



**Universidad
Zaragoza**

Trabajo Fin de Grado

**Implementación y testeo del
método de alineamiento temporal
de señales biomédicas basado en
autovalores**

Autor

Javier Preciado Garbayo

Director

Pablo Laguna Lasasa

Escuela de Ingeniería y Arquitectura

2015

RESUMEN

Implementación y testeo del método de alineamiento temporal de señales biomédicas basado en autovalores

Una aspecto importante a la hora de obtener una estimación robusta de una señal biológica repetitiva es el alineamiento de los diferentes segmentos de la señal donde encontramos la periodicidad, para su posterior promediado.

Actualmente existen métodos de alineamiento eficaces basados en la correlación. En este TFG proponemos un nuevo método de alineamiento basado en los autovalores y autovectores de la matriz de autocorrelación. A partir de estos nuevos datos construiremos una función de coste. Será nuestra forma de ver cual de los retardos aplicados corrige nuestra situación inicial de no alineamiento, ya que la función se maximiza a medida que nos acercamos al alineamiento ideal.

Más adelante se considera una alternativa en el cálculo de la matriz de autocorrelación, un cambio de orientación entre filas y columnas que justificaremos a lo largo del capítulo 2. De esta manera obtenemos una solución que da con los desalineamientos de forma más eficiente en cálculo.

Tras comentar de forma breve las bases del método de autovalores procederemos a compararlo en prestaciones con métodos basados en la correlación, como son el método de Woody y el método de correlación. Los parámetros escogidos serán los siguientes: diferentes SNR (relación señal a ruido), diferente número de realizaciones y diferente rango de desalineamiento. Veremos de forma adicional variaciones en la amplitud de la onda a alinear (anexo).

Finalmente, veremos como responde cada método ante igualdad de condiciones, observando el error en muestras del desalineamiento real aplicado y el estimado por cada método, así como su rango intercuartil. Todo esto se justificará en el capítulo de resultados.

Agradecimientos

Me gustaría agradecer a mi director, Pablo, por su paciencia, esfuerzo y dedicación. Durante el tiempo que se ha ido realizando este trabajo me ha ayudado a progresar en mi forma de trabajar y a conseguir una mayor madurez.

A todos profesores que participaron en mi formación y a los miembros del grupo de investigación BSICoS, Biomedical Signal Interpretation & Computational Simulation, en especial a mis compañeros de laboratorio, Alba y Jesús, que para cualquier problema o duda me encontraron solución.

No puedo olvidarme de mis compañeros de clase, en especial, mi grupo de "Esmochados", hace ya unos años que empecé esta aventura en solitario, ahora que estamos acabando ya tengo una magnífica familia aquí en Zaragoza :)

A mi familia, en especial mis padres y abuelos, mi hermano Ángel y mi hermana Cristina, por su esfuerzo realizado y apoyo durante estos años y toda mi vida.

A Beatriz Giraldo de la UPC por compartir las señales con las que fue realizado este TFG, BIOSPIN - Biomedical Signal Processing and Interpretation (IBEC) Departamento de Ingeniería de Sistemas, Automática e Informática Industrial (ESAI, UPC).

*Un abrazo enorme a todos ellos,
Javier*

Índice general

1. Introducción	7
1.1. Motivación	8
1.1.1. Interés Fisiológico	10
1.2. Objetivos	11
1.3. Contexto	11
1.4. Organización de la memoria	11
2. Métodos de alineamiento a usar	12
2.1. Alineamiento mediante autovalores	13
2.1.1. Matriz de autocorrelación, para relación intra-realización	13
2.1.2. Matriz de autocorrelación, para relación inter-realización	20
2.2. Alineamiento mediante método de correlación	25
2.3. Alineamiento mediante método de Woody	30
3. Resultados	33
3.1. Descripción	34
3.1.1. Método de alineamiento mediante autovalores	34
3.1.2. Método de correlación y método de Woody	43
4. Conclusiones y líneas futuras	54
4.1. Líneas futuras	56
Referencias	57
Índice de figuras	59
Índice de cuadros	63
A. Estimador de máxima verosimilitud	64
B. Entorno alineamiento	65
C. Variabilidad de amplitud	67

Capítulo 1

Introducción

Actualmente en un mundo donde todo se dirige a la era digital el campo de la biomedicina no es una excepción. Cada vez tenemos mayor capacidad para observar los diferentes procesos que tienen lugar en el cuerpo humano. Existe la necesidad de transmitir todos estos datos a otros entornos como por ejemplo entornos médicos para que ellos realicen una correcta interpretación clínica de los fenómenos que ocurren en nuestro organismo. Para ello todas estas señales las debemos cuantificar, codificar, transmitir e interpretar con la mayor precisión posible.

Debemos tener en cuenta que no captaremos de forma ideal estas señales, tendremos de forma aditiva varias fuentes de ruido en nuestras medidas, como puede ser el ruido producido por el cuerpo humano, más concretamente, interferencias de señales generadas por otros órganos adyacentes o de la propia red eléctrica que alimenta nuestros sensores de captación o incluso por movimientos corporales. Es importante que la información llegue lo más fideligna posible al intérprete. Por ello necesitaremos de etapas previas que procesen adecuadamente estas medidas.

Por ejemplo las señales desencadenadas del sistema nervioso u órganos vitales como el corazón pueden darnos información acerca de trastornos del sueño o enfermedades cardiovasculares. Una de las formas de obtener alguna de estas señales, concretamente señales que tienen un comportamiento pseudorepetitivo (una misma forma de onda se repite de forma aproximada en el tiempo) es la segmentación de esta señal y su posterior medida sobre un promedio de repeticiones en lugar de sobre una única realización. Dependiendo de las características de la señal observada podremos tener una señal más realzada y con menor ruido.

Es importante en esto último que las señales a promediar se encuentren bien alineadas. Y un método de alineamiento podría ser el que proponemos en este TFG.

1.1. Motivación

Hoy en día el desarrollo de la ingeniería junto con la medicina es uno de los pilares más importantes de la sociedad. Cada vez vivimos en una sociedad con mayor número de facilidades debido al avance en el desarrollo de la tecnología donde gozamos de comodidades o delegamos habilidades y tareas a otros dispositivos como por ejemplo smartphones que simplifican nuestra vida. Esto nos da facilidades para separar las cosas cotidianas de nuestro día a día y centrarnos en mayor medida en las tareas más relevantes o simplemente para disponer de más tiempo libre.

De la misma forma que vemos avances en el desarrollo de nuestra rutina, en el entorno sanitario también ha habido grandes avances en las últimas décadas. Gracias a nuevos equipos que permiten tratamientos más eficaces o fármacos y vacunas que aumentan la esperanza de vida de la población.

Debido a la necesidad de dar diagnósticos cada vez más exactos o anticiparnos en mayor medida a enfermedades en un entorno tan complejo como es el ser humano, una fuente posible de mejora se puede encontrar en la obtención de estas señales biológicas del sujeto de forma más exacta y con menor distorsión. Para ello podemos incluir alguna etapa de forma adicional en el procesado de estas señales.

Algunas de estas etapas de procesado hacen posible que las medidas sean más fiables y eviten pasar por alto aspectos importantes para el diagnóstico del paciente debido a que parte de la señal de interés quede distorsionada, o embebida en el ruido.

Centrándonos más en nuestro problema concreto, se trata de una aportación a la obtención de unas medidas robustas a partir de una señal pseudorepetitiva (con un ruido aleatorio superpuesto en una forma de onda determinista) mediante un nuevo alineamiento a usar, que tiene su origen en la literatura [1] y [2]. Proponemos un nuevo método que pueda ser una alternativa más eficiente a otros ya existentes como son el método de correlación o el método de Woody.

En el desarrollo de este TFG comprobaremos como de fiable es este nuevo método frente a los ya existentes. De resultar más o menos eficaz que los métodos existentes explicaremos, gracias a los resultados obtenidos en simulación, cual es la fuente que mejora o empeora respecto a los otros métodos de alineamiento.

Veremos como en el caso de tener un alineamiento que diste de la situación ideal, obtendremos una penalización en nuestra señal de interés en forma de filtrado paso bajo. Este fenómeno lo ilustraremos en las siguientes figuras.

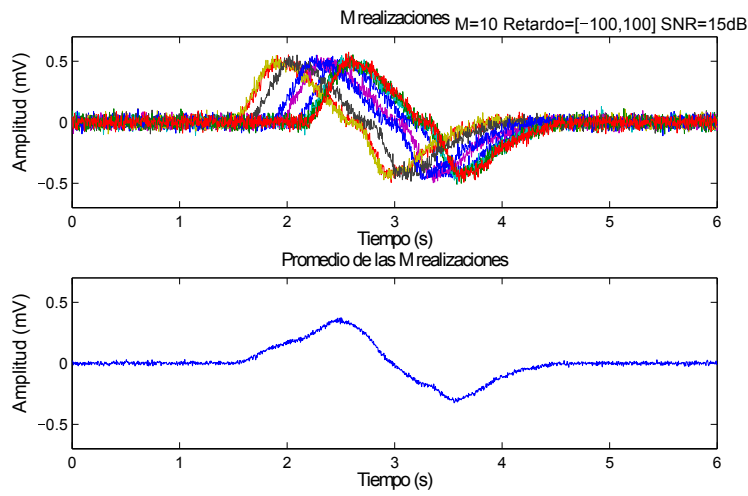


Figura 1.1: Realizaciones mal alineadas

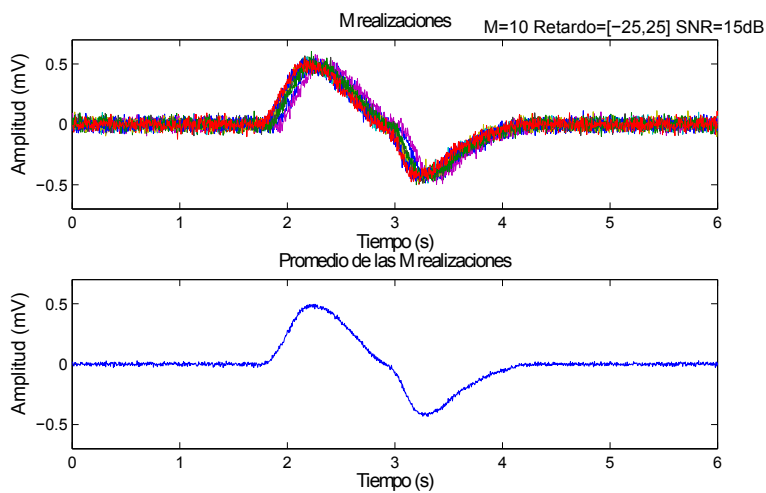


Figura 1.2: Realizaciones próximas a estar alineadas

Donde M es el número de realizaciones, *Retardo* nos indica el rango de los posibles desalineamientos en muestras de cada realización y SNR es la relación en dB entre potencia de señal y potencia de ruido.

1.1.1. Interés Fisiológico

Una situación favorable y de interés para poder aplicar nuestro método de alineamiento es encontrar señales biológicas que tengan un comportamiento repetitivo en el tiempo. De esta manera podremos segmentar estas señales y tratar de obtener una señal más limpia de ruido y de mayor calidad tras realizar el promedio.

En el cuerpo humano encontramos varios procesos biológicos que desencadenan señales que cumplen esta particularidad que nos interesa. El corazón desencadena señales como el electrocardiograma (ECG) donde se observa la actividad eléctrica que tiene lugar en el corazón, los pulmones al realizar el ciclo respiratorio donde vemos la variación de la impedancia de la caja torácica (debido al volumen de aire presente en los pulmones), la actividad cerebral que se produce ante un estímulo externo detectado por los sentidos (potenciales evocados).

De los ejemplos citados anteriormente escogeremos la señal correspondiente al ciclo respiratorio, siguiendo los trabajos ya desarrollados en [1] y [2].

Vemos a continuación una figura donde se ilustra el comportamiento repetitivo del ciclo respiratorio.

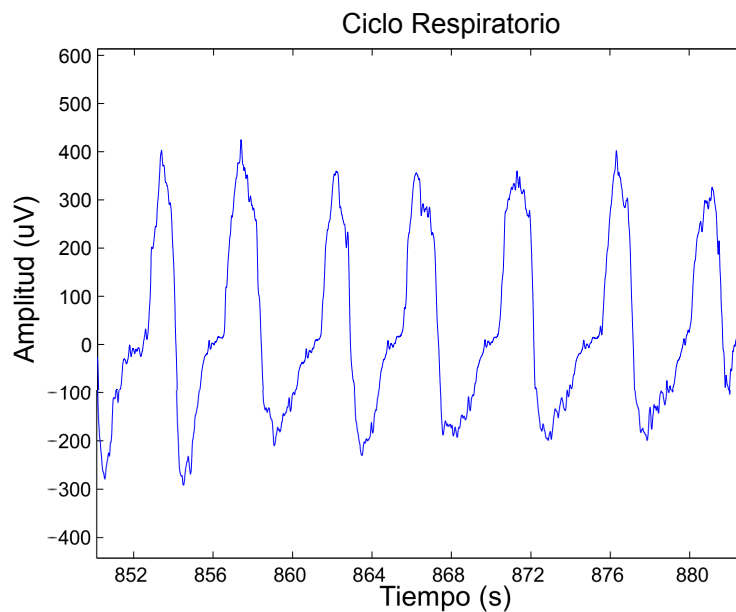


Figura 1.3: Fragmento del ciclo respiratorio de un paciente sano, muestreado a 250 Hz y con un periodo aproximadamente de 2,5 segundos.

1.2. Objetivos

Dada la importancia para el análisis clínico, las señales captadas por los sensores serán posteriormente interpretadas por personas cualificadas, como por ejemplo médicos y por ello deben ser tratadas, si ello es posible, para realzar la información clínica.

Por ello pretendemos mejorar la calidad de la señal a estudiar, es decir, mejorar su relación señal a ruido aprovechando su aparición de forma repetitiva. Para ello promediaremos a lo largo de los segmentos, cancelando un ruido que deberá tener media nula. Necesitaremos de un buen alineamiento previo al promediado.

El objetivo principal es proponer y estudiar un método de alineamiento eficiente en cálculo. Además compararemos los resultados obtenidos de este método de alineamiento (basado en autovalores y autovectores) con otros métodos de alineamiento ya existentes (método de correlación y método de Woody).

1.3. Contexto

Este TFG se ha desarrollado en el Departamento de Ingeniería Electrónica y Comunicaciones (DIEC) de la Universidad de Zaragoza. Además este trabajo ha sido desarrollado junto al grupo *Biomedical Signal Interpretation and Computational Simulation Group* (BSICoS) Aragón Institute of Engineering Research (I3A). Zaragoza University CIBER-BBN, Instituto de Salud Carlos III. Spanish Ministry of Science and Innovation.

1.4. Organización de la memoria

Explicamos a continuación de forma breve los puntos destacados de esta memoria:

- En el capítulo 2 explicamos en detalle las bases de cada método de alineamiento (de autovalores, de correlación y de Woody) que posteriormente implementaremos en la plataforma Matlab para su testeo.
- En el capítulo 3 redactaremos acerca de los resultados obtenidos de la simulación realizada en Matlab. Nos centraremos en comentar aspectos relativos al error entre la estimación obtenida de los retardos y los que realmente fueron aplicados. Para ello habremos hecho uso de varios parámetros en simulación que explicaremos en este capítulo.
- En el último capítulo de esta memoria trataremos de destacar los resultados más relevantes de este TFG. También consideraremos de forma aditiva la posibilidad de una versión iterativa del método de autovalores para las situaciones de mayor error.
- Para finalizar destacaremos los aspectos más importantes de este TFG. También nombraremos posibles mejoras y aspectos a considerar para futuras ampliaciones de este TFG.

Capítulo 2

Métodos de alineamiento a usar

La señal repetitiva la podemos modelar en cada realización, i , como una parte repetitiva $s(n)$ más un ruido aleatorio $v_i(n)$ independiente de cada realización i .

$$x_i(n) = s(n - \theta_i) + v_i(n), \quad n = 0, \dots, N - 1; \quad i = 1, \dots, M; \quad (2.1)$$

donde cada realización consta de nuestra señal de interés $s(n)$, con una energía E_s y forma fijas, que en este TFG se corresponde con un segmento de señal respiratoria de un paciente sano, muestreado a 250 hz. El desalineamiento θ_i es el retardo perteneciente a la realización i -ésima que debemos estimar mediante nuestro método de alineamiento. Consideraremos que la media de los desalineamientos de las M realizaciones es nula y se distribuye de manera uniforme. Además tenemos un conjunto de M realizaciones de N muestras cada una de ellas. En cada realización tenemos superpuesto un ruido aleatorio $v_i(n)$ que supondremos de distribución normal y media nula.

El conjunto de realizaciones lo podemos denotar mediante la siguiente matriz \mathbf{X} , formada por vectores columna \mathbf{x}_i definidos a continuación.

$$\mathbf{x}_i = \begin{bmatrix} x_i(\Delta) \\ x_i(\Delta + 1) \\ \vdots \\ x_i(\Delta + N - 1) \end{bmatrix} \quad (2.2)$$

En las columnas o realizaciones tendremos un instante de comienzo de las muestras diferente debido al instante de inicio del muestreo Δ . De todas las muestras disponibles de cada realización, N muestras formarán parte de nuestra observación. Vemos el aspecto final de la matriz \mathbf{X} que agrupa las M realizaciones.

$$\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_M] \quad (2.3)$$

2.1. Alineamiento mediante autovalores

Este método de alineamiento hace uso de los autovalores y autovectores de la matriz de autocorrelación del proceso.

Explicamos brevemente que es un vector y un valor propio (autovalor y autovector). Un vector propio es aquel que al multiplicar una matriz por este vector nos da el mismo vector y un escalar. Este escalar es el autovalor asociado al vector propio

$$\mathbf{X}\mathbf{v} = \lambda\mathbf{v} \quad (2.4)$$

donde \mathbf{v} se corresponde con un vector propio de la matriz \mathbf{X} y λ es el valor propio asociado al vector propio \mathbf{v} .

