.00	$\frac{\pi}{4}$
Señal 3	14.30	500.00	3.913	100.00	$\frac{\pi}{9}$
Señal 4	33.30	700.00	5.478	120.03	$\frac{\pi}{3}$

Cuadro 7: Valores de las componentes resonantes que se van a ajustar y caracterizar mediante los métodos SVD para estudiar la superposición de dos componentes en un espectro formado por cuatro de ellas.

Señal	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
Señal 1	10.00	163.56	1.280	40.00	$\frac{\pi}{6}$
Señal 2	20.00	164.00	1.283	70.00	$\frac{\pi}{4}$
Señal 3	14.30	162.00	1.268	100.00	$\frac{\pi}{9}$
Señal 4	33.30	700.00	5.478	120.03	$\frac{\pi}{3}$

Cuadro 8: Valores de las componentes resonantes que se van a ajustar y caracterizar mediante los métodos SVD para estudiar la superposición de tres componentes en un espectro formado por cuatro de ellas.

Los resultados que se obtuvieron se muestran en la siguientes tablas:

Señales	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
4,2	10.00	163.56	1.280	40.00	0.5236
	20.00	164.00	1.283	70.00	0.7854
	14.30	500.00	3.913	100.00	0.3491
	33.30	700.00	5.4781	120.00	1.0472
4,2 + r	15.32	163.63	1.280	107.86	0.6985
	14.56	499.95	3.913	100.36	0.3634
	33.39	700.07	5.479	120.08	1.0417
	14.07	-524.67	-4.106	0.6696	-0.3370
4,3	10.00	163.56	1.280	40.00	0.5236
	20.00	164.00	1.283	70.00	0.7854
	14.30	162.01	1.267	100.00	0.3491
	33.30	700.0	5.478	120.00	1.0472
4,3 + r	17.28	162.79	1.274	207.45	0.5302
	33.07	700.07	5.479	120.70	1.0410
	17.31	100.29	0.785	1.60	2.8397
	35.28	-355.67	-2.783	0.64	0.8089

Cuadro 9: Resultados de la estimación de parámetros mediante el método HSVD para el caso de dos y tres componentes superpuestas en simulaciones de señales con cuatro componentes, en ausencia y presencia ruido.

Señales	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
4 2	10.00	163.56	1.280	40.00	0.5236
	20.00	164.00	1.283	70.00	0.7854
	14.30	500.00	3.913	100.00	0.3491
	33.30	700.00	5.4781	120.00	1.0472
4 2 + r	19.19	163.82	1.282	111.00	0.7543
	14.80	500.00	3.913	102.19	0.3415
	35.511	699.92	5.477	124.81	1.0528
	2.99	-605.61	-4.739	0.6013	-0.4983
4 3	10.00	163.56	1.280	40.00	0.5236
	20.00	164.00	1.283	70.00	0.7854
	14.30	162.01	1.267	100.00	0.3491
	33.30	700.0	5.478	120.00	1.0472
4 3 + r	17.39	162.03	1.268	154.21	0.5157
	38.29	163.09	1.276	56.35	-0.4661
	35.46	700.29	5.479	124.94	1.0416
	10.78	99.19	0.776	1.72	-3.0312

Cuadro 10: Resultados de la estimación de parámetros mediante el método LPSVD para el caso de dos y tres componentes superpuestas en simulaciones de señales con cuatro componentes, en ausencia y presencia ruido.

Señales	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
4 2	10.00	165.51	1.295	32.03	-1.0618
	20.00	165.95	1.299	156.74	1.2048
	14.30	501.95	3.928	91.36	-0.0540
	33.30	701.95	5.493	119.87	0.8619
4 2 + r	10.97	165.50	1.295	98.31	-0.8427
	14.37	501.94	3.928	92.05	-0.0516
	33.61	701.92	5.493	121.33	0.8660
	22.25	166.12	1.300	147.87	1.3255
4 3	10.00	165.51	1.295	43.73	0.9490
	20.00	165.95	1.299	208.23	2.5755
	14.30	163.95	1.283	320.28	-0.2821
	33.30	701.95	5.493	117.94	1.8707
4 3 + r	16.03	164.66	1.288	234.03	0.5665
	33.48	702.01	5.494	117.90	0.8694
	7.39	164.82	1.290	99.85	-2.0920
	1.26	-140.82	-1.102	0.63	-2.2162

Cuadro 11: Resultados de la estimación de parámetros mediante el método SELF-SVD para el caso de dos y tres componentes superpuestas en simulaciones de señales con cuatro componentes, en ausencia y presencia ruido.

Se puede ver como en ausencia de ruido los métodos son capaces de resolver la superposición de señales, pero fallan a la hora de analizarla en espectros con presencia de ruido. Por tanto, a la hora de analizar un espectro real, ninguno de los tres métodos será capaz de resolver una superposición de componentes.

También se ha probado la eficacia del método SELF-SVD para analizar las componentes presentes en una banda frecuencial concreta dentro de un espectro. Los resultados que se obtuvieron fueron los siguientes:

Señal	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad
1	9.99	165.51	1.295	35.38	0.0073
1 + r	8.68	165.23	1.293	33.77	0.0509

Cuadro 12: Resultados obtenidos para el ajuste de la señal 1 mediante la selección de una banda frecuencial a través del método SELF-SVD en un espectro formado por cuatro componentes.

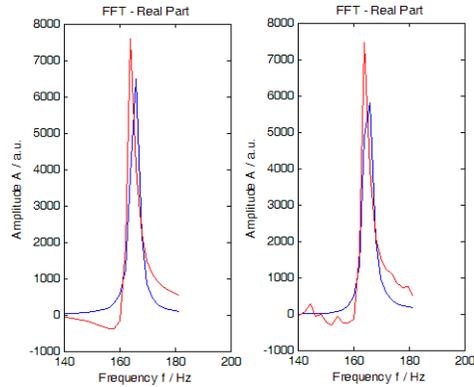


Figura 9: Representación gráfica de los ajustes obtenidos para la señal 1 utilizando una banda de frecuencias concreta a través del método SELF-SVD para un espectro simulado de 4 componentes en presencia (dch.) y ausencia de ruido (izq.).