A continuación introducimos un método desarrollado en la literatura, en el cual se centra este TFG. Se puede ver que tiene una gran demanda computacional en [1], como queda patente en la implementación aquí hecha y que se explica en 2.1.1. Lo implementamos también de forma más eficiente en el apartado 2.1.2, donde nos ayudaremos en una equivalencia matemática que se puede encontrar en la literatura, [2]. Mencionamos que esta última fuente se encuentra en revisión, aun no esta publicada en ninguna revista científica.

2.1.1. Matriz de autocorrelación, para relación intra-realización

Para la obtención de la matriz de correlación podemos realizar las siguientes operaciones.

$$\mathbf{R}_x \triangleq E[\mathbf{x}\mathbf{x}^T] \approx \frac{1}{M}\mathbf{X}\mathbf{X}^T \quad (2.5)$$

Donde \mathbf{X} es la matriz formada por el conjunto de realizaciones (2.3) y \mathbf{x} un vector de una realización genérica (2.2). Nosotros haremos uso de la estimación de \mathbf{R}_x mediante $\hat{\mathbf{R}}_x = \frac{1}{M}\mathbf{X}\mathbf{X}^T$. Cabe destacar que la matriz de autocorrelación tiene $\dim(N \times N)$ y además se tiene una autocorrelación intra-realización, fijando el eje temporal miramos relación entre muestras de la misma realización.

Podemos ver el aspecto de cada uno de los coeficientes de la matriz $\hat{\mathbf{R}}_x = \frac{1}{M}\mathbf{X}\mathbf{X}^T$ de forma más detallada.

$$\mathbf{X}\mathbf{X}^T = \begin{pmatrix} \sum_{k=1}^M x_k(0)x_k(0) & \sum_{k=1}^M x_k(0)x_k(1) & \cdots & \sum_{k=1}^M x_k(0)x_k(N-1) \\ \sum_{k=1}^M x_k(1)x_k(0) & \sum_{k=1}^M x_k(1)x_k(1) & \cdots & \sum_{k=1}^M x_k(1)x_k(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^M x_k(N-1)x_k(0) & \cdots & \cdots & \sum_{k=1}^M x_k(N-1)x_k(N-1) \end{pmatrix} \quad (2.6)$$

Definimos ahora los siguientes vectores, nuestra señal de interés \mathbf{s} y el ruido de una realización genérica \mathbf{v}_i . Cada realización será, $\mathbf{x}_i = \mathbf{s} + \mathbf{v}_i$

$$\mathbf{s} = [s(0) \ s(1) \ \cdots \ s(N-1)]^T \quad (2.7)$$

$$\mathbf{v}_i = [v_i(0) \ v_i(1) \ \cdots \ v_i(N-1)]^T \quad (2.8)$$

Una vez comentado el entorno que rodea a nuestro problema nos situamos primeramente en un caso de alineamiento perfecto, es decir, $\theta_i = 0 \ \forall i$, nos encontraremos en el siguiente caso:

$$E[\mathbf{xx}^T] = E[(\mathbf{ss}^T) + (\mathbf{sv}_i^T) + (\mathbf{v}_i\mathbf{s}^T) + (\mathbf{v}_i\mathbf{v}_i^T)] \quad (2.9)$$

Tras desarrollar los productos y aplicar el operador esperanza obtenemos el siguiente resultado en (2.10). Debemos recordar que el ruido es de media nula e incorrelado consigo mismo, por esto último tenemos una matriz diagonal como segundo término en (2.10):

$$\mathbf{R}_x = \mathbf{ss}^T + \sigma_v^2 \mathbf{I} \quad (2.10)$$

A continuación obtenemos los autovalores y autovectores

$$\lambda_i = \begin{cases} E_s + \sigma_v^2, & i = 1; \\ \sigma_v^2, & i = 2, \dots, N \end{cases} \quad (2.11)$$

donde E_s , la energía de \mathbf{s} , se corresponde con el producto $\mathbf{s}^T\mathbf{s}$. Los autovectores en este caso son \mathbf{s} y cualquier conjunto de vectores pertenecientes al espacio N-dimensional ortogonal a \mathbf{s} y ortogonales entre ellos mismos.

Vemos además que en estos autovalores y autovectores no encontramos información relevante en cuanto al desalineamiento. Cosa lógica debido a que las realizaciones ya estaban perfectamente alineadas.

En (2.12) se define la siguiente función ratio de autovalores de la matriz de autocorrelación (2.5), será nuestra función de costes, con ella nos daremos cuenta en que medida alineamos de forma correcta las diferentes realizaciones.

$$\Lambda = \frac{\lambda_1}{\sum_{i=2}^N \lambda_i} \quad (2.12)$$

Al final de esta sección 2.1.1 explicaremos porque la motivación de construir esta función ratio de autovalores de esta manera.

Nuestro caso real que nos ocupa no será encontrarnos las realizaciones ya alineadas sino que tendremos un cierto grado de desalineamiento en cada una de las realizaciones, es decir, $\theta_i \neq 0 \ \forall i$ con gran probabilidad. Para este caso resolvemos (2.5).

Antes deberemos realizar una aproximación mediante el desarrollo en serie de Taylor. Para ello usaremos nuestra función (2.1) en el dominio continuo del tiempo. Centramos el desarrollo en t y supondremos que θ_i es de media nula y de un valor pequeño para tener una aproximación sin demasiado error del desarrollo en serie de Taylor. De esta manera tendremos fuera de la dependencia temporal a la variable θ .

$$s(t - \theta_i) \approx s(t) - \theta_i s'(t) = \hat{s}_T(t - \theta_i), \quad (2.13)$$

Y por lo tanto la expresión final derivada de (2.1) será la siguiente:

$$x_i(t) \approx s(t) - \theta_i s'(t) + v_i(t), \quad (2.14)$$

A continuación podremos ver varias gráficas para apreciar el grado de distorsión provocado en $s(t)$ ó $s(n)$ debido a diferentes θ_i .

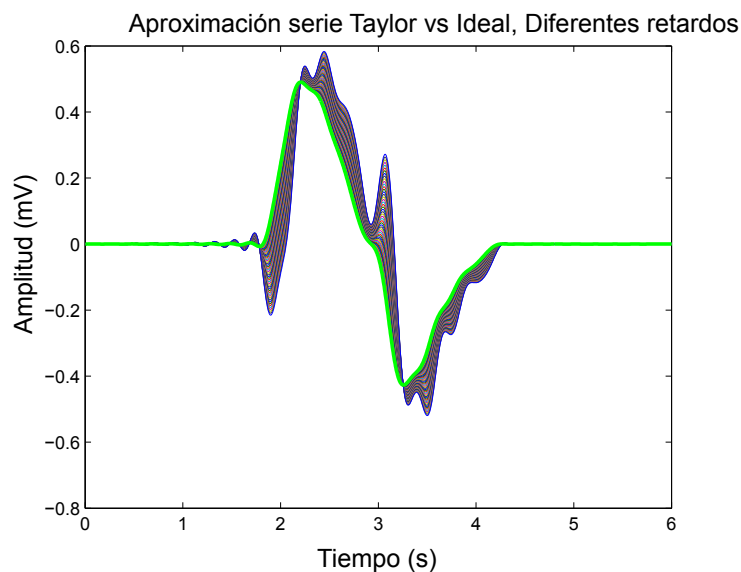


Figura 2.1: Desalineamiento de 4 a 160 ms, en verde la señal original $s(t)$ y alineada. En diferentes colores $\hat{s}_T(t)$ para $\theta_i = 4, 8, \dots, 160$ ms

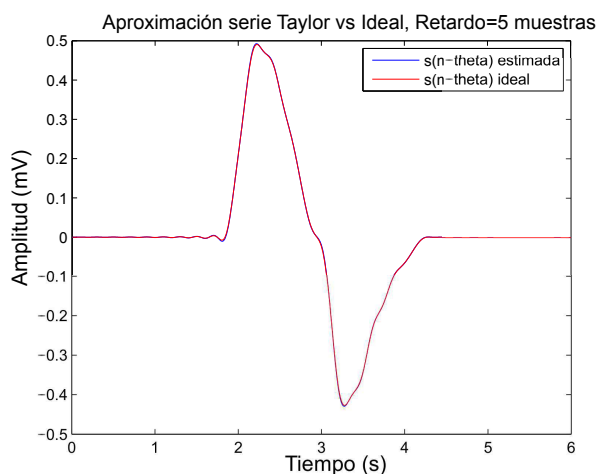


Figura 2.2: Desalineamiento de 20 ms ó 5 muestras, en rojo la señal original desplazada $s(t - 0,02)$. En azul $\hat{s}_T(t)$ para $\theta_i = 20$ ms

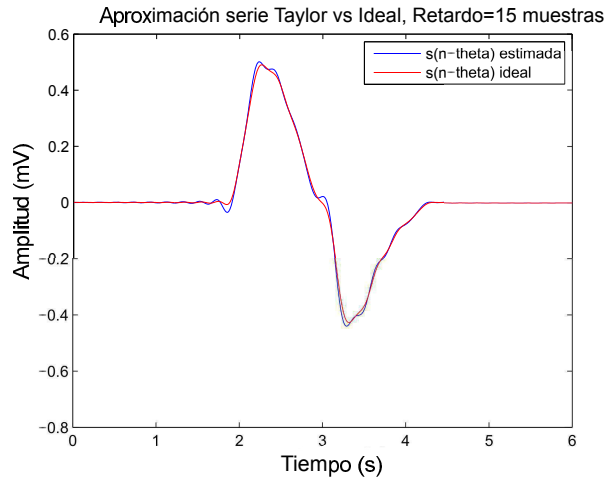


Figura 2.3: Desalineamiento de 60 ms ó 15 muestras, en rojo la señal original desplazada $s(t - 0,06)$. En azul $\hat{s}_T(t)$ para $\theta_i = 60$ ms

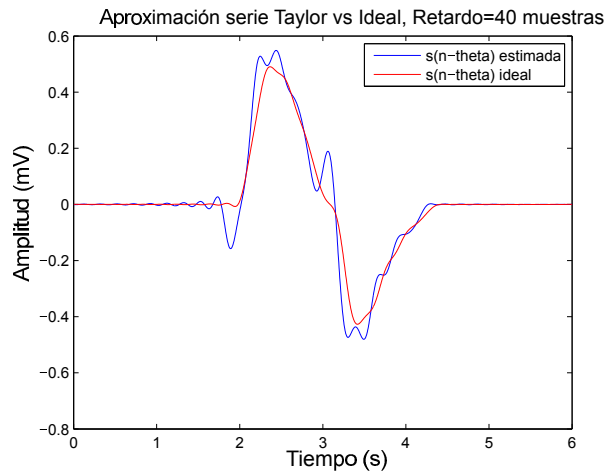


Figura 2.4: Desalineamiento de 160 ms ó 40 muestras, en rojo la señal original desplazada $s(t - 0,16)$. En azul $\hat{s}_T(t)$ para $\theta_i = 160$ ms

Podemos comprobar el grado de distorsión en la forma de onda de $s(t)$. En las gráficas vemos representados casos discreto en tiempo. Teniendo en cuenta la frecuencia de muestreo, 250 Hz, estamos observando desalineamiento en tiempo desde 4 ms para 1 muestra de desalineamiento hasta 160 ms para las 40 muestras (Figura 2.1). Vemos además caso concretos de desalineamientos de varias muestras, (Figuras 2.2 a 2.4).

Habiendo concretado ya las consecuencias por aproximar mediante la serie de Taylor de 2 términos procederemos a obtener la matriz de autocorrelación.

$$\begin{aligned} \mathbf{R}_x &= E[\mathbf{xx}^T] = E[(\mathbf{s} - \theta_i \mathbf{s}' + \mathbf{v}_i)(\mathbf{s} - \theta_i \mathbf{s}' + \mathbf{v}_i)^T] = \\ &= E[\mathbf{ss}^T - \theta_i \mathbf{ss}'^T + \mathbf{sv}_i^T - \theta_i \mathbf{s}'\mathbf{s}^T + \theta_i^2 \mathbf{s}'\mathbf{s}'^T - \theta_i \mathbf{s}'\mathbf{v}_i^T + \mathbf{v}_i \mathbf{s}^T - \theta_i \mathbf{v}_i \mathbf{s}'^T + \mathbf{v}_i \mathbf{v}_i^T] \end{aligned} \quad (2.15)$$

Teniendo en cuenta que $\mathbf{s} \perp \mathbf{s}'$ y que el ruido \mathbf{v}_i es de media nula $E[\mathbf{v}_i] = 0$ y esta incorrelado tanto con la señal de interés así como con su derivada. Además \mathbf{s} es ortogonal con \mathbf{s}' . Realizamos la comprobación de esta última, $\mathbf{s} \perp \mathbf{s}'$.

$$\int_{-\infty}^{\infty} s(t) s(t)' dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\Omega) (-j\Omega) S^*(\Omega) d\Omega = -\frac{1}{2\pi} \int_{-\infty}^{\infty} j\Omega |S(\Omega)|^2 d\Omega = 0, \quad (2.16)$$

Debemos tener en cuenta que nuestra señal $s(t)$ es de duración finita y real, por lo tanto su transformada de Fourier tendrá simetría hermítica.

$$S(j\Omega) = S^*(-j\Omega) \quad (2.17)$$

Quitando términos que se anulan al realizar productos obtenemos la siguiente expresión

$$\mathbf{R}_x = \mathbf{ss}^T + \sigma_\theta^2 \mathbf{s}'\mathbf{s}'^T + \sigma_v^2 \mathbf{I}, \quad (2.18)$$

y los autovalores son los siguientes

$$\lambda_i = \begin{cases} E_s + \sigma_v^2, & i = 1; \\ \sigma_\theta^2 E_{s'} + \sigma_v^2, & i = 2; \\ \sigma_v^2, & i = 3, \dots, N. \end{cases} \quad (2.19)$$

donde $E_{s'} = \mathbf{s}'^T \mathbf{s}'$ y $E_s = \mathbf{s}^T \mathbf{s}$.

Los autovectores que nos dan estos autovalores son, \mathbf{s} , \mathbf{s}' y cualquier conjunto de vectores perteneciente al espacio N-dimensional ortogonal a los dos anteriores y ortogonales entre ellos mismos.

Podemos construir la siguiente función ratio a partir de la definida en (2.12). Definimos primeramente $\boldsymbol{\theta}$, el vector de retardos

$$\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \dots \ \theta_M]^T, \quad (2.20)$$

ahora la función ratio,

$$\Lambda(\boldsymbol{\theta}) = \frac{\lambda_1(\boldsymbol{\theta})}{\sum_{i=2}^N \lambda_i(\boldsymbol{\theta})} \approx \frac{E_s}{\sigma_\theta^2 E_{s'} + (N-1)\sigma_v^2} \quad (2.21)$$

Donde hemos aproximado $E_s + \sigma_v^2$ por E_s ya que $\lambda_1 \gg \lambda_N$ o $E_s \gg \sigma_v^2$ debido a que el ruido es mucho menor que la señal.

Volviendo con la función de coste, $\boldsymbol{\theta}$ será un vector columna donde se encuentra el desalineamiento de cada una de las M realizaciones. Asumiremos además que $s(n - \theta_i)$ esta totalmente contenida dentro de la observación, es decir, el

valor de Δ es de un valor pequeño en comparación de la duración de $s(n)$. Añadido a esto asumiremos que el desalineamiento en muestras es pequeño también en términos de la duración de $s(n)$. Teniendo en cuenta ambos efectos, Δ y θ_i , no se contempla un desalineamiento mayor de 100 muestras, correspondientes con 0,4 segundos. (Un pulso de presión de sangre del corazón tarda menos de 0,4 segundos en llegar a las yemas de los dedos de las manos [3]), nuestro caso es un ciclo respiratorio donde medimos el flujo de una duración media próxima a 2,5 segundos en reposo. No se contempla grandes variaciones en el periodo debido a que la situación es próxima a ser estacionaria y por tanto habrá poco ruido y poca imprecisión en el alineamiento original.

Volviendo a la función, en este caso si que vemos una dependencia del desalineamiento en nuestra función de coste. Concretamente aparece un termino en forma de varianza del vector de deslaineamiento θ , σ_θ^2 . Podemos tratar de probar diferentes combinaciones de alineaminetos con el objetivo de hacer disminuir este valor y por la tanto aumentar el valor de la función de coste (2.21). Vemos entonces que maximizando esta función lo que realmente estamos haciendo es tratar de dar con los desalineaminetos, haciendo que la varianza de θ disminuya y aproximando a un valor fijo al conjunto de desalineamientos θ , en este caso a su media, aproximadamente nula por la distribución asumida del desalienamiento.

Concluimos entonces que mediante el resultado obtenido en (2.21), maximizando esta función estaremos dando con una combinación de retardos que alinea nuestras M realizaciones. El problema que podemos ver a priori es que solo podemos hacer prueba y error para medir σ_θ^2 . Esto quiere decir que deberemos probar todas las posibilidades para asegurarnos de tener el máximo global de la función de coste y no un máximo local. Dependiendo del valor de N y M esto puede ser un inconveniente en lo que a cómputo se refiere.

Vemos un par de figuras que describimos a continuación.

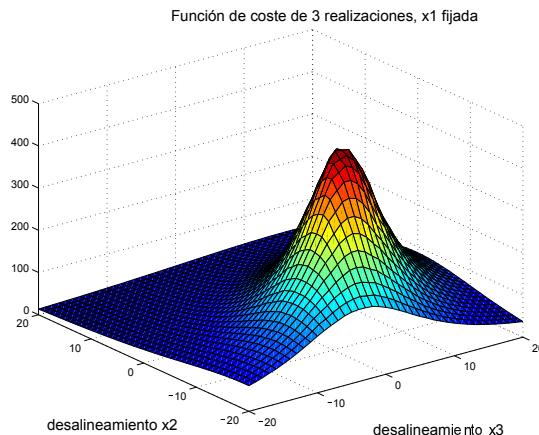


Figura 2.5: SNR de 25 dB

Mostramos varios ejemplo del resultado de realizar una simulación por medio de fuerza bruta, observamos el valor de la función de coste. En cada ejemplo tenemos 3 realizaciones: \mathbf{x}_1 , \mathbf{x}_2 y \mathbf{x}_3 , de las cuales una dejaremos fija y variaremos las otras 2 restantes.

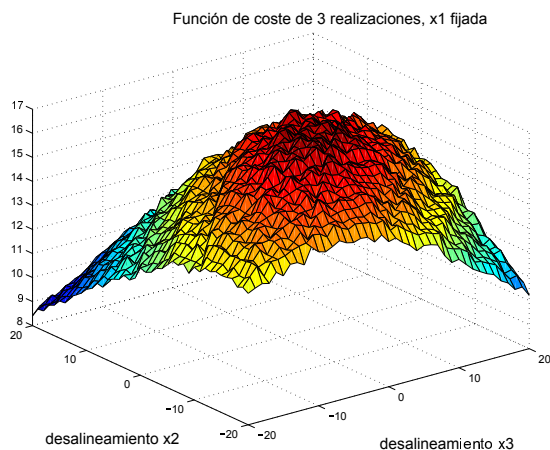


Figura 2.6: SNR de 10 dB

Lo primero que nos llama la atención es ver como a mayor relación señal a ruido en nuestras realizaciones, en este caso 10 ó 25 dB, vemos un máximo global más definido. Este nos indica a lo largo de las dos realizaciones de las tres que componen nuestra observación cual ha de ser el desplazamiento a aplicar en cada caso para corregir el desalineamiento.

En estos ejemplos ilustrados únicamente hemos considerado un rango de movilidad de $\Delta = 20$ posiciones o muestras sobre 2 realizaciones. De aumentar la búsqueda en más realizaciones, M , nuestro problema crece de forma exponencial como vemos con la siguiente expresión, $(2\Delta + 1)^M$ desplazamientos posibles a probar. Las cifras obtenidas conformen Δ o sobre todo M crecen nos motiva a buscar otras soluciones alternativas más eficientes.

2.1.2. Matriz de autocorrelación, para relación inter-realización

Motivados por la baja eficiencia en tiempo de cálculo de maximizar la función de coste vista en el apartado anterior se han estudiado alternativas más eficientes. Si consideramos una forma alternativa de la matriz de correlación inicial (2.21), recurriendo a la literatura [2].

Primeramente nos fijaremos en un método de alineamiento ya existente, concretamente el método de Woody. Comentamos brevemente en que consiste el método de Woody [4], es similar al método de correlación pero realizado de forma iterativa.

Inicialmente estimamos la señal a alinear mediante el promedio de todas las realizaciones desalineadas, correlando cada una de las M realizaciones con esta estimación. Observamos donde se encuentra el máximo en la correlación cruzada y tendremos la información necesaria para estimar el desalineamiento en esa realización. A medida que avancemos en iteraciones la señal estimada será más fiel a la ideal, dado que promediamos cada vez con señales mejor alineadas. Llegado ese momento nuestro algoritmo habrá alcanzado la convergencia (los desalineamientos calculados en ese momento serán los óptimos).

Se puede entender la relación entre nuestro método y el método de Woody con la siguiente consideración: el método de Woody es un método que involucra a la correlación entre el conjunto de realizaciones y un filtro adaptado de respuesta impulsional \mathbf{h} . Si asumimos que nuestras realizaciones se encuentran perfectamente alineadas, el filtro adaptado se puede obtener de maximizar la norma euclídea de la correlación cruzada al cuadrado de \mathbf{h} y todas las realizaciones \mathbf{x}_i ,

$$\|\mathbf{X}^T \mathbf{h}\|_2^2, \quad (2.22)$$

debemos tener en cuenta además que $\|\mathbf{h}\|_2$ sea distinto de cero. Es bien sabido que la expresión (2.22) el vector \mathbf{h} que maximiza se corresponde con el mayor autovector de $\mathbf{X}\mathbf{X}^T$ [5], y por lo tanto obtendremos el autovalor asociado a este autovector λ_1 ,

$$\max_{\mathbf{h}} \left(\|\mathbf{X}^T \mathbf{h}\|_2^2 \right) = \lambda_1. \quad (2.23)$$

Aunque el método de Woody emplea como aproximación de \mathbf{h} un promedio del conjunto de las M realizaciones, ensemble. La estimación de \mathbf{h} obtenida maximizando (2.23) no difieren mucho entre si en caso de encontrarse todas las realizaciones bien alineadas. Podemos ver entonces que el método de Woody trata de alinear las realizaciones y la forma de darnos cuenta de que procedemos de forma correcta y damos con el desalineamiento es ver que obtenemos el mayor autovalor de $\mathbf{X}\mathbf{X}^T$, reducir el valor de la expresión (2.23), quiere decir que nos alejamos de la situación de alineamiento ideal.

Vemos que a diferencia del método de autovalores donde la función de coste tiene en cuenta a todos ellos (2.21) en el método de Woody solo se busca el

mayor de ellos, λ_1 .

A continuación nos fijamos en una versión optimizada del método de Woody que mejora nuestro problema de alineamiento. El método de Woody de máxima verosimilitud o *ML-Woody* (*Maximum Likelihood Woody*).

Vemos la función de coste de este nuevo método,

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \Lambda_{\Delta}(\boldsymbol{\theta}), \quad (2.24)$$

donde la función-log de máxima verosimilitud es obtenida de [6]

$$\Lambda_{\Delta}(\boldsymbol{\theta}) = \sum_n \sum_{i=1}^M \sum_{k>i}^M x_k(n + \theta_k) x_i(n + \theta_i). \quad (2.25)$$

Esta función la podemos identificar sumando todos los elementos de la parte triangular superior de la siguiente matriz [7], llegamos ya a la que será nuestra matriz de autocorrelación inter-realización,

$$\hat{\mathbf{R}}_x^{\bullet} = \frac{1}{N} \mathbf{X}^T \mathbf{X} \quad (2.26)$$

donde $\hat{\mathbf{R}}_x^{\bullet}$ describe la relación inter-realización mientras que la matriz de autocorrelación vista en el primer apartado (2.1.2), $\hat{\mathbf{R}}_x = \frac{1}{M} \mathbf{X} \mathbf{X}^T$ describe la relación intra-realización.

Veamos de forma más detallada los diferentes coeficientes de la matriz inter-realización.

$$\mathbf{X}^T \mathbf{X} = \begin{pmatrix} \sum_{k=0}^{N-1} x_0(k)x_0(k) & \sum_{k=0}^{N-1} x_0(k)x_1(k) & \cdots & \sum_{k=0}^{N-1} x_0(k)x_M(k) \\ \sum_{k=0}^{N-1} x_1(k)x_0(k) & \sum_{k=0}^{N-1} x_1(k)x_1(k) & \cdots & \sum_{k=0}^{N-1} x_1(k)x_M(k) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} x_M(k)x_0(k) & \cdots & \cdots & \sum_{k=0}^{N-1} x_M(k)x_M(k) \end{pmatrix} \quad (2.27)$$

Observamos en un mismo instante temporal genérico k , pese a ello no quiere decir que el desalineamiento de cada una de las M realizaciones sea el mismo en cada caso podrá ser diferente.

Usando la aproximación (2.13), centrándonos en un instante temporal n a lo largo de las M realizaciones. En este caso podremos formar nuestra nueva matriz de autocorrelación que a diferencia de la primera matriz $\hat{\mathbf{R}}_x^{\bullet}$ tiene $\dim(M \times M)$ y busca la relación inter-realización a lo largo de los diferentes términos de la matriz (2.27).

Expresando la señal en un instante n , y apilando en realizaciones tenemos

$$\mathbf{x}(n) \approx s(n)\mathbf{1} - s'(n)\boldsymbol{\theta} + \mathbf{v}(n), \quad (2.28)$$

siendo $\mathbf{x}(n)$ un vector columna y $\mathbf{v}(n)$ otro vector columna de ruido de tamaño M de las mismas características ya definidas en (2.1), $\boldsymbol{\theta}$ un vector columna de longitud M con los desalineamientos de todas las realizaciones, y $\mathbf{1}$ un vector columna de todo unos de longitud M , vemos el aspecto de estos vectores,

$$\mathbf{x}(n) = [\mathbf{x}_1(n) \ \mathbf{x}_2(n) \ \cdots \ \mathbf{x}_M(n)]^T, \quad (2.29)$$

$$\mathbf{v}(n) = [v_1(n) \ v_2(n) \ \cdots \ v_M(n)]^T, \quad (2.30)$$

$$\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \cdots \ \theta_M]^T, \quad (2.31)$$

$$\mathbf{1} = [1 \ 1 \ \cdots \ 1]^T, \quad (2.32)$$

donde se cogen las M realizaciones de un mismo instante temporal n .

La nueva matriz de autocorrelación también la podemos formar de la siguiente manera, siguiendo un procedimiento análogo a (2.5)

$$\mathbf{R}_x^\bullet = E[\mathbf{x}(n)\mathbf{x}(n)^T] \approx \frac{1}{N}\mathbf{X}^T\mathbf{X}. \quad (2.33)$$

Nosotros seguiremos trabajando con la estimación de \mathbf{R}_x^\bullet , es decir $\hat{\mathbf{R}}_x^\bullet = \frac{1}{N}\mathbf{X}^T\mathbf{X}$. Procedemos a la resolución de la expresión (2.33),

$$\begin{aligned} \mathbf{R}_x^\bullet = E[\mathbf{x}(n)\mathbf{x}(n)^T] &= E[(s(n)\mathbf{1} - s'(n)\boldsymbol{\theta} + \mathbf{v}(n))(s(n)\mathbf{1} - s'(n)\boldsymbol{\theta} + \mathbf{v}(n))^T] = \\ &= \frac{E_s}{N}\mathbf{1}\mathbf{1}^T + \frac{E_{s'}}{N}\boldsymbol{\theta}\boldsymbol{\theta}^T + \sigma_v^2\mathbf{I}, \end{aligned} \quad (2.34)$$

donde los autovectores de nuestra nueva matriz de autocorrelación son, $\mathbf{1}$, $\boldsymbol{\theta}$ y cualquier conjunto de vectores pertenecientes al espacio M-dimensional ortogonal a los dos anteriores y ortogonales entre ellos mismos.

Veamos los autovalores asociados a estos autovectores

$$\lambda_i = \begin{cases} \frac{E_s M}{N} + \sigma_v^2, & i = 1; \\ \frac{E_{s'}}{N} E_\theta + \sigma_v^2, & i = 2; \\ \sigma_v^2, & i = 3, \dots, M. \end{cases} \quad (2.35)$$

Recordamos que $\boldsymbol{\theta}^T\boldsymbol{\theta} = E_\theta$.

Vemos que el mayor de los autovalores λ_1 se corresponde con $\frac{E_s M}{N} + \sigma_v^2$ y va asociado al siguiente autovector normalizado ya $\frac{1}{\sqrt{M}}$. Lo podemos identificar como la suma de todos los elementos de la matriz \mathbf{R}_x^\bullet

$$\frac{\mathbf{1}^T}{\sqrt{M}}\mathbf{R}_x^\bullet\frac{\mathbf{1}}{\sqrt{M}} = \lambda_1\frac{\mathbf{1}^T}{\sqrt{M}}\frac{\mathbf{1}}{\sqrt{M}} = \lambda_1. \quad (2.36)$$

Podemos aproximar con $\hat{\mathbf{R}}_x^\bullet$ y haciendo uso de su simetría, el valor de λ_1 . Vemos a continuación siguiendo el desarrollo de (2.36)

$$\frac{\mathbf{1}^T}{\sqrt{M}} \hat{\mathbf{R}}_x^\bullet \frac{\mathbf{1}}{\sqrt{M}} = 2\Lambda_\Delta(\boldsymbol{\theta}) + \text{tr} \left\{ \hat{\mathbf{R}}_x^\bullet \right\} = \lambda_1. \quad (2.37)$$

Análogamente al igual que en el método Woody comentado al inicio de este capítulo, la revisión de máxima verosimilitud de este método de alineamiento (2.24) vuelve a ser equivalente a maximizar λ_1 . Debemos tener en cuenta que $\text{tr} \left\{ \hat{\mathbf{R}}_x^\bullet \right\}$ no se ve afectada por la variación de $\boldsymbol{\theta}$.

Un factor importante que incorpora *ML-Woody* respecto al método de Woody es la no necesidad de estimar un filtro adaptado \mathbf{h} . Esta diferencia es importante porque la estimación de este filtro será mas imprecisa a medida que tengamos menor número de realizaciones [6], o estas se encuentren demasiados desalineadas antes de realizar su promedio.

Certificamos ahora que nuestra búsqueda de una matriz de autocorrelación más eficiente ha concluido. Ya tenemos de forma directa la información que buscábamos, los desalineamientos de cada una de nuestras M realizaciones en el autovector $\boldsymbol{\theta}$.

Nos surge un nuevo problema, el autovector $\boldsymbol{\theta}$, por formar parte de una base generado a través de nuestra matriz de autocorrelación estará en su versión normalizada que denotaremos de la siguiente forma $\|\mathbf{v}_2\| = \mathbf{v}_2^T \mathbf{v}_2 = 1$. Por lo tanto tendremos que desnormalizarlo y además tendremos otro problema, deberemos determinar el signo de $\boldsymbol{\theta}$ que se simplificó al normalizarlo.

Busquemos este factor de escala de $\boldsymbol{\theta}$, sabemos que

$$\mathbf{R}_x^\bullet \boldsymbol{\theta} = \lambda_2 \boldsymbol{\theta}$$

y que $\boldsymbol{\theta}$ estará normalizado, por lo tanto

$$\beta \mathbf{v}_2 = \boldsymbol{\theta}$$

si hacemos

$$\boldsymbol{\theta}^T \boldsymbol{\theta} = \mathbf{v}_2^T \mathbf{v}_2 \beta^2 = \beta^2 = E_\theta \approx M \sigma_\theta^2 \quad (2.38)$$

recordamos que el vector de desalineamientos tiene media nula, por ello la varianza se corresponde con la potencia. Podemos poner este factor de escala en función del segundo autovalor.

$$\lambda_2 = \frac{E_{s'}}{N} E_\theta + \sigma_v^2 \quad (2.39)$$

Como sabemos que

$$\beta^2 = E_\theta \approx M \sigma_\theta^2 \quad (2.40)$$

y por otro lado despejamos E_θ de (2.39)

$$E_\theta = \frac{(\lambda_2 - \sigma_v^2)N}{E_{s'}} \quad (2.41)$$

ya tenemos por fin nuestro factor de escala β

$$\beta = \sqrt{E_\theta} = \sqrt{\frac{(\lambda_2 - \sigma_v^2)N}{E_{s'}}} \quad (2.42)$$

Siguiendo ahora con la función de coste definida en (2.21), ya que tenemos una dependencia muy similar en los autovalores que en la caso de la matriz de autocorrelación \mathbf{R}_x . La interpretación de esta función de coste no cambia en el caso de tener la matriz de \mathbf{R}_x^\bullet , se sigue comportando de la misma manera, una función que se maximiza a medida que nos acercamos a la situación de alineamiento perfecto.

Vemos una figura con unos resultados finales,

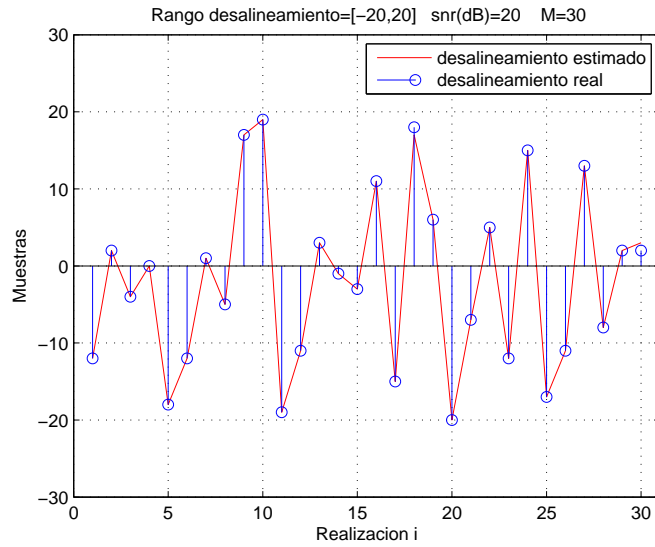


Figura 2.7: Vemos el desalineamiento en muestras real (azul) y el estimado a partir del segundo autovector ya desnormalizado de \mathbf{R}_x^\bullet (rojo) para cada una de las 30 realizaciones.

2.2. Alineamiento mediante método de correlación

El método de alineamiento de la correlación es la forma más obvia de estimar el retardo. Podemos ver en los anexos A y B como la correlación es un estimador apropiado en presencia de ruido gaussiano. No es un método único en su forma de implementar, en este caso lo hemos implementando de la siguiente forma.

Inicialmente, de forma experimental, obtenemos una estimación de nuestra señal a alinear $\hat{s}(n)$ mediante un promedio a lo largo de las realizaciones

$$\hat{s}(n) = E[\mathbf{x}(n)] = \frac{1}{M} \sum_{i=1}^M x_i(n). \quad (2.43)$$

Al tener el ruido de $x(n)$ media nula (definido en 2.1) en promedio tenderá a desaparecer, mientras que la forma de nuestra señal $s(n)$ dependiendo de los desalineamientos θ_i de las M realizaciones producirá una cierta degradación de nuestra señal de interés. Si el desalineamiento es considerable respecto de la evolución de $s(n)$ obtendremos un efecto de filtrado paso bajo en mi estimación de la señal de interés. Ilustramos este fenómeno en varias figuras.