Se puede ver como el método SELF-SVD funciona adecuadamente y es capaz de estimar los parámetros de las señales presentes en la banda de frecuencias seleccionada. La precisión de esta estimación empeora al añadir ruido y al aumentar el número de componentes del espectro, lo que concuerda con el resto de resultados que se habían obtenido en los anteriores estudios.

Por último se ha realizado un estudio sobre los errores de precisión de los métodos de ajuste. Para ello, primero se llevó a cabo un estudio sobre como varía el error a medida que se varía la SNR. Se obtuvieron los siguientes resultados:

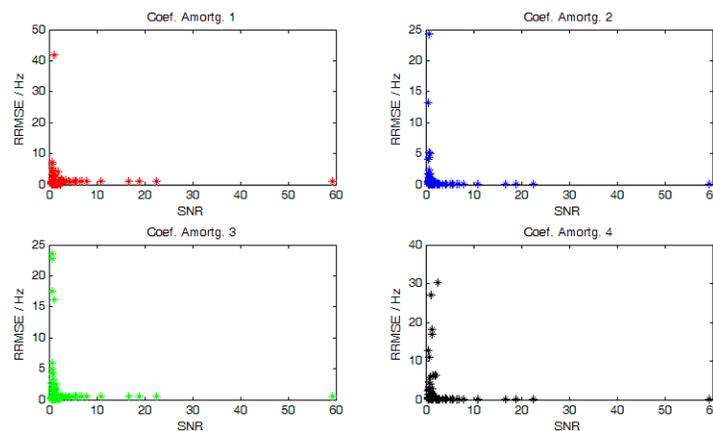


Figura 10: Representación gráfica de los valores del error en la estimación de los factores de amortiguación de un espectro simulado con 4 componentes en función de la SNR.

Se ve, como era de esperar, que el error que se produce en la estimación aumenta al disminuir la SNR. El error se ha definido como:

$$RRMSE = \sqrt{\left(\frac{1}{s}\right) \frac{\sum_{i=1}^s (\bar{x} - x_i)^2}{\bar{x}^2}} \quad (53)$$

donde x_i es el valor estimado del parámetro, \bar{x} hace referencia al valor medio del mismo y s es el número de simulaciones realizadas.

Por último, se realizaron 1000 simulaciones de una señal, con las cuatro componentes descritas al inicio de esta sección, a la que se añadía ruido blanco generado aleatoriamente. Después, se realizó estadística con las estimaciones de cada parámetro para ver si seguían una distribución de tipo Gaussiano, para así poder estimar el error, en el caso de espectros reales donde la muestra de los mismos es bastante limitada, mediante la desviación estándar de la misma. Se obtuvo lo siguiente:

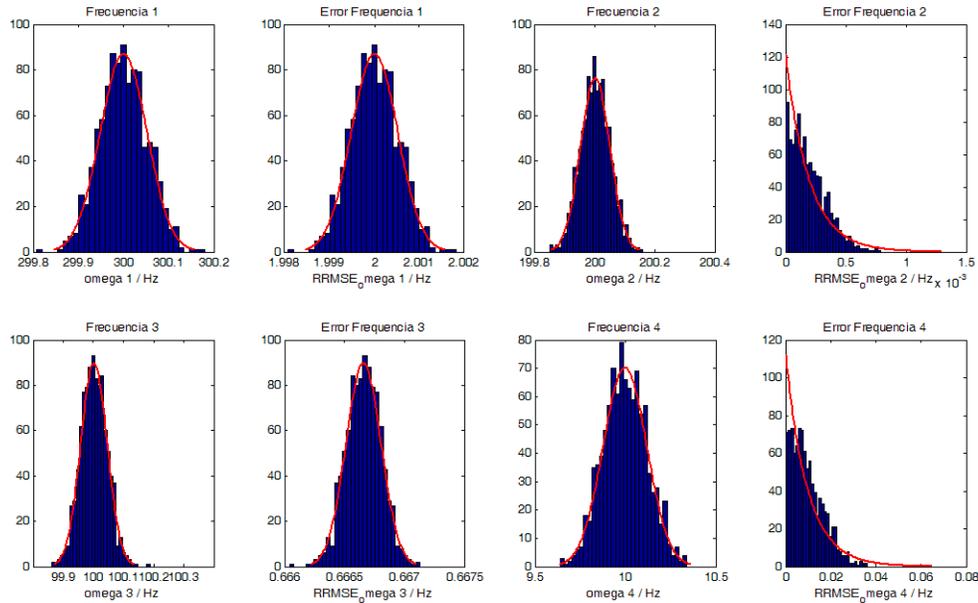


Figura 11: Histogramas de las estimaciones de las frecuencias de las 4 componentes del espectro simulado y los errores de las mismas, ajustados a Gaussianas.

Se puede ver como los resultados obtenidos para cada parámetro se pueden describir mediante una Gaussiana, y por tanto, ésto será lo que se use para estimar el error de las estimaciones obtenidos en el estudio de los espectros

reales. El hecho de que los histogramas del error de las frecuencias 2 y 4 aparezcan de esa manera es debido a la forma de definir el error según la Ec.(53) y el valor de los mismos.

3.1.2. Espectros reales

A la vista de los resultados obtenidos en el apartado anterior, se desarrolló una estrategia para analizar los espectros de datos *in vivo* de hipocampos humanos, que consta de tres partes. Primero se realizó un análisis inicial aplicando el método LPSVD sobre todos ellos, para determinar en cada uno el posible número de componentes que los integran. Después, se procedió al análisis de todos ellos con los métodos HSVD y SELF-SVD, introduciendo como número de componentes el obtenido mediante el primer estudio con el método LPSVD. A continuación, se realizó una comparación entre los resultados obtenidos con los tres métodos para determinar que estimaciones de coeficientes coincidían en los tres estudios y seleccionar éstas como posibles componentes reales presentes en los mismos. Por último, se llevó a cabo un análisis con el método HSVD, introduciendo como número de componentes el resultado de la comparación anterior, e incrementando algo más este valor.

A continuación se va a mostrar una tabla en la que se muestran las estimaciones de los parámetros obtenidos para cada componente que aparecía reflejada en el análisis con los tres métodos. En esta tabla va a aparecer una nueva columna para mostrar una frecuencia en ppm que se ha calculado como si se hubiese tomado como referencia el pico el agua para recopilar los datos el espectro. Esto responde a que en la mayor parte de estudios de este tipo se toma como referencia el pico del agua, y de esta forma se pueden comparar estos resultados con los que aparecen en [35]. Se van a mostrar únicamente los datos obtenidos mediante el método HSVD ya que ha demostrado ser el más preciso en las estimaciones y es el método que se va a emplear en los estudios posteriores:

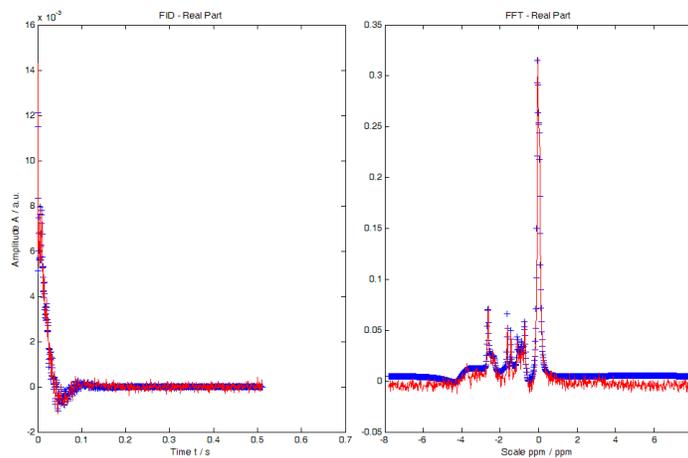


Figura 12: Representación gráfica de un ajuste (azul) a una señal *in vivo* en el dominio del tiempo (arriba dcha.) y en el de la frecuencia (arriba izc.).

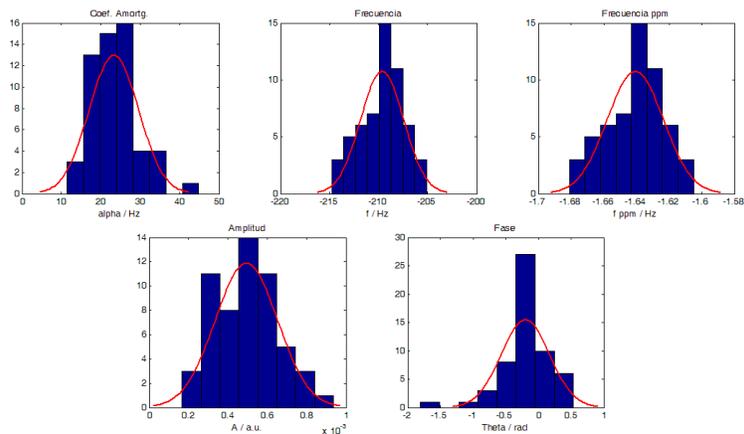


Figura 13: Histogramas de las estimaciones de los parámetros de una componente resonante de los espectros *in vivo* ajustados a una Gaussiana.

	NS	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad	$\nu_{ppm}^{H_2O}/\text{ppm}$
1	58	44.18	-340.88	-2.667	0.00094	0.3809	4.740
2	57	23.40	-209.63	-1.640	0.00049	-0.1982	3.712
3	57	29.22	-187.47	-1.467	0.00054	0.2341	3.539
4	50	39.03	-95.19	-0.745	0.00193	-0.2185	2.817
5	51	44.26	-143.94	-1.126	0.00059	0.2775	3.198
6	45	43.75	8.09	0.063	0.00508	0.1823	2.009
7	63	32.70	0.03	0.0002	0.01178	0.4340	2.072
8	45	43.44	-8.667	-0.068	0.00068	-0.3139	2.140
9	32	37.97	85.65	0.670	0.00030	0.9594	1.402
10	28	56.23	-117.45	-0.920	0.00078	-0.0838	2.991
11	32	111.40	-476.41	-3.728	0.00056	0.0486	5.800
12	29	188.90	-298.03	-2.332	0.00149	-0.5036	4.404
13	34	98.23	-75.50	-0.591	0.00147	-0.6196	2.663
14	18	88.25	-104.10	-0.815	0.00146	-0.1020	2.887
15	16	64.54	25.377	0.199	0.00074	0.7542	1.873
16	20	100.05	-26.71	-0.209	0.00636	-0.5520	2.281
17	19	26.56	107.63	0.842	0.00006	0.11359	1.230
18	16	123.93	-48.85	-0.382	0.00229	-0.20967	2.454
19	32	88.27	-78.07	-0.611	0.00117	-0.76335	2.683

Cuadro 13: Resultados obtenidos tras los análisis de los datos *in vivo* mediante el método HSVD. NS hace referencia al número de espectros donde aparece la componente resonante.

Esta mismo proceso se realizó tanto con el método LPSVD, como con el SELF-SVD. A la luz de los datos obtenidos, se puede estimar que el número de componentes resonantes oscila entre siete u ocho, ya que el resto de posibles componentes aparece en menos de la mitad de los espectros estudiados. Éstas pueden deberse a coincidencias a la hora de que los métodos generen componentes resonantes para intentar modelar el resto del espectro, o ser realmente componentes que están presentes en algunos de los espectros debido a algún motivo físico del sujeto bajo estudio. Esto requeriría un estudio por parte de los doctores, ya que los datos se anonimizan antes de ser analizados. Tomando esto como punto de partida, se procedió a realizar el análisis de los espectros con el método HSVD, indicando que hay presentes en los mismos desde 7 hasta 10 componentes. Los resultados obtenidos se muestran en las siguientes tablas:

	NS	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad	$\nu_{ppm}^{H_2O}/\text{ppm}$
1	62	67.65	-339.32	-2.655	0.00146	0.2974	4.727
2	56	19.72	-209.24	-1.637	0.00040	-0.3432	3.709
3	56	27.24	-186.07	-1.456	0.00047	-0.0910	3.528
4	45	79.64	-93.44	-0.731	0.00149	-0.5895	2.803
5	45	37.46	-10.54	-0.083	0.00685	-0.2864	2.154
6	63	34.36	0.638	0.005	0.00964	0.3691	2.067
7	35	131.30	-136.68	-1.070	0.00200	0.0000	3.142

Cuadro 14: Resultados obtenidos tras los análisis de los datos *in vivo* mediante el método HSVD suponiendo 7 componentes. NS hace referencia al número de espectros donde aparece la componente resonante.