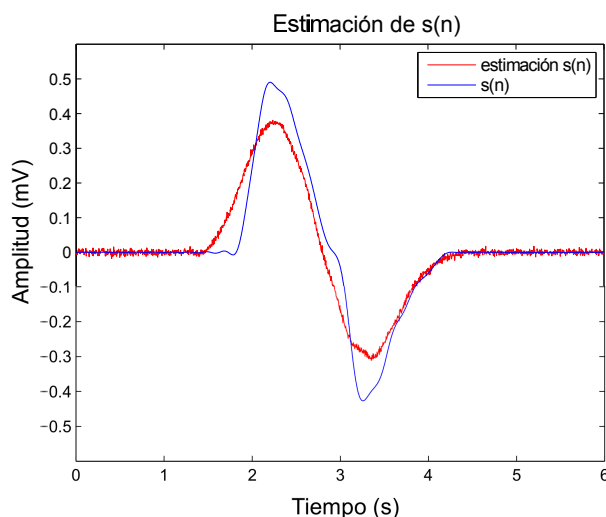


Figura 2.8: Rango de desalineamiento de $[-100,100]$ muestras, $M=30$

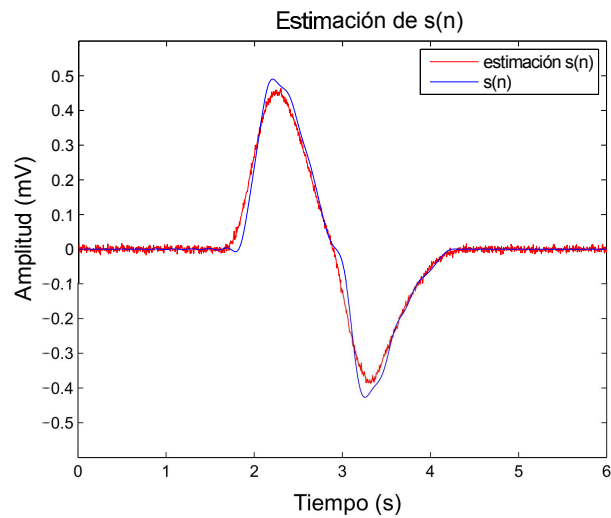


Figura 2.9: Rango de desalineamiento de $[-50,50]$ muestras, $M=30$

Además del rango de desalineamiento otro factor que influye notablemente en nuestra estimación de $s(n)$ es el número de realizaciones, M . Una M elevada hace que la media de los desalineamientos sea más próxima a cero. De tener un número bajo de realizaciones nos supondría estimar una forma de $s(n)$ con una cierta latencia de los máximos o deformación respecto de la forma perfectamente alineada, pese a que la distribución de los desalineamientos sea uniforme y de media nula. Ilustramos este fenómeno con varias figuras.

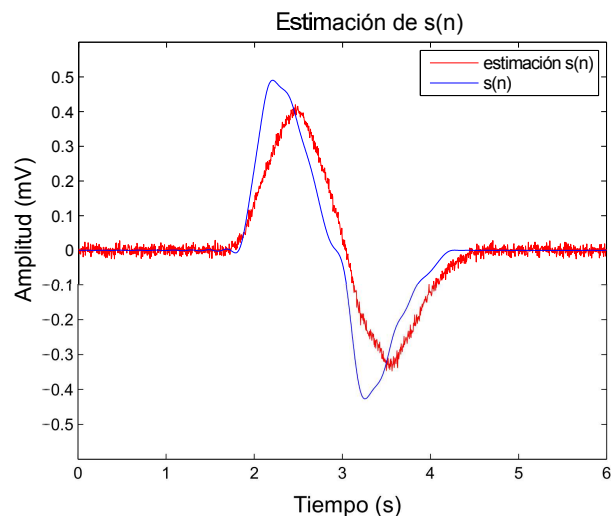


Figura 2.10: Rango de desalineamiento de $[-100,100]$ muestras, $M=10$

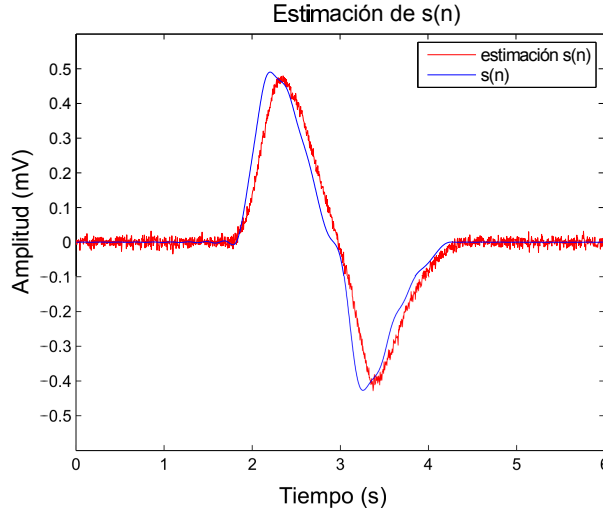


Figura 2.11: Rango de desalineamiento de $[-50,50]$ muestras, $M=10$

Apreciamos este fenómeno de latencia que cuantificaremos mediante un sesgo, ϵ , al observar una cierta latencia respecto de los máximos de la señal ideal y la estimada. No será un problema clínico debido a que los θ_i no son inicialmente muy grandes, dado que la señal se encontrará correctamente segmentada. Unido a este ϵ vemos una mayor distorsión en la forma de la señal de interés, esto se debe a que al tener un M menor la estimación de $s(n)$ es menos robusta. Este fenómeno será un riesgo que aparecerá cuando M sea bajo y un gran desalineamiento este presente. Se tendrá en cuenta en simulación.

Una vez conocida como estimamos la señal con la que correlaremos el resto y los posibles efectos de esta estimación en simulación, procederemos a explicar las funciones a usar. En este caso solo necesitaremos usar la correlación entre una señal de referencia, que la estimaremos como ya dijimos anteriormente, y nuestro conjunto de realizaciones.

Realizaremos M correlaciones cruzadas, una para cada realización

$$r_{x_i}(k) = \frac{1}{2N-1} \sum_n \hat{s}(n-k)x_i(n), \quad i = 1, \dots, M. \quad (2.44)$$

En este caso hemos realizado una operación equivalente, la convolución, habiendo reflejado antes una de las dos señales en el dominio temporal.

Acto seguido procedemos a buscar el máximo en cada $r_{x_i}(k)$, para que valor de k es la función (2.44) máxima. Denotaremos a este índice como k_{max} . Esta distancia en muestras nos dará una estimación del desalineamiento para cada realización. Debemos tener en cuenta que la correlación nos introduce un cierto retardo, ya que paso a tener $2N-1$ muestras, con restar N será suficiente para obtener una nueva referencia.

Cabe recordar que es posible que tengamos un cierto error en el alineamiento debido al sesgo, denotado numéricamente como ϵ . Esta latencia o sesgo la podremos ver como un offset que será el mismo en todas las realizaciones. Por ello aplicamos la siguiente corrección a las M realizaciones,

$$\theta_i = k_{i_{max}} - N - \epsilon, \quad i = 1, \dots, M. \quad (2.45)$$

Ilustramos una figura detallando la latencia o sesgo, ϵ

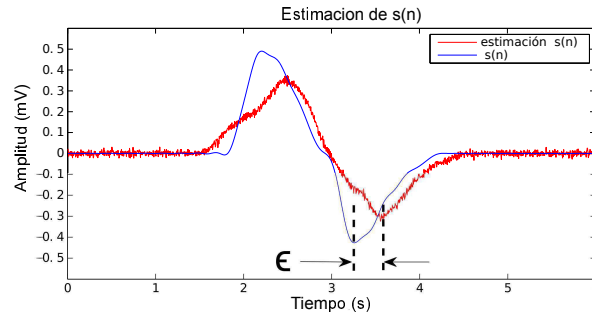


Figura 2.12: Observamos el efecto de la latencia respecto los máximos

Para finalizar mostramos un diagrama de bloques,

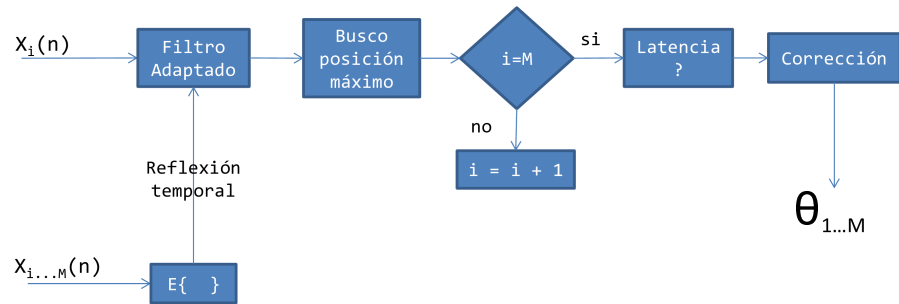


Figura 2.13: Diagrama de bloques método de correlación

Vemos de forma secuencial la implementación del método. El bloque corrección se incluye para corregir la latencia calculada previamente, añadido a la corrección del desplazamiento de N muestras producido al realizar la correlación cruzada. En este caso se ha optado para implementar la correlación cruzada una convolución con la estimación de $s(n)$ (2.43) reflejada temporalmente siendo una operación equivalente.

2.3. Alineamiento mediante método de Woody

El origen de este método proviene de C.D. Woody [4], tiene un gran parecido con el método planteado en el apartado 2.2. Podemos decir de forma resumida que se trata de una versión iterativa del método descrito en 2.2.

En este caso hemos desarrollado el método a partir de otra fuente [8]. Comentamos a continuación los aspectos más destacables para llevar a cabo su implementación. Dentro del apartado 2.1 ya hicimos alguna referencia hacia este método de alineamiento [6].

Se demuestra que el estimador de retardo óptimo de máxima verosimilitud en presencia de ruido gaussiano es, (Anexo B)

$$\hat{\theta}_i = \arg \max_{\theta_i} \left(\sum_{n=\theta_i}^{\theta_i+D-1} x_i(n)s(n-\theta_i) \right). \quad (2.46)$$

Donde $\hat{\theta}_i$ viene dado por la mayor correlación cruzada entre $s(n)$ y $x_i(n)$ conseguida de entre todas las posibles. A este valor obtenido de $\hat{\theta}_i$ le aplicaremos el factor corrector visto en el capítulo 2.2, concretamente la ecuación (2.45).

Podemos ver también esta maximización como una operación de filtrado,

$$y_i(\theta_i) = \sum_{n=\theta_i}^{\theta_i+D-1} x_i(n)h(\theta_i-n), \quad (2.47)$$

donde la respuesta al impulso del filtro $h(k)$ es igual a la reflexión en tiempo de $s(n)$,

$$h(k) = \begin{cases} s(D-1-k), & k=0, \dots, D-1; \\ 0, & \text{resto.} \end{cases} \quad (2.48)$$

Este filtro es también conocido como filtro adaptado y constituye un bloque importante en estimadores de alineamiento como ya hemos visto en capítulos anteriores.

En la práctica la forma de onda de $s(n)$ que define al filtro adaptado se obtiene mediante una estimación que proviene de la matriz \mathbf{X} . Si asumimos que inicialmente tenemos una variación de retardo no muy grandes la aproximación inicial vendrá de promediar la M realizaciones desalineadas, este caso será el nuestro, no tendremos gran rango de desalineamientos.

$$\hat{s}_a^{(0)}(n) = \frac{1}{M} \sum_{i=1}^M x_i(n). \quad (2.49)$$

Una vez obtenemos la primera aproximación de los M desalineamientos θ_i , $\hat{\theta}_i^{(1)}$, corregimos los desalineamientos, y obtenemos una nueva estimación de la señal $s(n)$.

$$\hat{s}_a^{(1)}(n) = \frac{1}{M} \sum_{i=1}^M x_i(n + \hat{\theta}_i^{(1)}). \quad (2.50)$$

Continuaremos obteniendo los M desalineamientos con una estimación cada vez más robusta de $s(n)$. A continuación plasmamos el valor de la estimación j -ésima de $s(n)$, $\hat{s}_a^{(j)}(n)$, obtenida a partir de los desalineamientos de la iteración j del método, $\hat{\theta}_i^{(j)}$,

$$\hat{s}_a^{(j)}(n) = \frac{1}{M} \sum_{i=1}^M x_i(n + \hat{\theta}_i^{(j)}). \quad (2.51)$$

Se decide que nuestro algoritmo habrá alcanzado la convergencia cuando la evolución de los desalineamientos obtenidos en sucesivas iteraciones diste en un determinado valor. Estableceremos un umbral, η_θ , que marcará la condición de convergencia y la obtención de los valores finales de los M desalineamientos θ_i . Esta condición será,

$$\max_i \left| \hat{\theta}_i^{(j)} - \hat{\theta}_i^{(j-1)} \right| \leq \eta_\theta, \quad j \geq 1, \quad (2.52)$$

donde j indica el número de iteraciones y η_θ será un número entero positivo y será un parámetro más de simulación en el método de Woody.

Presentamos a continuación un diagrama de bloques del método de Woody,

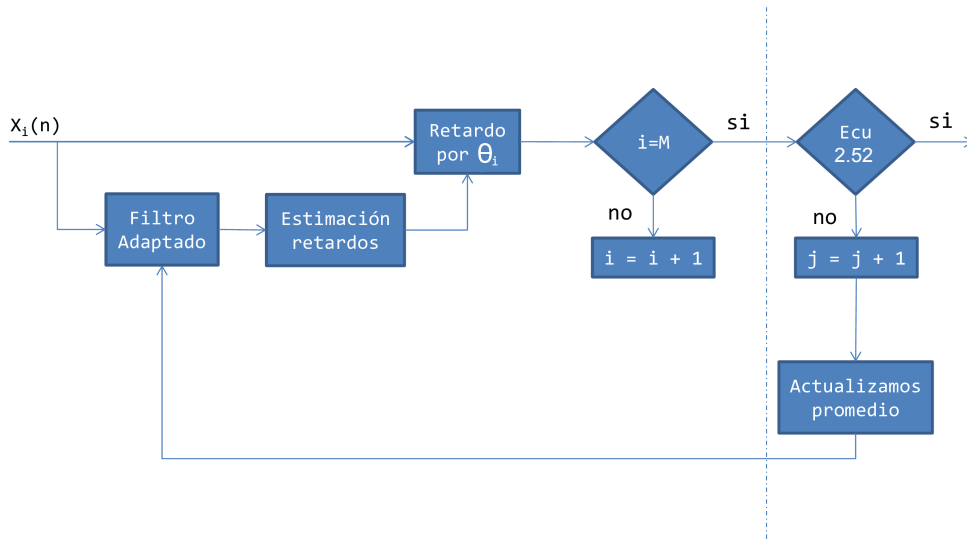


Figura 2.14: Diagrama de bloques del método de Woody. El procedimiento iterativo finaliza al cumplir (2.52).

De forma implícita se incluye en este diagrama de bloques 2.14 la figura 2.13 realizada en el método de correlación.

Ilustramos con una figura el comportamiento de este método en simulación,

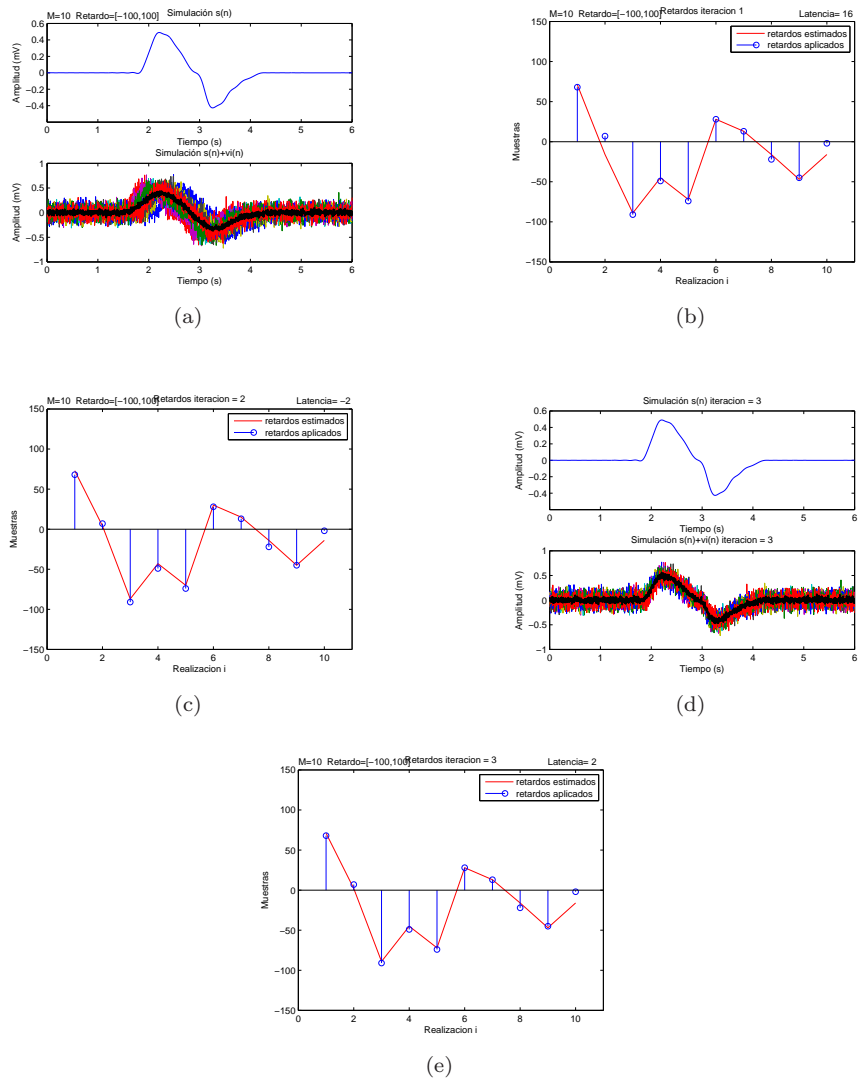


Figura 2.15: El error respecto la estimación del desalineamiento y el retardo real evoluciona de la siguiente forma en media: 1ª iteración 8'93, 2ª iteración 5'64 y 3ª iteración 4'42 muestras en cada caso. Vemos destacado en negro en las figuras 2.15a y 2.15d el promedio de las M realizaciones.

Capítulo 3

Resultados

Nos proponemos testear este nuevo método de alineamiento descrito en el apartado 2.1. Para ello nos centraremos en varios aspectos (parámetros) del entorno de simulación que podremos controlar:

- M , el número de realizaciones de nuestro ensemble o conjunto de realizaciones. Veremos la evolución en el error de alineamiento entre los desplazamientos estimados y el desalineamiento real provocado al inicio de la simulación.
- SNR, será otro parámetro con el cual controlaremos la potencia de ruido disponible en cada simulación. Definimos a este parámetro como $10 \lg(\frac{P_s}{P_v})$, un cociente de potencias. Donde P_s se corresponde con la potencia de señal, $s(n)$ y P_v se corresponde con la potencia de ruido, $v(n)$.
- θ , nuestro vector de desalineamiento generado mediante una distribución uniforme de media nula. Determinaremos su rango de valores de forma determinista, permitiendo en un caso desalineamientos de unas pocas muestras o de hasta ± 100 muestras. Teniendo en cuenta la frecuencia de muestreo, 250 Hz, estamos tratando con tiempos desde los ± 4 ms hasta los ± 400 ms.
- $s(n)$, la forma de onda de la señal a alinear. En este caso tiene unas características concretas y forma fijas. Posteriormente con la superposición del ruido y el retardo se distorsionará. Consideraremos alguna posible modificación en su amplitud en el anexo C.