	NS	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad	$\nu_{ppm}^{H_2O}/\text{ppm}$
1	58	58.59	-339.67	-2.658	0.00130	0.2744	4.730
2	58	20.00	-209.61	-1.640	0.00039	-0.2500	3.712
3	58	26.21	-186.25	-1.458	0.00045	-0.0556	3.530
4	48	53.98	-93.93	-0.735	0.00099	-0.4232	2.807
5	38	43.71	-10.20	-0.080	0.00741	-0.2626	2.152
6	63	3.42	0.35	0.003	0.00984	0.3103	2.069
7	44	79.63	-137.29	-1.074	0.00111	0.0030	3.146
8	11	30.72	7.42	0.058	0.00573	0.0054	2.014

Cuadro 15: Resultados obtenidos tras los análisis de los datos *in vivo* mediante el método HSVD suponiendo 8 componentes. NS hace referencia al número de espectros donde aparece la componente resonante.

	NS	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad	$\nu_{ppm}^{H_2O}/\text{ppm}$
1	57	49.71	-340.30	-2.663	0.00110	0.3553	4.735
2	57	21.03	-209.94	-1.643	0.00044	-0.1185	3.715
3	58	27.68	-186.75	-1.461	0.00051	0.2404	3.533
4	54	71.14	-141.87	-1.110	0.00102	0.4912	3.182
5	50	58.30	-92.27	-0.722	0.00107	-0.3797	2.794
6	41	59.63	-11.48	-0.090	0.00680	-0.2460	2.162
7	63	33.45	0.332	0.003	0.01018	0.2286	2.069
8	13	64.47	-107.40	-0.840	0.00096	0.1392	2.912
9	14	38.85	7.49	0.059	0.00642	-0.0605	2.013

Cuadro 16: Resultados obtenidos tras los análisis de los datos *in vivo* mediante el método HSVD suponiendo 9 componentes. NS hace referencia al número de espectros donde aparece la componente resonante.

	NS	α/Hz	ν/Hz	ν_{ppm}/ppm	$A/\text{a.u.}$	ϕ_0/rad	$\nu_{ppm}^{H_2O}/\text{ppm}$
1	56	44.03	-341.24	-2.671	0.00097	0.4993	4.742
2	58	23.50	-210.05	-1.648	0.00049	-0.0926	3.716
3	58	27.68	-186.75	-1.461	0.00051	0.2404	3.533
4	54	61.13	-141.72	-1.109	0.00075	0.2919	3.181
5	50	41.11	-94.86	-0.742	0.00074	-0.0635	2.814
6	42	43.52	-10.03	-0.078	0.00727	-0.3589	2.150
7	63	31.58	0.31	0.002	0.16670	0.3101	2.070
8	20	50.41	-115.66	-0.905	0.00063	-0.2757	2.977
9	18	32.77	8.41	0.066	0.00428	-0.5403	2.006

Cuadro 17: Resultados obtenidos tras los análisis de los datos *in vivo* mediante el método HSVD suponiendo 10 componentes. NS hace referencia al número de espectros donde aparece la componente resonante.

A la vista de los resultados obtenidos se puede decir que el número de componentes promedio presentes en los espectros de hipocampos es de ocho o nueve, ya que para estos valores es para los que se consigue un mejor ajuste. También se puede ver como la supuesta componente asociada a la neurogénesis no aparece en ninguno de los análisis. Se procedió a realizar un ajuste en la banda frecuencial en la que debería aparecer esta componente con la ayuda del método SELF-SVD, pero no se obtuvo ningún resultado satisfactorio.

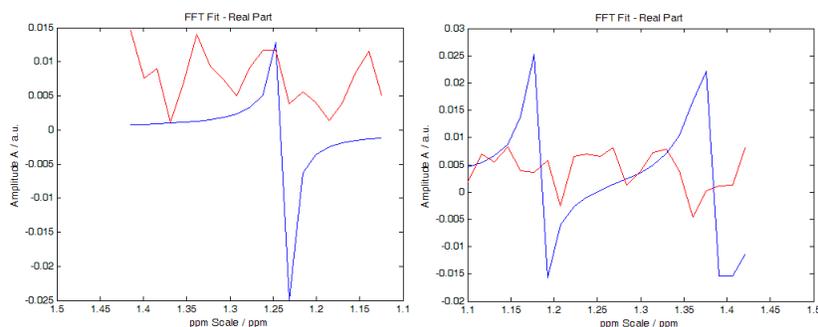


Figura 14: Muestra de un ajuste (azul) mediante el método SELF-SVD de un espectro real en la banda de frecuencias donde se esperaba encontrar el biomarcador asociado a la neurogénesis, suponiendo una o dos componentes.

Por último, si se comparan los valores obtenidos para $\nu_{ppm}^{H_2O}$ con la bibliografía, [35], se puede ver como la mayoría de componentes pueden asociarse con metabolitos que se espera que estén presentes en espectros de hipocampos, como el N-acetylaspartate (NAA) (2.00 ppm), Cholina (cho) (3.5 ppm), Creatina (cre) (3.8 ppm), y otros.

Aunque en esta memoria tan sólo se muestren una parte de los datos y gráficas obtenidos al realizar los diferentes test a los que se han sometido los métodos para comprobar su eficiencia y robustez, así como los de los estudios de los datos *in vivo*, debido al espacio limitado de la misma, los resultados completos, así como todas las gráficas asociadas a los mismos, aparecen recogidos en el recurso electrónico que se adjunta con esta memoria.

4. Conclusiones

A partir de todos los resultados que se han obtenido durante la realización de este proyecto, se pueden sacar una serie de conclusiones tanto sobre los métodos empleados para realizar los análisis de los datos, como de los propios datos.