Además valoraremos la veracidad de los resultados obtenidos mediante la repetición de la simulación bajo los mismos parámetros durante 100 veces y luego promediaremos a lo largo de las 100 repeticiones. Obtendremos además los boxplot (diagramas de caja con sus percentiles al 25 % y 75 % y mediana) correspondientes de cada una de nuestras simulaciones, observaremos así la varianza, parámetro fundamental de un estimador.

3.1. Descripción

A continuación comentaremos los resultados obtenidos de las simulaciones realizadas en el entorno Matlab. Haremos mayor hincapié en ver los resultados en error expresado en muestras entre la estimación $\hat{\theta}_i$ y el retardo real provocado θ_i . Se expresará el error mediante la raíz del valor cuadrático medio o *MSE* (*Mean Square Error*), para obtener un resultado final en término de muestras. Vemos la expresión,

$$\text{Error}(muestras) = \text{MSE}(muestras) = \mathcal{E} = \sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{\theta}_i - \theta_i)^2}. \quad (3.1)$$

También mostraremos la distribución estadística en forma de boxplot para cada una de las simulaciones realizadas (cada simulación consta de 100 repeticiones bajo la misma clase de parámetros).

Trataremos además de justificar los resultados obtenidos, tanto en caso de obtener unas prestaciones excelentes en lo que a error se refiere así como si los resultados obtenidos no son los esperados. En este último caso se buscará cual es la fuente que origina que los resultados obtenidos no sean los esperados.

Y por último buscaremos que conjunto de parámetros de todos los aplicados nos conduce a la situación con mejor resultado final, es decir menor error posible y mínima varianza posible.

A continuación los vectores para cada parámetro,

$$\mathbf{M} = \{5 \ 10 \ 15 \ 20 \ 25 \ 30 \ 50 \ 75 \ 100\} \text{ (realizaciones)}, \quad (3.2)$$

$$\text{SNR} = \{-10 \ -5 \ 0 \ 5 \ 10 \ 15 \ 20 \ 25\} \text{ (dB)}, \quad (3.3)$$

$$\boldsymbol{\theta} = \{\pm 25 \ \pm 50 \ \pm 75 \ \pm 100\} \text{ (rango en muestras)}. \quad (3.4)$$

3.1.1. Método de alineamiento mediante autovalores

Para comenzar mencionaremos el tiempo medio llevado a cabo para cada simulación. Una simulación comprende el barrido de parámetros en su totalidad para M y SNR, para θ escogeremos uno de los 4 rangos posibles de desalineamiento (3.4) solamente en una simulación. Además se llevará a cabo 100 repeticiones para cada una de las combinaciones posibles de parámetros mencionada anteriormente. La duración de una simulación comprende un tiempo de 3-4 minutos para este método de alineamiento.

La fase de simulación fue realizada con el siguiente portátil, *Dell studio 1555*, procesador *Intel Core 2 Duo T 6500 2.1 Ghz*.

Viendo la gran cantidad de datos posibles obtenida de cada simulación comentaremos los puntos más destacables:

1. A medida que el rango de θ aumenta, el error aumenta de forma aproximadamente proporcional en la mayoría de los casos manteniendo el resto de parámetros fijos. Esto posiblemente tenga lugar debido a que la aproximación realizada en (2.13) cada vez es más errática, dando lugar una forma de $s(n)$ cada vez más distorsionada conforme θ aumenta.

Ilustramos este fenómeno con varias figuras,

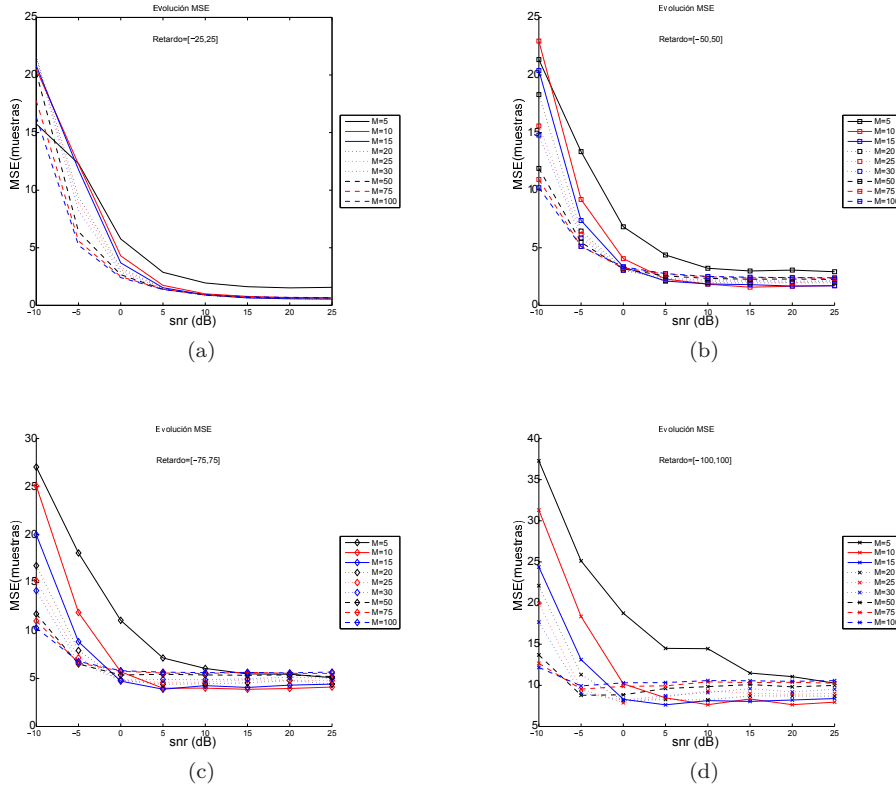
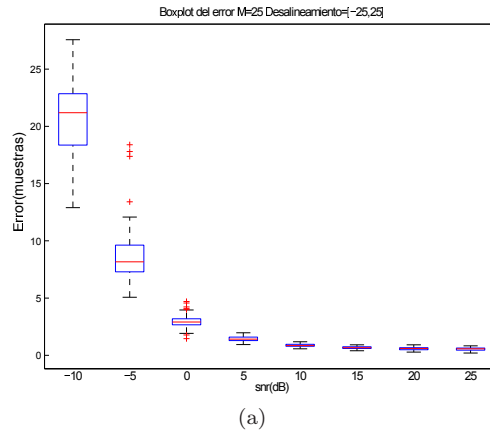


Figura 3.1: Observamos el error en 4 figuras para los 4 rangos de desalineamiento en muestras disponibles: $\theta = [-25, 25]$ 3.1a, $\theta = [-50, 50]$ 3.1b, $\theta = [-75, 75]$ 3.1c y $\theta = [-100, 100]$ 3.1d. En cada gráfica tenemos para cada M diferente un barrido en SNR.

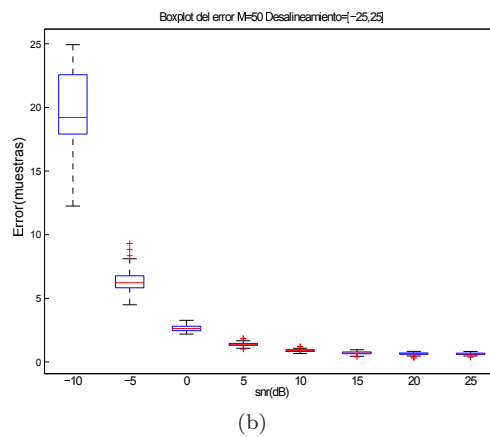
Podemos ver que independientemente del número de realizaciones M con SNR creciente el error decrece hasta alcanzar un valor fijo para los rangos de θ más bajos. A medida que θ crece en rango, el error aumenta y su variabilidad entre diferente M también lo hace.

Otro resultado que vemos al aumentar θ es que la variabilidad en el error aumenta, de forma más notable a medida que nos distanciamos. Como conclusión de este suceso podemos decir que la fiabilidad del estimador de retardos baja conforme θ aumenta.

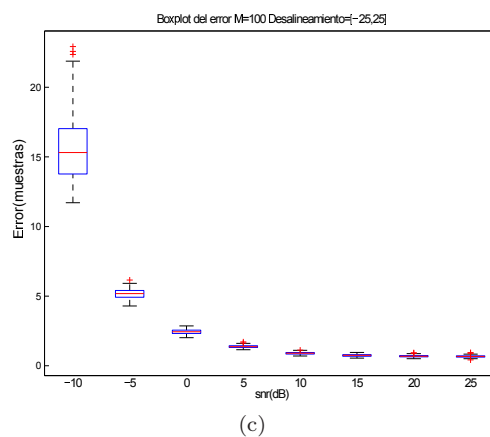
Lo vemos de forma gráfica diferentes en boxplot del error,



(a)



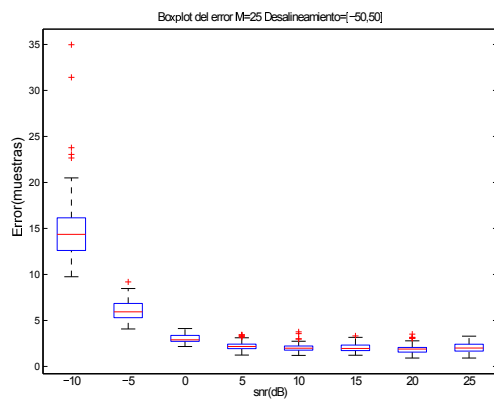
(b)



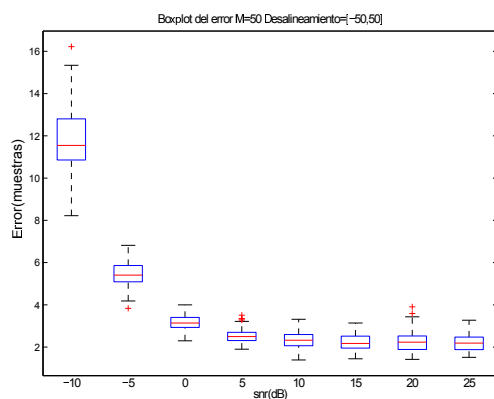
(c)

Figura 3.2: Observamos los boxplot para un rango de desalineamiento θ de ± 25 muestras: $M = 25$ 3.2a, $M = 50$ 3.2b, $M = 100$ 3.2c. En cada gráfica tenemos para un M concreto un barrido en SNR.

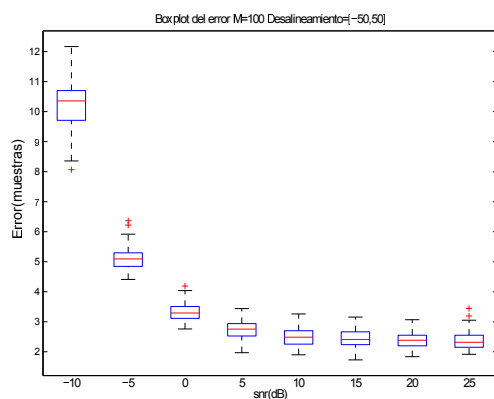
Vemos el mismo comportamiento con un rango desalineamiento mayor, ilustramos el fenómeno para cotas de θ igual a ± 50 y ± 100 muestras,



(a)



(b)



(c)

Figura 3.3: Observamos los boxplot para un rango de desalineamiento θ de ± 50 muestras: $M = 25$ 3.3a, $M = 50$ 3.3b, $M = 100$ 3.3c. En cada gráfica tenemos para un M concreto un barrido en SNR.

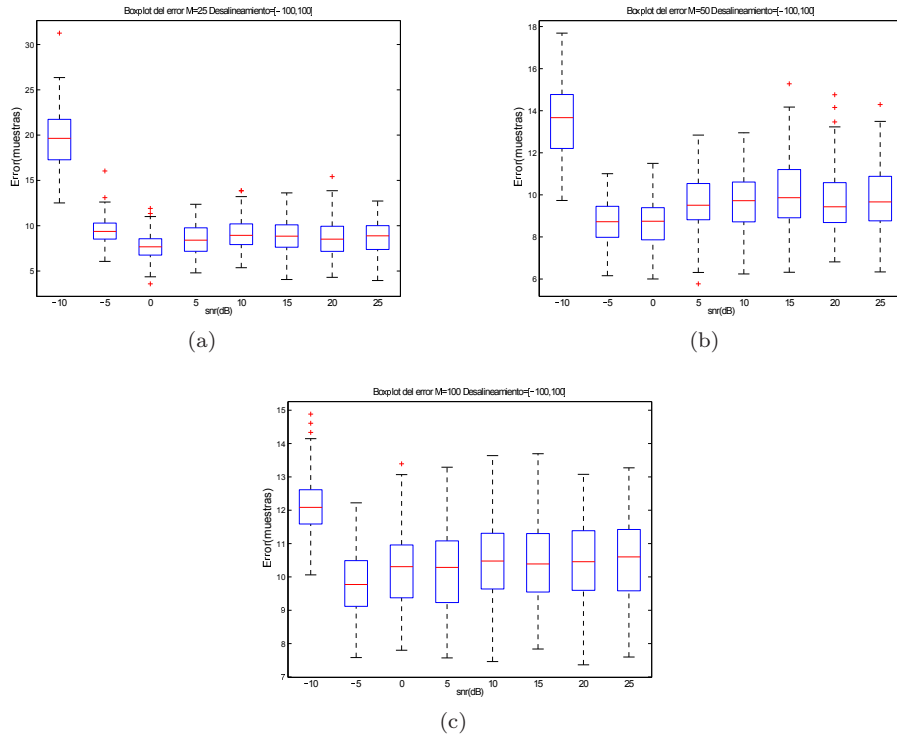


Figura 3.4: Observamos los boxplot para un rango de desalineamiento θ de ± 100 muestras: $M = 25$ 3.4a, $M = 50$ 3.4b, $M = 100$ 3.4c. En cada gráfica tenemos para un M concreto un barrido en SNR.

Podemos observar como las distribuciones del error tienen percentiles que abarcan mayor cantidad de muestras conforme θ aumenta su valor. Se puede ver también como un mayor M compacta más la distribución del error reduciendo su variabilidad, haciendo que el estimador converja en el entorno de un valor de rango más reducido.

En las figuras anteriores hemos visto como θ es un parámetro muy importante en los resultados obtenidos por nuestro estimador de desalineamiento. Esto se debe a que por la aproximación asumida en (2.13) los vectores que conforman la matriz de autocorrelación obtenida en (2.34), distan de lo que nos dan los resultados teóricos. Una forma de darnos cuenta de que esto sucede con total seguridad sin mirar el desalineamiento es tratar de obtener el autovector que nos da el mayor de los autovalores de la matriz de autocorrelación obtenida en el capítulo 2.1.2.

Solamente necesitaremos conocer el número de realizaciones M , ya que sabemos que el autovector es un conjunto de unos cuya norma es \sqrt{M} . Observamos su distorsión,

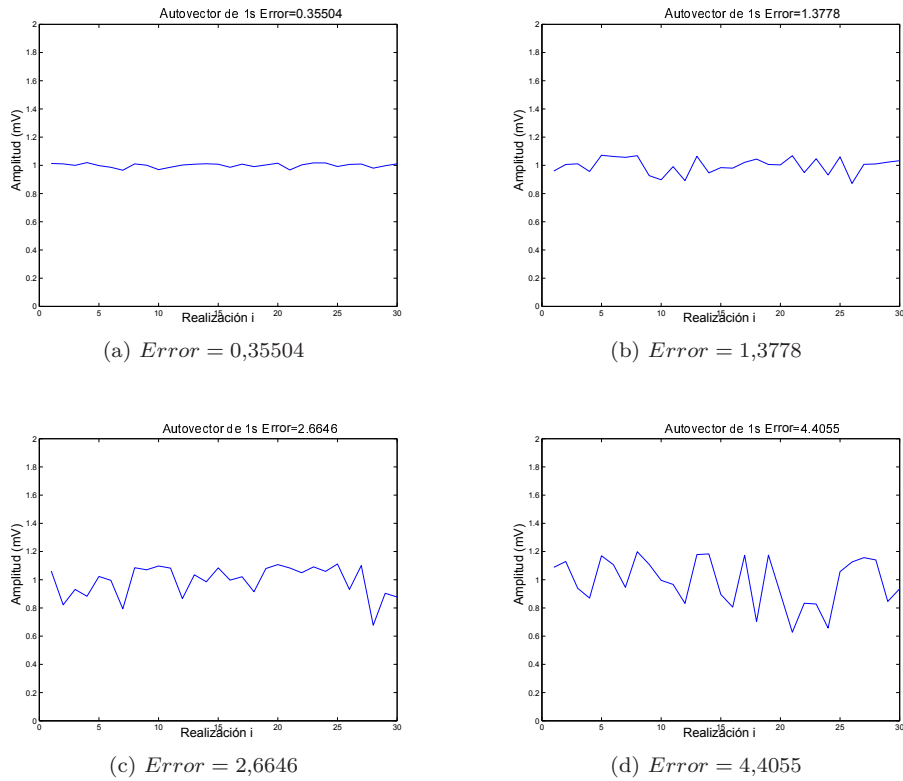


Figura 3.5: Observamos la distorsión del autovector **1** desnormalizado para diferentes θ con M y SNR fijas: $\theta = \pm 25$ 3.5a, $\theta = \pm 50$ 3.5b, $\theta = \pm 75$ 3.5c y $\theta = \pm 100$ 3.5d. Vemos además el error respecto el autovector ideal de forma numérica. Este aumenta de forma proporcional con θ al igual de lo que sucedía con el error de alineamiento.

Analizar este autovector nos puede dar una idea aproximada de la robustez del estimador, no por ello quiera decir que podamos extrapolar el error de este autovector con el valor del error en el alineamiento. Solamente se ha observado que a medida que θ aumenta en rango, los autovectores sufren mayor distorsión. Lo reflejamos en una tabla,

Cuadro 3.1: Vemos el error en muestras, donde fijamos $M = 30$ y SNR = 10 dB.

θ (rango en muestras)	1	θ
$[-25, 25]$	0,40	0,89
$[-50, 50]$	1,26	2,08
$[-75, 75]$	3,14	3,75
$[-100, 100]$	4,86	12,26

Podemos ver que el error aumenta en mayor medida en el caso del vector

de alineamiento, pero en ambos casos sigue una tendencia creciente. Nos aporta una ligera información de como puede ir el error de alineamiento.

- Otro parámetro que nos aporta información, es el número de realizaciones a alinear, M . A medida que nos encontramos ante una población mayor los resultados obtienen soluciones más robustas. Esto se puede deber a que las distribuciones de algunos de los parámetros son más fieles al modelo teórico descrito durante el capítulo 2.1.

Ilustramos en unas figuras,

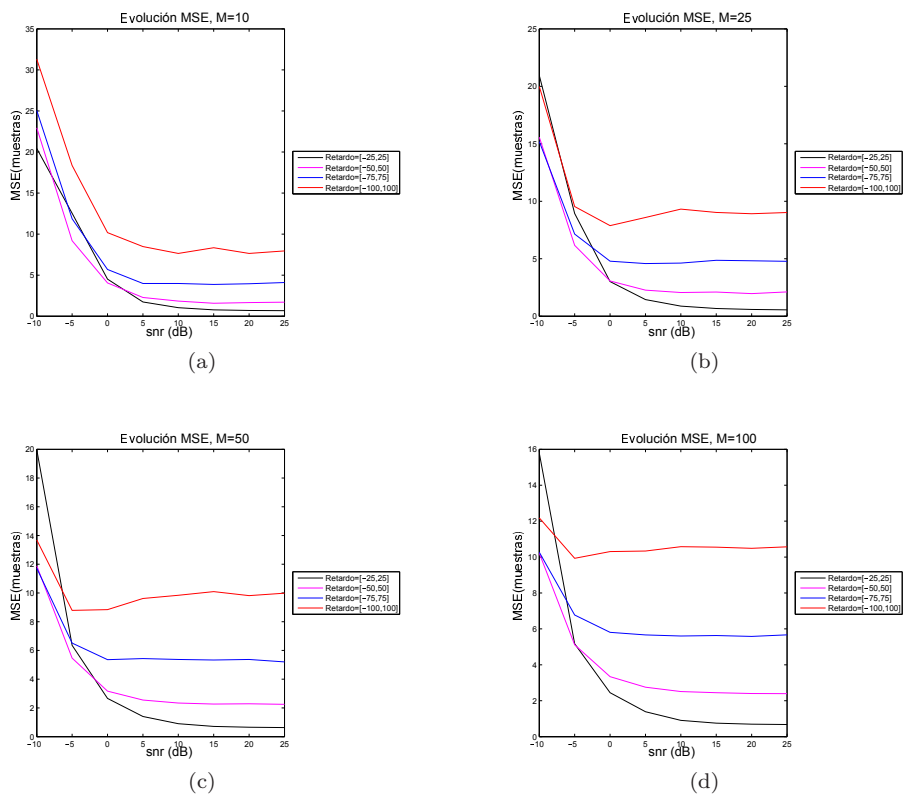


Figura 3.6: Observamos para los 4 rangos de θ , ± 25 , ± 50 , ± 75 y ± 100 muestras el error de alineamiento: $M = 10$ 3.6a, $M = 25$ 3.6b, $M = 50$ 3.6c y $M = 100$ 3.6d. En cada gráfica tenemos para un M concreto un barrido en SNR.

Se puede ver que las diferencias más aparentes ocurren en las SNR más bajas entre las diferentes figuras, para SNR más alta los valores de las 4 figuras tienden hacia valores más próximos habiendo menos diferencia en el error para los diferentes M . Además podemos ver un comportamiento similar en las 4 figuras, todos los trazados siguen una tendencia descendente hasta converger a un valor constante a medida que la relación señal a ruido crece, entre los diferentes trazados se puede ver una distancia significativa. Esto vuelve a deberse a que a medida que θ crece la estimación

de los autovalores es menos próxima a la situación descrita en la teoría y por tanto más errónea, esta situación la describimos en párrafos anteriores.

Otro fenómeno que se puede ver afectado por M es la estimación de la norma del autovector θ (2.42), ya que para obtener las diferentes variables que lo componen nos servimos de promedios de las realizaciones. Un menor M o un amplio valor de θ nos hará cometer mayor error.

Observando los boxplot, se puede apreciar que el error converge en mayor medida hacia un valor o hacia el entorno de un valor conforme M aumenta. Un M bajo hace que la solución del estimador tenga los percentiles del 25% y 75% a mayor distancia de su mediana.

Observamos unas figuras,

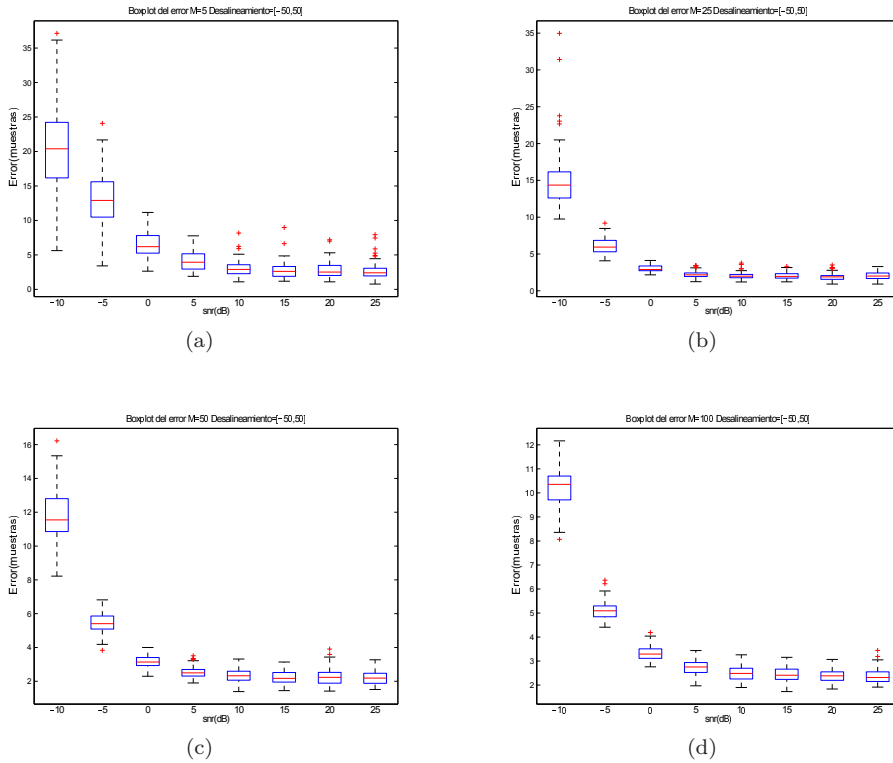


Figura 3.7: Observamos los boxplot para un rango de desalineamiento θ de ± 50 muestras: $M = 5$ 3.7a, $M = 25$ 3.7b, $M = 50$ 3.7c y $M = 100$ 3.7d. En cada gráfica tenemos para un M concreto un barrido en SNR.

El comportamiento visto en la figura (3.7) lo podemos extender para el resto de rangos de retardos.

3. En el caso de la SNR vemos una tendencia bastante clara, a medida que

esta aumenta su valor, las prestaciones del estimador de alineamiento aumentan, es decir, el error disminuye tanto en valor como en su distribución estadística como podemos ver en los boxplot de la figura (3.7) convergiendo hacia un valor numérico concreto en un entorno más cercano en muestras.

En cuanto a la estimación de algunas variables como el factor de escala, no es un factor tan determinante (a partir de los 0 dB en adelante el error no sufre cambios drásticos en la mayoría de las simulaciones). Sin embargo si nos encontramos ante un número de realizaciones por encima de las 20-25 o unos desalineamientos entre estas de no demasiadas muestras (con mayor M podremos asumir mayor desalineamiento o tener menor desalineamiento nos permite tener un M más bajo), aparecerá un sesgo.

Resumiendo la idea anterior, en caso de tener unas condiciones de pocas realizaciones, gran desalineamiento presente o una SNR negativa se observa un error en media en la estimación de los alineamientos. A esto suceso lo denotamos como sesgo o latencia, ϵ , lo describimos al final del capítulo 2.2. Lo vemos en unas figuras.

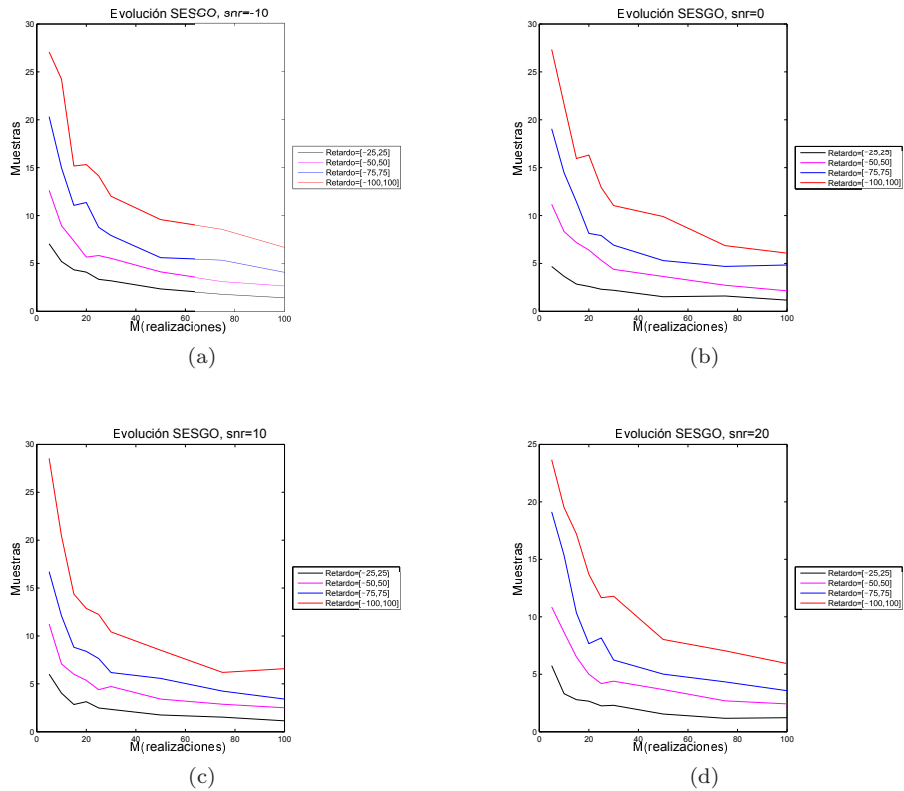


Figura 3.8: Observamos el sesgo del error para los 4 rangos de θ : SNR = -10dB 3.8a, SNR = 0 dB 3.8b, SNR = 10 dB 3.8c y SNR = 20 dB 3.8d. En cada gráfica tenemos para un SNR concreto un barrido en M .

3.1.2. Método de correlación y método de Woody

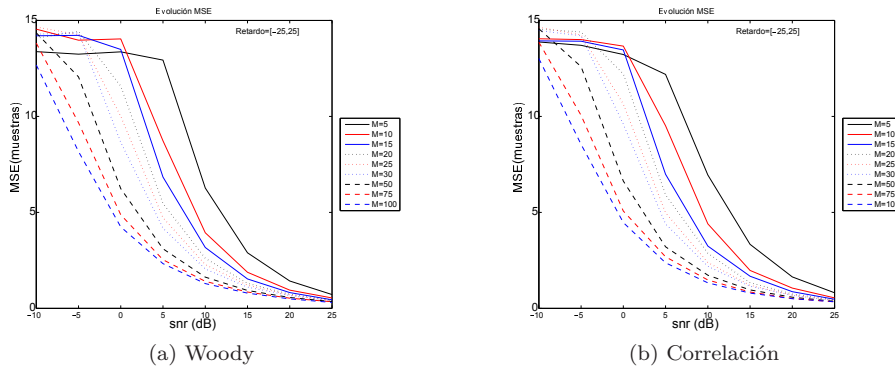
En ambos métodos observamos una tendencia similar en los resultados obtenidos. Siendo el método de Woody el que mejores resultados obtiene. Esto sucede debido a que podríamos decir que el método de correlación es una etapa que esta incluida en el método de Woody.

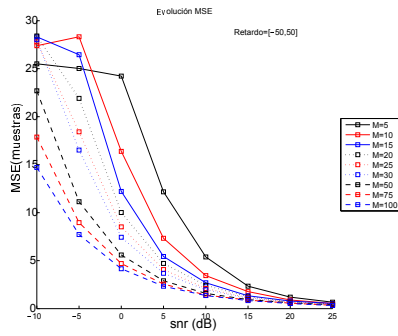
Sin embargo debemos mencionar que el tiempo de cómputo en una simulación entre los dos métodos difiere de forma notable. Esto es algo que se ha comprobado en numerosas ocasiones durante el desarrollo de las simulaciones de este TFG. Debido a la variación de parámetros (diferente número de realizaciones, diferentes SNR y rango de desalineamientos cada vez más amplios) y de las diferentes variables disponibles (M , SNR y θ). Además debemos mencionar las 100 repeticiones de cada simulación.

Con todos estos parámetros una simulación para el método de Woody comprendía un tiempo de 20-25 minutos mientras que en el caso de método de correlación este tiempo bajaba a la mitad aproximadamente, 10-13 minutos.

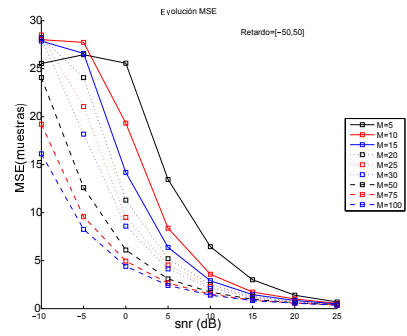
La fase de simulación fue realizada en las mismas condiciones y portátil que en el apartado 3.1.1.

Una vez explicada de forma introductoria algunas de las características de la simulación pasamos a comentar los resultados más destacados. Comentaremos los resultados obtenidos del método de Woody ya que estos se encuentran (ligeramente en algunos casos) en mejor lugar que en el caso del método de correlación. Ilustramos en una figura la similitud de ambos.

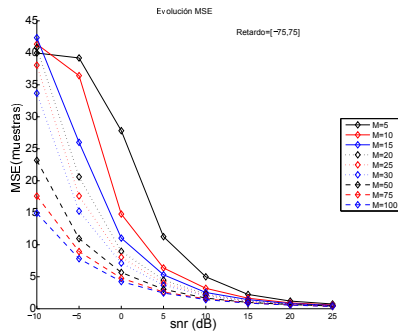




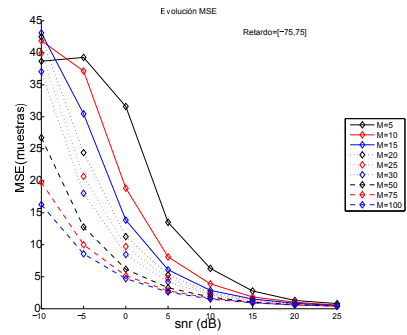
(c) Woody



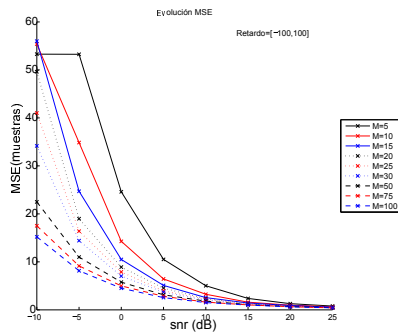
(d) Correlación



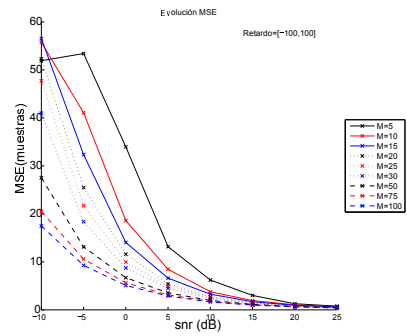
(e) Woody



(f) Correlación



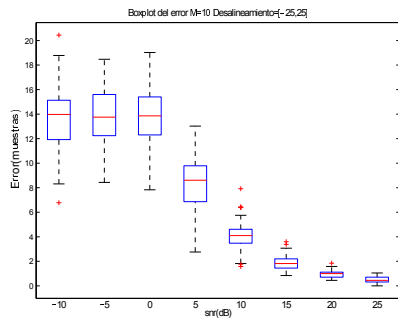
(g) Woody



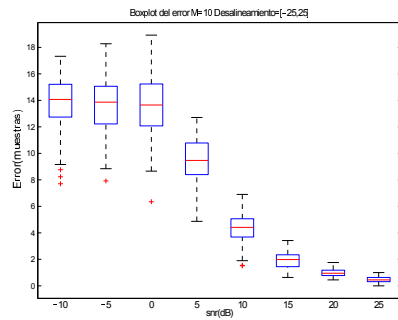
(h) Correlación

Figura 3.9: Observamos el error para los 4 rangos de $\theta \pm 25, \pm 50, \pm 75$ y ± 100 : $\theta = [-25, 25]$ 3.9a, $\theta = [-25, 25]$ 3.9b, $\theta = [-50, 50]$ 3.9c, $\theta = [-50, 50]$ 3.9d, $\theta = [-75, 75]$ 3.9e, $\theta = [-75, 75]$ 3.9f $\theta = [-100, 100]$ 3.9g y $\theta = [-100, 100]$ 3.9h. En cada gráfica tenemos para un M concreto un barrido en SNR.