Respecto a los métodos SVD que se han empleado, también se pueden sacar una serie de conclusiones tanto a nivel colectivo como individual. Por un lado, este tipo de métodos han demostrado ser una herramienta útil para llevar a cabo análisis en muestras de datos sobre las que no se tiene información previa o cuando se pretende buscar nuevas componentes resonantes que puedan estar presentes en las señales. Esto se debe a que son relativamente rápidos en su ejecución, al ser no-iterativos, y no requieren la inclusión de mucha información previa sobre los datos por parte del usuario. Esto permite obtener de forma rápida una primera estimación sobre el tipo de espectro que se está estudiando. En el trabajo realizado se ha modificado la implementación habitual de estos métodos para tratar de hacerlos más interactivos y permitir al usuario poder determinar más parámetros de funcionamiento de los mismos, pero procurando siempre que esta clase de valores a introducir sean fácilmente determinables. También ha quedado constatado que ninguno de estos tres métodos es capaz de determinar la presencia de superposición de componentes en la señal cuando hay ruido presente en la misma. Esto corrobora una de las críticas respecto al estudio publicado en [30], donde se argumentaba que la componente resonante asociada al biomarcador de la neurogénesis podría tratarse de una superposición de señales de lípidos y que este tipo de métodos no serían capaces de determinar este hecho.

En cuanto a los resultados que se han obtenido al someter los espectros simulados a los métodos SVD, se han podido determinar una serie de características de los mismos. En la mayoría de los test que se han realizado, el método HSVD ha demostrado ser el más preciso a la hora de realizar las estimaciones de los parámetros de las componentes resonantes que integran la señal y el más rápido en su ejecución. Esto último, se debe a que es el método que menos transformaciones de los datos realiza para llevar a cabo el análisis. Por su parte el LPSVD necesita calcular las raíces de un polinomio para llegar a las estimaciones, y el SELF-SVD requiere realizar la transformada de Fourier de los datos, ya que trabaja en el dominio de la frecuencia.

En todos ellos, para un nivel de ruido moderado, afecta en mayor medida a la precisión de las estimaciones el número de componentes que forma la señal que el ruido en sí.

Por su parte el método LPSVD presenta la ventaja de determinar de forma automatizada una estimación de las componentes resonantes que integran la señal MRS. En la implementación que se ha realizado, se puede pasar como dato el orden del polinomio para determinar los polos de la señal, el cuál esta relacionado con el número de señales como se indicó en la descripción del método en esta memoria. Así todo, siempre ha de ser de mayor que el número de señales presentes, por lo que siempre se obtiene una estimación sobreestimada del número de componentes. Aunque, como se también se ha podido observar, siempre es mejor esta opción a la hora de realizar el análisis, que la de subestimar el número de las mismas. Siempre se puede combinar esta estimación con la observación de la gráfica de valores singulares de la señal para realizar una estimación aún mas certera.

El método SELF-SVD ha demostrado que es el menos eficiente de los tres estudiados, ya que presenta la peor precisión en las estimaciones. Parte de esta falta de precisión se debe a trabajar en el dominio de la frecuencia, donde al tener que realizar una transformación de los datos originales para poder trabajar en ella, ya se introduce una incertidumbre adicional que se arrastra a lo largo de todo el análisis. Por otra parte, la posibilidad de poder trabajar únicamente con una banda de frecuencias concreta puede resultar muy útil para ahorrar tiempo de estudio de los espectros cuando se conoce de antemano la posición en el mismo de la componente que se desea caracterizar. De esta manera se puede realizar un estudio de una banda de frecuencias alrededor de la posición de la componente deseada, y determinar de forma rápida la presencia o ausencia de la misma en el espectro. Si está presente, se puede proceder a un estudio más concreto por medio de otros métodos, ya que la estimación de los parámetros de esa componente mediante este método no es suficientemente precisa.

Por tanto, cuando se plantea el análisis de un conjunto de datos espectrales sobre los que no se posee información previa, la forma de proceder más segura y efectiva sería, primero realizar una estimación del número de componentes que integran la señal, mediante el método LPSVD, que procura un método automático para determinar el mismo, combinado con una visualización de la representación de los valores singulares de la señal. Una vez

conocido, o estimado, el número de componentes, la forma más rápida y simple de conocer una estimación de los parámetros que las definen es realizar un análisis con el método HSVD. Si, por otro lado, se quiere determinar únicamente la presencia o ausencia de una determinada señal sobre la que se conoce su posición en el espectro, la forma más rápida sería aplicar el método SELF-SVD sobre una banda de frecuencias alrededor la misma.

El análisis llevado a cabo sobre los espectros reales también ha dado luz a una serie de aspectos. Primero se ha podido determinar, tras realizar un análisis con los tres métodos sobre los 63 espectros de hipocampos izquierdos humanos, que no se ha podido determinar la presencia de una componente resonante a 1.28 ppm en ninguno de ellos. Este resultado, sumado al que se ha comentado previamente sobre los métodos SVD, apoya las teorías que ponen en tela de juicio los resultados expuestos en [30].

Se ha podido ver como a la hora de llevar a cabo la estimación de los parámetros de las componentes en los espectros reales, los errores en las mismas son bastante elevados. Ésto, a parte de las posibles contribuciones debidas a la presencia de ruido y a la desviación del perfil Lorentziano supuesto para las componentes, se debe a los conjuntos de datos con los que se trabaja. El set de espectros de hipocampos procede, en este caso, de un grupo control y de un grupo que presenta una serie de patologías neurológicas. En ambos grupos cada espectro pertenece a un sujeto de prueba distinto. Ésto hace que, salvo la frecuencia característica de la componente, que depende tan sólo del entorno del átomo sobre el que actúa la resonancia, el resto de parámetros sean variables de una persona a otra. El coeficiente de amortiguación da una idea del entorno químico en el que se encuentra la sustancia a la que hace referencia la componente, y éste, aunque debería ser similar, no tiene porque ser exacto en todos los sujetos, lo que introduce una incertidumbre extra. Asimismo, la amplitud da una idea de la concentración de la sustancia, la cuál también varía de un sujeto a otro. Por último, la fase inicial depende de la orientación del vector magnetización al inicio de la precisión, el cuál no es el mismo exactamente en cada realización. Para mejorar el rendimiento de estos métodos que se han desarrollado, lo ideal sería que se pudiese disponer de un set de datos más amplio y que procedan de un único sujeto, reduciendo así la incertidumbre debida a las diferentes condiciones que se dan al trabajar con varios sujetos. De esta manera se podría realizar un análisis que proporcionase estimaciones de parámetros de componentes

más precisas.

Habitualmente, en los estudios de este tipo se utilizan métodos de pre-procesado para eliminar las llamadas componentes de base. Éstas presentan un valor del factor de amortiguación más elevado que el de las componentes asociadas a metabolitos. Como en este estudio no se han suprimido éstas antes de llevar a cabo el análisis de los datos, pueden extraerse las estimaciones de dichas componentes para realizar un estudio de las mismas y ver si pueden tener alguna influencia en alguna de las patologías que experimentan algunos de los sujetos sometidos a estudio.

En este trabajo se pretendía haber usado el método basado en wavelets de autocorrelación, que aparece comentado en la introducción, para buscar la componente asociada a la neurogénesis. Ésto no se ha podido realizar, ya que se requiere tener una base de datos de espectros lo suficientemente amplia para poder definir esta wavelet a partir de la estadística de las estimaciones de los parámetros de dicha componente. Debido a que no se ha conseguido encontrar en ninguno de los espectros la componente asociada a la neurogénesis, no se dispone ni de las estimaciones de los parámetros de la misma, ni de estadística. Así todo, este procedimiento parece el más prometedor para su introducción en el ámbito clínico a este respecto, ya que proporcionaría una forma rápida y precisa de estimar la presencia de una componente concreta en un espectro. Además, las propiedades de compresión y dilatación de este tipo de wavelets las hace menos susceptibles a las variaciones experimentales que se han comentado anteriormente.

Por último, todos los métodos que se han implementado se han recogido en un interfaz gráfico (GUI) de MATLAB, que permite un uso rápido y sencillo de los mismos.

5. Agradecimientos

Se agradece su colaboración en el desarrollo de todos los ámbitos del proyecto a Diego Herranz Muñoz. También se agradece al Dr. Benedicto Crespo Facorro y a Roberto Roiz Santiañez, integrantes del Laboratorio de Neuroimagen asociado al Instituto de Formación e Investigación Marqués de Valdecilla (IFIMAV) y al área de Psiquiatría del Hospital Universitario Marqués de Valdecilla, por su colaboración en el desarrollo del proyecto y la facilitación de los datos MRS de hipocampos de cerebros humanos, procedentes de sujetos de prueba anónimos. Asimismo, se agradece a Rafael Marco de Lucas, fundador del Instituto de Análisis y Técnicas de Estudio del Cerebro (IATEC), por ser el artífice de cohesionar todas las partes implicadas en este proyecto y hacerlo posible.

6. Bibliografía

1. Marinette van der Graaf, *In vivo magnetic resonance spectroscopy: basic methodology and clinical applications*, European Biophysics Journal (2010), Vol. 39 pp. 527-540, DOI:10.1007/s00249-009-0517-y
2. R.J. Ordidge, M.R. Bendall, R.E. Gordon, A. Connelly, *Volume selection for in vivo biological spectroscopy*, In: Magnetic resonance in biology and medicine. New Delhi: Tata McGraw-Hill; 1985. pp. 387397.
3. J. Frahm, H. Bruhn, M.L. Gyngell, K.D. Merboldt, w. Hänicke, R. Sauter, *Localized high-resolution proton NMR spectroscopy using stimulated echoes: initial applications to human brain in vivo*, Magn Reson Med. 1989;9:7993. doi: 10.1002/mrm.1910090110.
4. Pravat K. Mandal, *In vivo proton magnetic resonance spectroscopic signal processing for the absolute quantitation of brain metabolites*, European Journal of Radiology (2011), DOI:10.1016/j.ejrad.2011.03.076
5. A. van den Boogaart, *Quantitative Data Analysis of In Vivo MRS Data Sets*, Magnetic Resonance in Chemistry (1997), Vol.35 pp. 146-152, DOI:10.1002/(SICI)1097-458X(199712)35:13;S146::AID-OMR196;3.0.CO;2-X
6. J.B. Pouillet, D.M. Sima, S. Van Huffel, *MRS signal quantitation: A review of time- and frequency-domain methods*, Journal of Magnetic Resonance, December 2008, Vol. 195 pp. 134-144, DOI:10.1016/j.jmr.2008.09.005
7. Filip Jiru, *Introduction to post-processing techniques*, European Journal of Radiology (2008), , Vol.67 pp. 202-217, DOI:10.1016/j.ejrad.2008.03.005
8. J.W. van der Veen, R. de Beer, P.R. Luyten, D. van Ormondt, *Accurate quantification of in vivo ^{31}P NMR signals using the variable projection method and prior knowledge*, Magnetic Resonance in Medicine, January 1988, Vol. 6 no. 1 pp. 92-98, DOI:10.1002/mrm.1910060111
9. L. Vanhamme, A. van den Boogaart, S. Van Huffel, *Improve method for accurate and efficient quantification of MRS data with use of prior knowledge*, Journal of Magnetic Resonance, November 1997, Vol. 129 pp. 35-43, DOI:10.1006/jmre.1997.1244