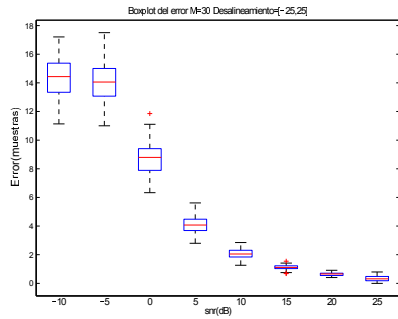
Sucede de la misma forma con los boxplot, lo vemos para un rango de θ de ± 25 muestras,



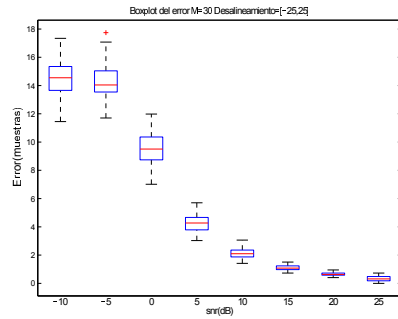
(a) Woody



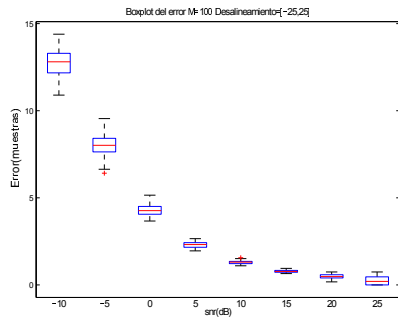
(b) Correlación



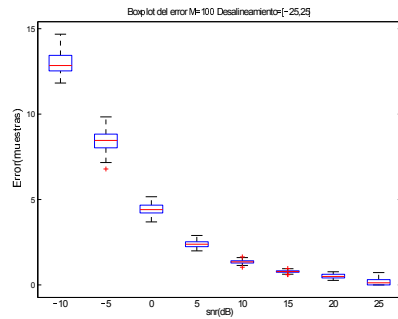
(c) Woody



(d) Correlación



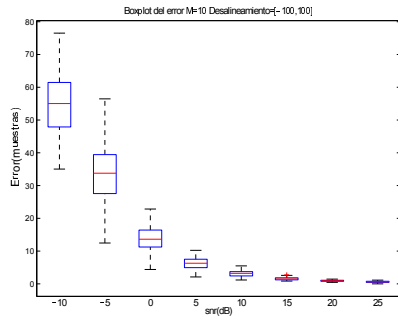
(e) Woody



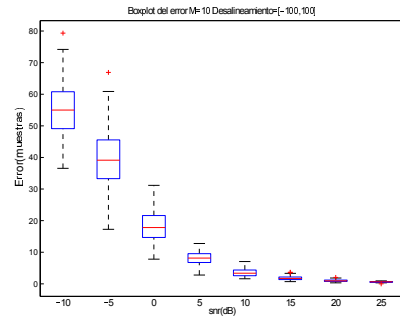
(f) Correlación

Figura 3.10: Observamos los boxplot para un rango de desalineamiento θ de ± 25 muestras: $M = 10$ 3.10a, $M = 10$ 3.10b, $M = 30$ 3.10c , $M = 30$ 3.10d, $M = 100$ 3.10e , $M = 100$ 3.10f. En cada gráfica tenemos para un M concreto un barrido en SNR.

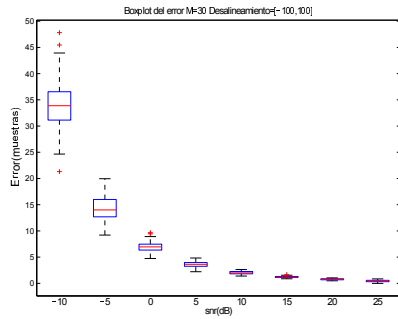
Aumentamos el rango de θ hasta las ± 100 muestras,



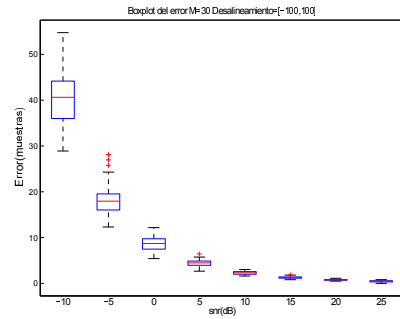
(a) Woody



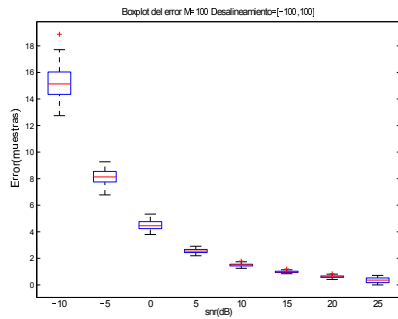
(b) Correlación



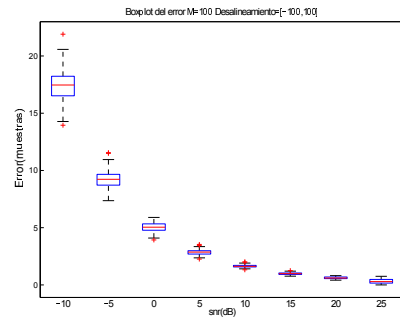
(c) Woody



(d) Correlación



(e) Woody



(f) Correlación

Figura 3.11: Observamos los boxplot para un rango de desalineamiento θ de ± 100 muestras: $M = 10$ 3.11a, $M = 10$ 3.11b, $M = 30$ 3.11c, $M = 30$ 3.11d, $M = 100$ 3.11e, $M = 100$ 3.11f. En cada gráfica tenemos para un M concreto un barrido en SNR.

Nos centramos de aquí en adelante en los resultados obtenidos del método de Woody.

1. Comenzamos describiendo la influencia del error a medida que θ aumenta. A diferencia de lo visto en el método de autovalores, donde para cada rango de θ se podía ver (figura 3.1) como el error aumentaba de forma notable para la mayoría de las combinaciones de parámetros de M y SNR conforme

el desalineamiento era mayor. En el caso del método de Woody esto no sucede así, podemos ver como para las SNR más altas (de alrededor 10 dB en adelante) y cualquier M el error no aumenta de forma tan notable y tiende hacia un valor inferior a las 2-3 muestras para todos rangos de θ . Lo vemos en unas figuras,

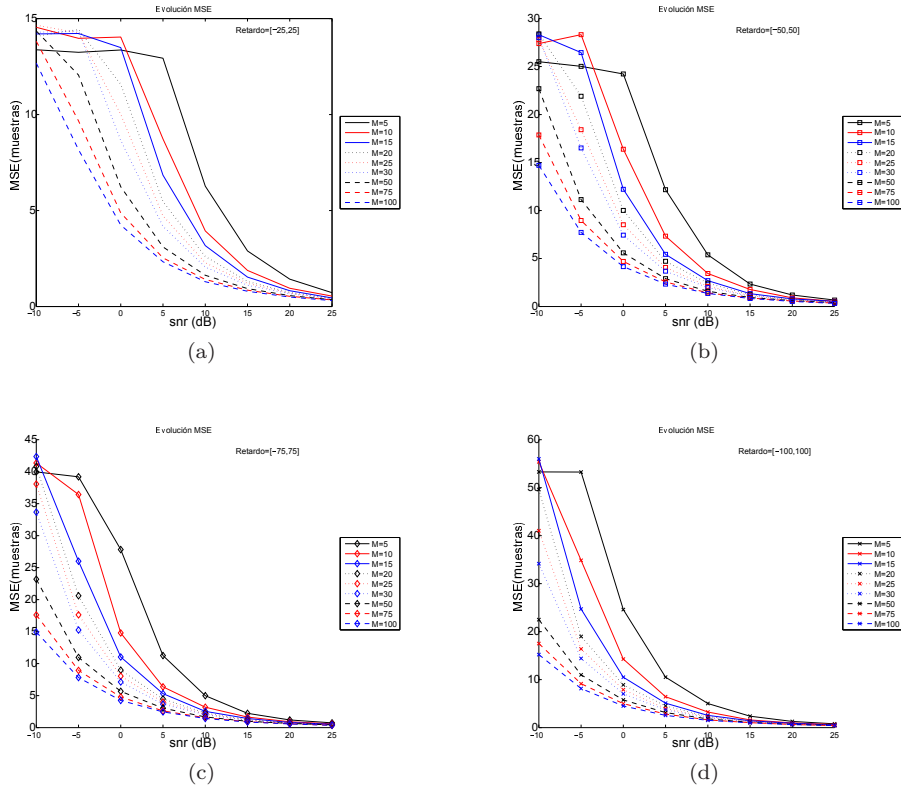
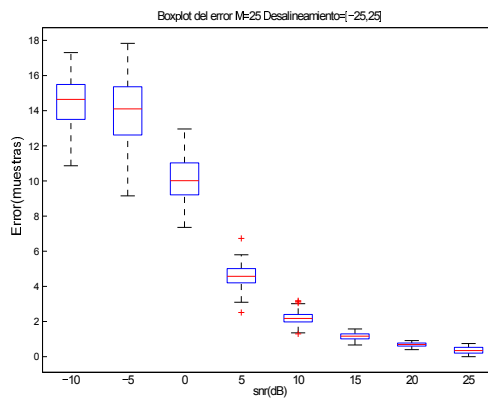


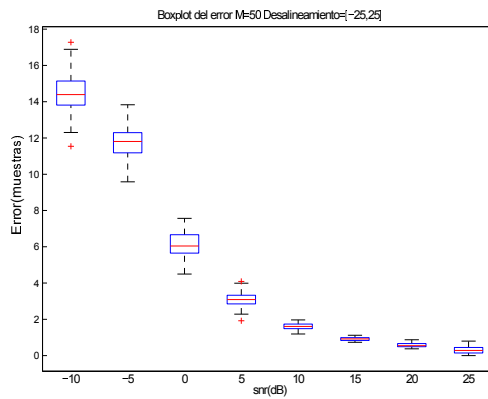
Figura 3.12: Observamos el error para 4 rangos de θ : $\theta = [-25, 25]$ 3.12a, $\theta = [-50, 50]$ 3.12b, $\theta = [-75, 75]$ 3.12c y $\theta = [-100, 100]$ 3.12d. En cada gráfica tenemos para un M concreto un barrido en SNR.

esto puede deberse a que la base de este método es la correlación y no es una función que se vea agravada cuando mayores desalineamientos están presentes (siempre y cuando la señal de interés esté contenida), además no tenemos una aproximación dentro de este método que se vea tan afectada por los desalineamientos en la misma medida que el método de autovalores. Lo único estimado para el inicio de este método es la señal $s(n)$, esta se ve más afectada por los otros dos parámetros, M y SNR.

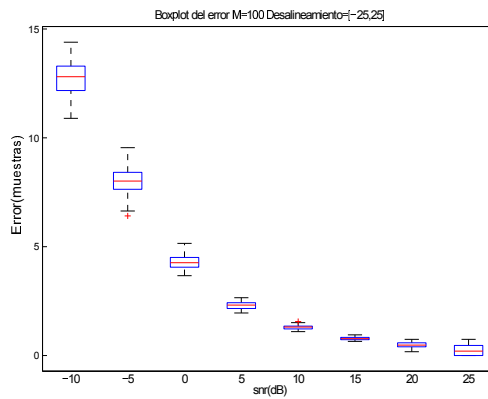
Pasamos a ver ahora la distribución estadística del error en diferentes boxplot, primero para un rango de $\theta = [-25, 25]$,



(a)



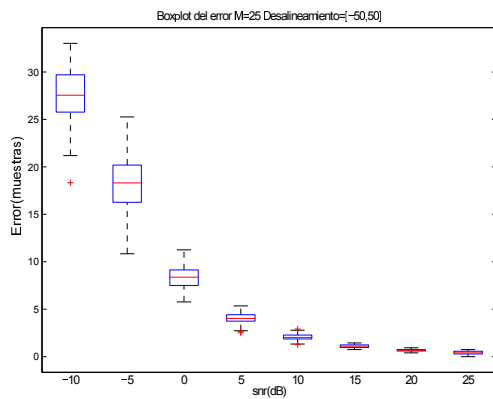
(b)



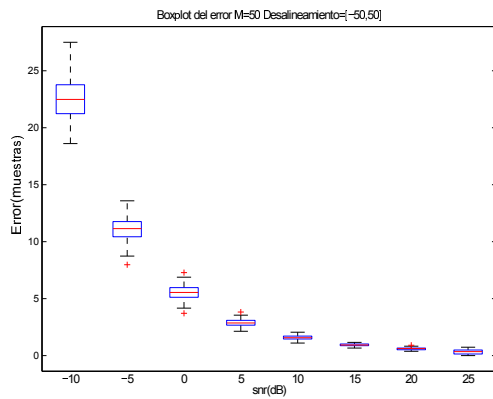
(c)

Figura 3.13: Observamos los boxplot para un rango de desalineamiento θ de ± 25 muestras: $M = 25$ 3.13a, $M = 50$ 3.13b, $M = 100$ 3.13c. En cada gráfica tenemos para un M concreto un barrido en SNR.

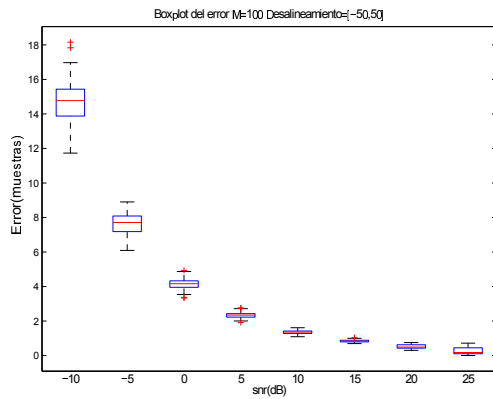
Vemos el mismo comportamiento con un rango desalineamiento mayor, ilustramos el fenómeno para cotas de θ igual a ± 50 y ± 100 muestras.



(a)

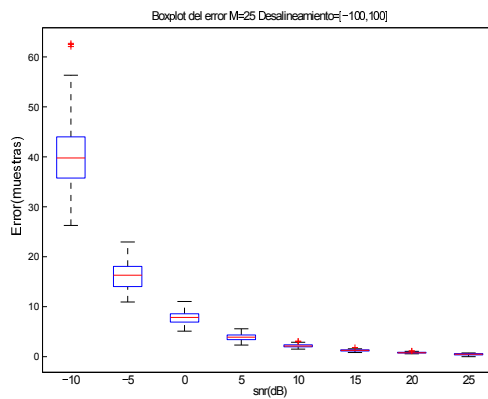


(b)

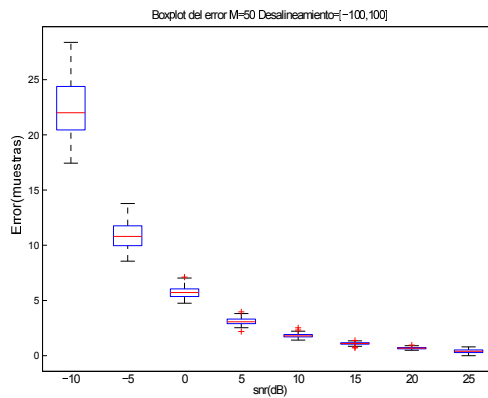


(c)

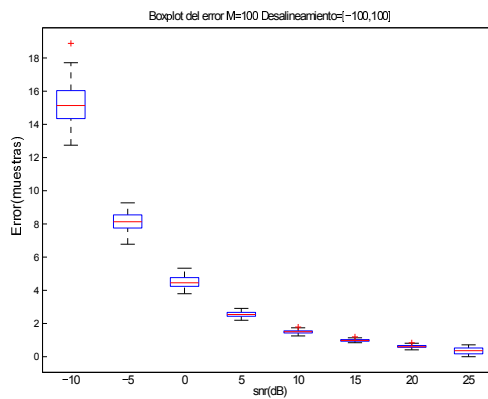
Figura 3.14: Observamos los boxplot para un rango de desalineamiento θ de ± 50 muestras: $M = 25$ 3.14a, $M = 50$ 3.14b, $M = 100$ 3.14c. En cada gráfica tenemos para un M concreto un barrido en SNR.



(a)



(b)



(c)

Figura 3.15: Observamos los boxplot para un rango de desalineamiento θ de ± 100 muestras: $M = 25$ 3.15a, $M = 50$ 3.15b, $M = 100$ 3.15c. En cada gráfica tenemos para un M concreto un barrido en SNR.

Vemos en los boxplot una situación en la que el método de autovalores es más óptimo que el método de Woody. Para valores de M bajos (30-25 realizaciones o menos) con una SNR también baja (0 dB o menor), el método propuesto en este TFG tiene mejores prestaciones en cuanto a

error cometido. Debemos mencionar también que aunque el error cometido por el método de autovalores es menor, este dista de ser la situación ideal de alineamiento como podemos ver en los boxplot del apartado anterior, podríamos aplicar una versión iterativa del método de autovalores. Vemos unas figuras para aclarar este fenómeno.

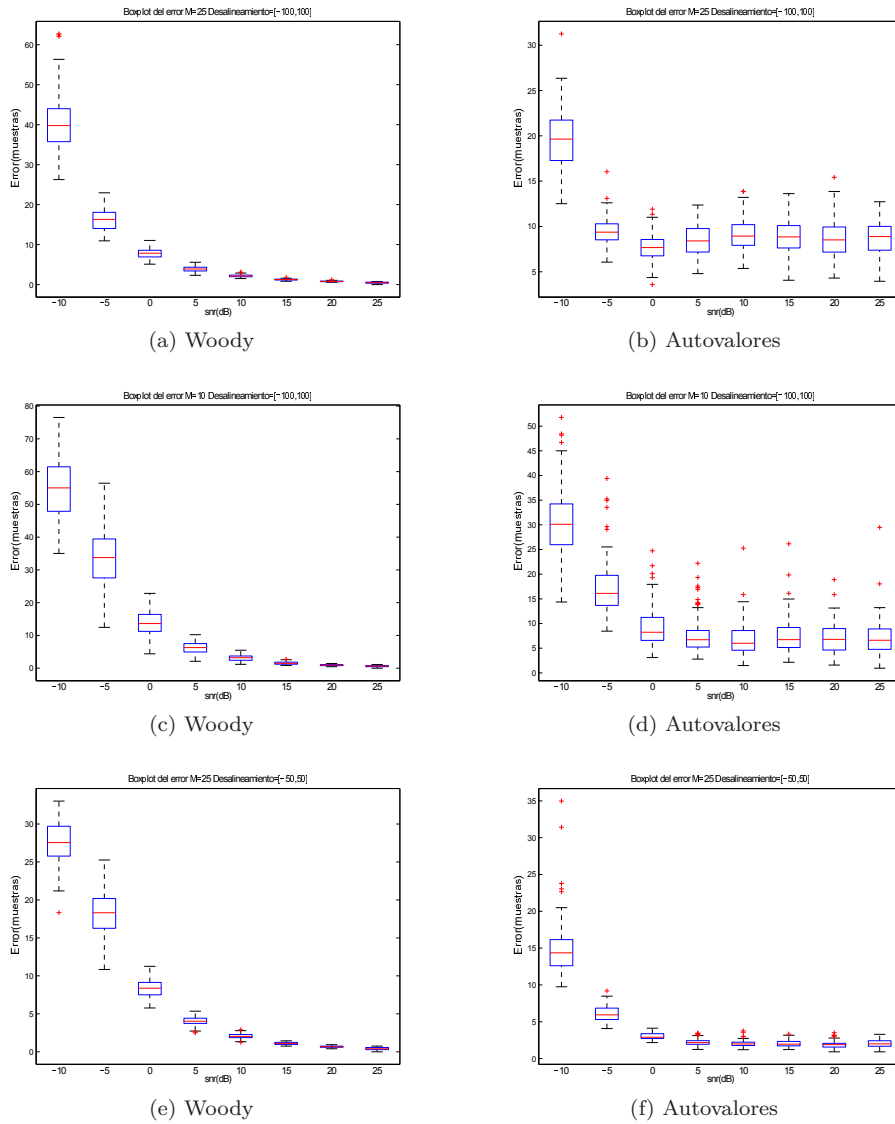


Figura 3.16: Observamos el boxplot del error para 2 rangos de $\theta \pm 50$ y ± 100 de diferentes métodos de alineamiento (Woody y autovalores): $\theta = [-100, 100]$ y $M = 25$ 3.16a, $\theta = [-100, 100]$ y $M = 25$ 3.16b, $\theta = [-100, 100]$ y $M = 10$ 3.16c, $\theta = [-100, 100]$ y $M = 10$ 3.16d $\theta = [-50, 50]$ y $M = 25$ 3.16e y $\theta = [-50, 50]$ y $M = 25$ 3.16f. En cada gráfica tenemos para un M concreto un barrido en SNR.