10. H. Ratiney, M. Sdika, Y. Coenradic, S. Cavassila, D. van Ormondt, D. Graveron-Demilly, *Time-domain semi-parametric estimation based on a metabolite basis set*, NMR in Biomedicine, February 2005, Vol. 18 pp. 1-13, DOI:10.1016/S1090-7807(03)00188-5
11. H. Barkhuijsen, R. de Beer, D. van Ormondt, *Improved Algorithm for Noniterative Time-Domain Model Fitting to Exponentially Damped Magnetic Resonance Signals*, Journal of Magnetic Resonance, July 1987, Vol. 73 pp. 553-557, DOI:10.1016/0022-2364(87)90023-0
12. S. Van Huffel, H. Chen, C. Decanniere, P. Van Hecke, *Algorithm for time-domain NMR data fitting based on total least squares*, Journal of Magnetic Resonance, October 1994, Vol. 110 pp. 228-237, DOI:10.1006/jmra.1994.1209
13. R. Kumaresan, D.W. Tufts, *Estimating the Parameters of Exponentially Damped Sinusoids and Pole-Zero Modeling in noise*, Acoustics: Speech and Signal Processing IEEE, December 1982, Vol. 30 pp. 833-840, DOI:10.1109/TASSP.1982.1163974
14. H. Barkhuijsen, R. de Beer, W.M. Bovee, J.H. Creyghton, D. van Ormondt, *Application of linear prediction and singular value decomposition (LPSVD) to determine NMR frequencies and intensities from the FID*, Magnetic Resonance in Medicine, February 1985, Vol. 2 pp. 86-89, DOI:10.1002/mrm.1910020111
15. S. Cavassila, S. Deval, C. Huegen, D. van Ormondt, D. Graveron-Demilly, *Cramér-Rao bounds, an evaluation tool for quantitation*, NMR in Biomedicine (2001), Vol. 14 pp. 278-283, DOI:10.1002/nbm.701
16. V.A. Mandelshtam, H.S. Taylor, A.J. Shaka, *Application of the filter diagonalization method to one- and two-dimensional NMR spectra*, Journal of Magnetic Resonance, August 1998, Vol. 133 pp. 304-312, DOI:10.1006/jmre.1998.1476
17. M. Cedervall, P. Stoica, R. Moses, *Mode-type algorithm for estimating damped, undamped or explosive modes*, Circuits, Systems and Signal Processing, May 1997, Vol. 16 pp. 349-362, DOI:10.1007/BF01246717
18. N. Sandgren, Y. Selen, P. Stoica, J. Li, *Parametric methods for frequency selective MR spectroscopy, a review*, Journal of Magnetic Resonance, June 2004, Vol. 168 pp. 259-272, DOI:10.1016/j.jmr.2004.03.011

19. P. Stoica, N. Sandgren, Y. Selen, L. Vanhamme, S. Van Huffel, *Frequency-Domain method based on the singular value decomposition for frequency-selective NMR spectroscopy*, Journal of Magnetic Resonance, November 2003, Vol. 165 pp. 80-88, DOI:10.1016/S1090-7807(03)00188-5
20. S.W. Provencher, *Estimation of metabolite concentrations from localized in vivo proton NMR spectra*, Magnetic Resonance in Medicine, June 1993, Vol. 30 pp. 672-679, DOI:10.1002/mrm.1910390607
21. J. Slotboom, C. Boesch, R. Keis, *Versatile frequency domain fitting using time domain models and prior knowledge*, Magnetic Resonance in Medicine, June 1998, Vol. 39 pp. 899-911, DOI:10.1002/mrm.1910390607
22. H. Serrai, L. Senhadji, J.D. de Certaines, J.L. Coatrieux, *Time-domain Quantification of Amplitude, Chemical Shift, Apparent Relaxation Time T_2^* and Phase by Wavelet-Transform Analysis. Application to Biomedical Magnetic Resonance Spectroscopy*, Journal of Magnetic Resonance, January 1997, Vol. 124 pp. 20-34, DOI:10.1006/jmre.1996.7492
23. D. Barache, J.P. Antoine, J.M. Dereppe, *The Continuous Wavelet Transform, an Analysis Tool for NMR Spectroscopy*, Journal of Magnetic Resonance, September 2009, Vol. 128 pp. 1-11, DOI:10.1006/jmre.1997.1214
24. A. Suvichakorn, H. Ratiney, A. Bucur, S. Cavassila, J.P. Antoine, *Toward a quantitative analysis of in vivo proton magnetic resonance spectroscopic signals using the continuous Morlet wavelet transform*, Measurement Science and Technology (2009), Vol. 20, DOI:10.1088/0957-0233/20/10/104029
25. A. Schuck Jr. , C. Lemke, A. Suvichakorn, J.P. Antoine, *Analysis of Magnetic Resonance Spectroscopic Signals with Data-based Autocorrelation Wavelets*, IEEE Engineering in Medicine and Biology Society, August 2010, pp. 855-858, DOI:10.1109/IEMBS.2010.5628034
26. S.W. Provencher, *Automatic quantitation of localized in vivo 1H spectra with LCMoDel*, NMR in Biomedicine, June 2001, Vol. 14 pp. 260-264, DOI:10.1002/nbm.698
27. A. Naressi, C. Couturier, J.M. Devos, M. Janssen, C. Mangeat, R. de Beer, D. Graveron-Demilly, *Java-based graphical user interface for the*

- MRUI quantitation package*, Magnetic Resonance Materials in Physics, Biology and Medicine (2001), Vol. 12 pp. 141-152, DOI:10.1007/BF02668096
28. D. Stefan, F. Di Cesare, A. Andrasescu, E. Popa, A. Lazariiev, E. Vescovo, O. Strbak, S. Williams, Z. Starcuk, M. Cabanas, D. van Ormondt, D. Graveron-Demilly, *Quantitation of magnetic resonance spectroscopy signals: The jMRUI software package*, Measurement Science and Technology (2009), Vol. 20, DOI:10.1088/0957-0233/20/10/104035
 29. J. Pouillet, D.M. Sima, A.W. Simonetti, B. De Neuter, L. Vanhamme, P. Lemmerling, S. Van Huffel, *An automated quantitation of short echo time MRS spectra in an open source environment: AQSES*, NMR in Biomedicine, August 2007, Vol. 20 pp. 493-504, DOI:10.1002/nbm.1112
 30. L.N. Manganas, X. Zhang, Y. Li, R.D. Hazel, S.D. Smith, M.E. Wagshul, F. Henn, H. Benveniste, P.M. Djuric, G. Enikolopov, M. Maletic-Savatic, *Magnetic Resonance Spectroscopy Identifies Neural progenitor Cells in the Live Human Brain*, Science, 9 November 2007, Vol. 318 no. 5852 pp. 980-985, DOI: 10.1126/science.1147851
 31. Jeffrey C. Hoch et al., *Comment on: Magnetic Resonance Spectroscopy Identifies Neural Progenitor Cells in the Live Human Brain*, Science, August 2008 , Vol. 321 p. 640b, DOI:10.1126/science.1153058
 32. Seth D. Friedman, *Comment on: Magnetic Resonance Spectroscopy Identifies Neural Progenitor Cells in the Live Human Brain*, Science, August 2008, Vol. 321 p. 640c, DOI:10.1126/science.1153484
 33. Jacobus F. A. Jansen et al., *Comment on: Magnetic Resonance Spectroscopy Identifies Neural Progenitor Cells in the Live Human Brain*, Science, August 2008, Vol. 321 p. 640d, DOI:10.1126/science.1153997
 34. Z. Dong, W. Dreher, D. Leibfritz, B.S. Peterson, *Challenges of Using MR Spectroscopy to Detect Neural Progenitor Cells In vivo*, American Journal of Neuroradiology, June-July 2009, Vol. 30 pp. 1096-1101, DOI:10.3174/ajnr.A1557
 35. K. Govindaraju, K. Young, A.A. Maudsley, *Proton NMR chemical shifts and coupling constants for brain metabolites*, NMR in Biomedicine (2000), Vol. 13 pp. 129-153, DOI:10.1002/1099-1492(200005)13:3<129::AID-NBM619>3.0.CO;2-V

7. Anexo I: GUI para MATLAB

Todos los métodos SVD que se han empleado en este proyecto se han reunido en un interfaz gráfico (GUI) de MATLAB, para facilitar su uso y el acceso a los resultados de sus análisis. En la implementación que se ha llevado a cabo se ha intentado mejorar la interacción entre el usuario y los métodos SVD, permitiendo determinar un mayor número de parámetros del ajuste antes de proceder al mismo.

Como habitualmente se desconoce a priori el número de señales presentes bajo estudio, todos los métodos se han implementado de tal manera que permitan visualizar una representación de los valores singulares de la matriz de datos para poder realizar una estimación del mismo. Para ello se ha implementado una opción mediante la cuál el método pueda trabajar sin necesidad de especificarle un número de señales determinado al inicio. También, en todos ellos se introducen como parámetros la frecuencia de transmisión del aparato con la que se ha realizado la adquisición de datos y la frecuencia de la señal tomada como referencia para realizar la misma, para que estos métodos puedan usarse para analizar los datos MRS independientemente del aparato y método por el que se hayan adquirido. También se ha de especificar que datos se van a analizar y el intervalo de muestreo de adquisición de los mismos. En el método LPSVD además se puede determinar el orden del polinomio característico para determinar los polos de la señal. En el método HSVD se puede indicarr el tamaño de la matriz de datos y el número de señales esperado en el espectro. En el método SELF-SVD se puede indicar el número de señales esperado y el rango frecuencial sobre el que se quiere llevar a cabo el estudio, pudiendo elegir también el análisis sobre el espectro completo. En todos los métodos se obtienen a la salida los resultados de las estimaciones del ajuste y la representación gráfica de los valores singulares y del ajuste, tanto en el dominio del tiempo como el de la frecuencia. En el dominio frecuencial se puede elegir representar los resultados tanto en Hz como en ppm. Por último existe la opción de indicar que las representaciones gráficas de los resultados aparezcan en figuras independientes de la aplicación. Esto responde a que si más adelante se quieren recopilar, de esta manera se consigue que presenten una mejor resolución. El aspecto que presenta este GUI de MATLAB es el siguiente:

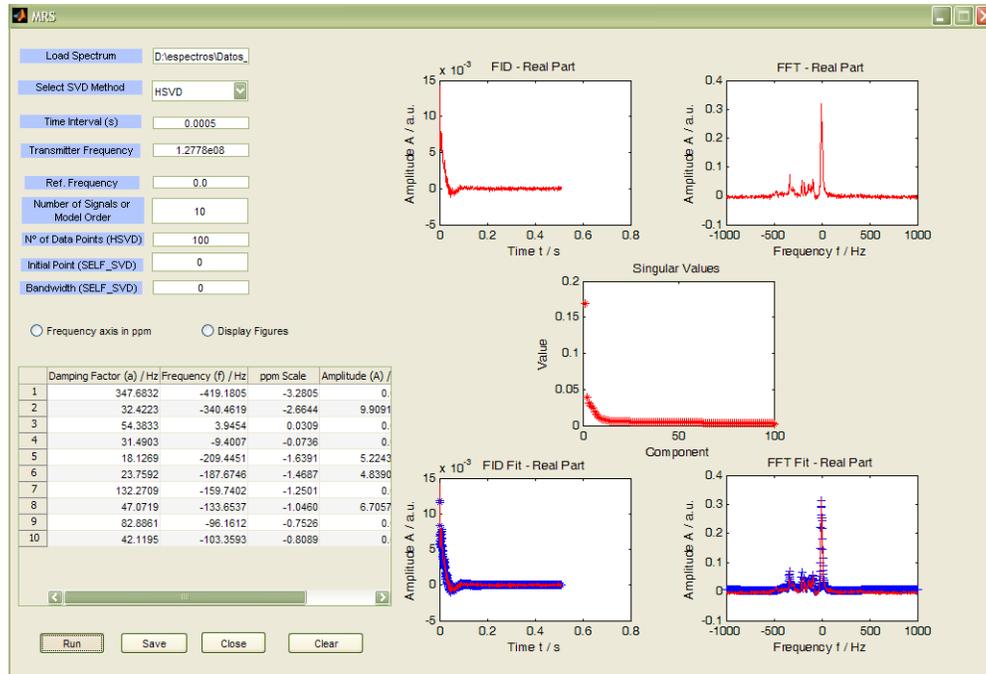


Figura 15: Muestra de la Interfaz de GUI creado en MATLAB para analizar datos MRS.

Como se puede ver en la imagen, el interfaz gráfico consta de una parte en la que el usuario puede introducir los parámetros necesarios para realizar el análisis, una tabla en las que se muestran los resultados del mismo y una sección gráfica en las que se representan la muestra de datos y los resultados obtenidos.

Una vez indicados los parámetros, se puede ejecutar la aplicación para llevar a cabo el análisis. En la tabla aparecen las estimaciones de los parámetros de las componentes y en el lado derecho aparecen las representaciones gráficas. Se muestra el aspecto que presentan los datos bajo estudio, tanto en el dominio del tiempo como en el de la frecuencia, los valores singulares para poder estimar el número de señales presentes, y los ajustes de la señal original, de nuevo tanto en el dominio del tiempo como en el de la frecuencia. Una vez concluido el análisis, se pueden guardar los resultados o eliminarlos para proceder a un nuevo análisis. Por último existe un botón para finalizar la aplicación.