Podemos decir que esta situación tan perjudicial para el método de Woody (gran desalineamiento, bajo M y SNR) afecta directamente al núcleo del método, la estimación de la señal de interés $s(n)$ con la que correlaremos todas las realizaciones. Aunque el método de Woody sea iterativo podría darnos lugar a pensar que a medida que pasen la iteraciones daríamos finalmente con una buena estimación de $s(n)$ y por tanto con los desalineamientos, pero unas condiciones iniciales tan pésimas hacen que el método de Woody converja a una solución de error considerable.

2. Para los otros dos parámetros, M y SNR, podemos decir que encontramos un comportamiento similar al descrito en el apartado anterior (3.1.1). A medida que ambos crecen, o alguno de ellos lo hace, el error disminuye y las distribuciones de los boxplot son más compactas.

Esto sucede debido a que una población de realizaciones mayor da lugar a una estimación de $s(n)$ más robusta (se reduce el ruido presente al promediar). Tener menor presencia de ruido en un número de realizaciones dado también facilita el obtener una estimación más próxima a la ideal. Este fenómeno viene descrito en la literatura, [8].

Podemos comprobar este comportamiento decreciente en el error en las figuras de este mismo apartado.

Volviendo a la fiabilidad del método de Woody, podemos decir que independientemente del desalineamiento presente, si M y SNR son elevadas el desalineamiento estimado converge a una solución de error bajo y además compacto ya que como podemos ver en los boxplot (figuras 3.13, 3.14 y 3.15) los percentiles se agrupan en un espacio inferior a una muestra.

Sin embargo este comportamiento en el método de autovalores se va alejando de la situación ideal conforme el desalineamiento presente aumenta. Podemos decir que hasta un rango de ± 50 muestras nos encontramos ante una situación aceptable, por encima de este rango el error queda estancado en valores cada vez mas alto independientemente de tener un M y SNR elevados (comprobamos en figura 3.1).

En cuanto el sesgo podemos decir que los tres métodos de alineamiento tiene un comportamiento similar al visto en la figura 3.8. Pocas realizaciones y gran desalineamiento presente, el sesgo aumenta. A medida que M aumenta o el desalineamiento disminuye el sesgo baja de forma notable.

A modo resumen observamos unas tablas donde agrupamos los errores medios del método autovalores respecto los del método de Woody, lo expresamos en %. Una cifra positiva destaca que el error medio es mayor con el método de autovalores, y esta cantidad esta referida al error medio obtenido con los mismos parámetros pero llevado a cabo con el método de Woody. A continuación la expresión matemática utilizada en los siguientes cuadros.

$$\frac{\mathcal{E}_{\text{autovalores}} - \mathcal{E}_{\text{woody}}}{\mathcal{E}_{\text{woody}}} 100 \quad (3.5)$$

Destacamos de nuevo que las celdas de los cuadros 3.2 y 3.3 donde obtenemos un valor negativo tras aplicar (3.5), el método de autovalores mejora en resultado frente a lo obtenido en el método de Woody.

M expresa el número de realizaciones y SNR los dB.

Cuadro 3.2: M.Autovalores vs M.Woody, Retardos=[-25,25] muestras

M \ SNR	-10	-5	0	5	10	15	20	25
5	26,4 %	-17,9 %	-54,7 %	-78,7 %	-72,2 %	-42,2 %	16,1 %	117,7 %
10	48,0 %	-10,1 %	-67,6 %	-79,7 %	-75,1 %	-60,2 %	-28,2 %	20,8 %
15	44,4 %	-27,6 %	-72,5 %	-76,0 %	-71,2 %	-59,6 %	-31,3 %	15,5 %
20	47,1 %	-31,2 %	-71,7 %	-73,0 %	-66,5 %	-49,6 %	-22,6 %	18,3 %
25	43,9 %	-37,3 %	-70,4 %	-68,7 %	-60,8 %	-42,4 %	-14,8 %	30,8 %
30	47,1 %	-43,6 %	-66,4 %	-65,7 %	-57,7 %	-40,8 %	-7,3 %	48,0 %
50	38,3 %	-46,1 %	-57,0 %	-54,8 %	-44,1 %	-22,8 %	12,6 %	71,5 %
75	28,1 %	-42,3 %	-47,9 %	-46,4 %	-35,7 %	-12,9 %	28,6 %	90,4 %
100	23,6 %	-35,8 %	-43,0 %	-39,8 %	-30,1 %	-4,1 %	37,9 %	93,0 %

Cuadro 3.3: M.Autovalores vs M.Woody, Retardos=[-50,50] muestras

M \ SNR	-10	-5	0	5	10	15	20	25
5	-16,3 %	-46,7 %	-71,8 %	-64,1 %	-40,2 %	27,4 %	155,1 %	333,4 %
10	-16,2 %	-67,5 %	-75,3 %	-68,8 %	-46,5 %	-11,0 %	85,6 %	213,3 %
15	-28,0 %	-72,2 %	-72,7 %	-61,2 %	-31,0 %	33,9 %	112,4 %	216,8 %
20	-35,6 %	-70,6 %	-68,0 %	-54,7 %	-18,7 %	59,0 %	178,1 %	309,8 %
25	-44,0 %	-66,5 %	-64,0 %	-44,3 %	-2,0 %	87,2 %	189,1 %	380,5 %
30	-47,2 %	-64,6 %	-59,1 %	-35,9 %	12,2 %	101,6 %	209,5 %	459,6 %
50	-47,6 %	-50,9 %	-43,3 %	-11,9 %	47,2 %	141,3 %	276,8 %	461,7 %
75	-38,9 %	-43,0 %	-30,8 %	5,9 %	74,1 %	169,5 %	324,3 %	556,6 %
100	-30,6 %	-33,7 %	-19,6 %	18,0 %	85,4 %	193,2 %	351,6 %	639,2 %

Podemos ver en estas tablas de forma más clara la relación entre el error medio cometido por ambos métodos que también hemos visto de forma numérica en varias figuras, para el método autovalores en la figura 3.1 y para el método de Woody en la figura 3.12.

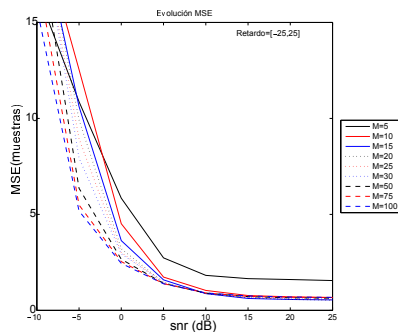
Capítulo 4

Conclusiones y líneas futuras

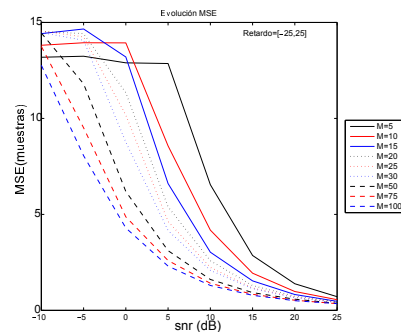
Tras llevar a cabo la implementación de los tres métodos de alineamiento en Matlab, autovalores, correlación y Woody hemos podido comprobar en el apartado de resultados los errores entre el retardo real y el estimado por cada método de alineamiento. También hemos podido comprobar la distribución del error viendo así la variabilidad del estimador para unos parámetros dados.

Se ha observado en un entorno acotado la fiabilidad del método de alineamiento basado en los autovalores de la matriz de autocorrelación del proceso, el cuál disminuye computacionalmente el tiempo necesario para la obtención de los desalineamientos de las M realizaciones respecto los otros dos métodos de alineamiento vistos, correlación y Woody. La eficiencia en cálculo era una de las principales apuestas en el momento de desarrollo de este nuevo método de alineamiento y hemos podido comprobar que cumple.

En cuanto a los resultados en forma de error, podemos decir que el método de autovalores esta muy próximo a los resultados obtenidos por los otros dos métodos o incluso en mejor situación (al menos mientras θ alcanza valores hasta ± 50 muestras). Lo vemos en unas gráficas.



(a) Autovalores



(b) Woody

Nos centramos en comparar las dos figuras siguientes, donde hacemos un zoom en la zona de interés.

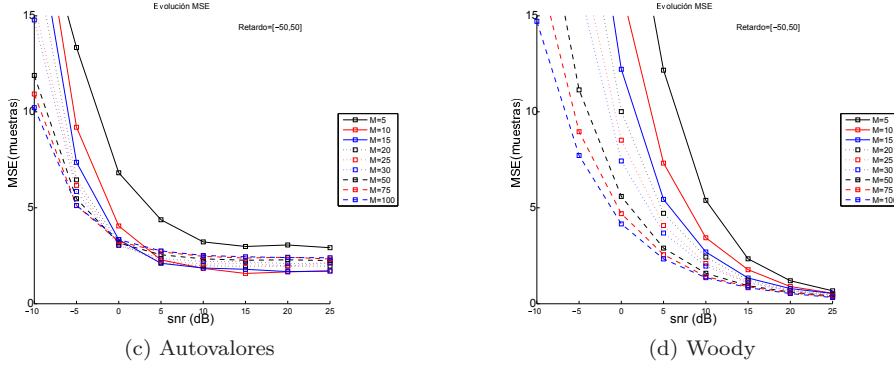


Figura 4.1: Observamos el error en 4 figuras para 2 rangos de desalineamiento: $\theta = [-25, 25]$ 4.1a, $\theta = [-25, 25]$ 4.1b, $\theta = [-50, 50]$ 4.1c y $\theta = [-50, 50]$ 4.1d. En cada gráfica tenemos para cada M diferente un barrido en SNR.

Siempre y cuando el desalineamiento no sea demasiado elevado, inferior a intervalos de $[-50, 50]$ muestras (± 200 ms) a la frecuencia de muestreo usada de 250 Hz, el método de autovalores esta próximo o mejora lo visto con el método de Woody. Hay situaciones donde encontramos a nuestro método superior también, un tanto más atípicas debido al gran ruido presente o bajo número de realizaciones incluidas. Pero debemos decir que en estas situaciones el error sigue siendo considerable, por encima de las 15-30 muestras en algunos casos.

Se ha considerado una implementación iterativa (2 iteraciones) del método de autovalores para los casos de mayor error, dando resultados dispares. Para desalineamientos dentro del intervalo $[-25, 25]$ muestras, con M entorno a las 30 realizaciones o menor obtenemos mejoras del 5 – 10 % respecto la versión no iterativa del método de autovalores, no es así para un mayor número de realizaciones ya que la situación llega a empeorar. Para rangos de θ superiores los resultados obtenidos empeoran o igualan lo obtenido con una iteración. En cualquier caso esto ha sido solo exploratorio y análisis más sistemáticos son necesarios para tomar conclusiones definitivas.

En cuanto la variabilidad del estimador podemos decir que de nuevo, con algún valor atípico en alguno de los boxplot, podemos ver una baja variabilidad en la solución dada. Siendo las soluciones con distribución estadística del error cada vez más compactas conforme la SNR aumenta, con el aumento de M se pueden ver ligeras mejoras también, pero menos determinantes. Encontramos por tanto un estimador que converge hacia el entorno reducido de un valor numérico. Debemos remarcar de nuevo que si un desalineamiento elevado esta presente, el error aumenta tanto en valor como en el tamaño de los percentiles vistos en los boxplot. Nos encontramos ante una situación más sensible para el método de autovalores.

En el caso del método de Woody o correlación hemos visto como θ no es un parámetro tan determinante como con el método de autovalores, siendo M y SNR parámetros que afectan de forma más directa al error obtenido. Esto se debe a que la única variable estimada para estos dos métodos, $s(n)$, que es el núcleo principal del método de Woody y correlación, se aleja del valor ideal a estimar llegando a estancar las estimaciones obtenidas de los θ_i en la sucesivas iteraciones de las M realizaciones. Obtenemos la convergencia de estos dos métodos hacia una solución de un error considerable como hemos podido ver en diferentes figuras. Para M y SNR bajos podemos ver los casos de mayor error.

Concluimos diciendo que bajo un entorno controlado que ha habido en simulación, este método de autovalores puede ser una solución eficaz como estimador de alineamiento siempre que el desalineamiento presente no sea demasiado elevado, alrededor de 200 milisegundos ó 50 muestras aproximadamente.

En un entorno real tendremos parámetros de simulación que aquí no fueron tenidos en cuenta, dejamos la inclusión de mayor causalística para una fase de testeo más exhaustiva en una línea futura.

4.1. Líneas futuras

Volvemos a mencionar el entorno controlado al que fue sometido cada uno de los métodos de alineamiento. Sería una buena idea comprobar estos resultados en un problema real de alineamiento donde tendremos una variabilidad de parámetros mayor a lo considerado en este TFG o parámetros aquí no considerados.

Se deja abierto el camino para comprobar las prestaciones del método ante una forma de $s(n)$ no fija a lo largo de las M realizaciones, con ligeras variaciones de amplitud o forma.

En este trabajo se ha considerado un tipo de señal, la respiratoria. Se podría comprobar los resultados de simulación con otras señales de origen biológico repetitivas en el tiempo, como la actividad eléctrica del corazón por ejemplo o de la actividad cerebral (potenciales evocados). Podemos ver otro tipo de señales biológicas aptas para este método de alineamiento en [8].

Para finalizar mencionaremos la posibilidad de realizar estimaciones alternativas que mejoren la obtención de algunas variables necesarias para la búsqueda de los desalineamientos. Ya que estas variables son fundamentales para la obtención de la norma del autovector de alineamiento θ .

Referencias

- [1] A. Garde, P. Laguna, B. F. Giraldo, R. Jane, and L. Sörnmo. Ensemble-based time alignment of biomedical signals. *BSI12*, 2012.
- [2] A. Garde, P. Laguna, B. F. Giraldo, R. Jane, and L. Sörnmo. Eigenvalue-based time alignment of biomedical signals. *manuscript in preparation*.
- [3] E. Gil, R. Bailon, J. M. Vergara, and P. Laguna. PTT variability for discrimination of sleep apnea related decreases in the amplitude fluctuations of PPG signal in children. *IEEE Transactions on Biomedical Engineering*, 57, 2010.
- [4] C. D. Woody. Characterization of an adaptative filter for the analysis of variable latency neuroelectric signals. *Med. Biol. Eng. Comp.*, 5:539–553, 1967.
- [5] T. K. Moon and W. C. Sterling. *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall, 2000.
- [6] A. Cabasson and O. Meste. Time delay estimation: A new insight into the Woody’s method. *IEEE Signal Proc. Letters*, 15:573–576, 2008.
- [7] F. Castells, P. Laguna, L. Sörnmo, A. Bollmann, and J.Millet Roig. Principal component analysis in ECG signal processing. *J. Adv. Signal Proc.*, 2007, 2007.
- [8] L. Sörnmo and P. Laguna. *Bioelectrical signal processing in cardiac and neurological applications*. Elsevier Academic Press, 2005.

Índice de figuras

1.1. Realizaciones mal alineadas	9
1.2. Realizaciones próximas a estar alineadas	9
1.3. Fragmento del ciclo respiratorio de un paciente sano, muestreado a 250 Hz y con un periodo aproximadamente de 2'5 segundos.	10
2.1. Desalineamiento de 4 a 160 ms, en verde la señal original $s(t)$ y alineada. En diferentes colores $\hat{s}_T(t)$ para $\theta_i = 4, 8, \dots, 160$ ms	15
2.2. Desalineamiento de 20 ms ó 5 muestras, en rojo la señal original desplazada $s(t - 0,02)$. En azul $\hat{s}_T(t)$ para $\theta_i = 20$ ms	15
2.3. Desalineamiento de 60 ms ó 15 muestras, en rojo la señal original desplazada $s(t - 0,06)$. En azul $\hat{s}_T(t)$ para $\theta_i = 60$ ms	16
2.4. Desalineamiento de 160 ms ó 40 muestras, en rojo la señal original desplazada $s(t - 0,16)$. En azul $\hat{s}_T(t)$ para $\theta_i = 160$ ms	16
2.5. SNR de 25 dB	18
2.6. SNR de 10 dB	19
2.7. Vemos el desalineamiento en muestras real (azul) y el estimado a partir del segundo autovector ya desnormalizado de \mathbf{R}_x^\bullet (rojo) para cada una de las 30 realizaciones.	24
2.8. Rango de desalineamiento de [-100,100] muestras, M=30	25
2.9. Rango de desalineamiento de [-50,50] muestras, M=30	26
2.10. Rango de desalineamiento de [-100,100] muestras, M=10	26
2.11. Rango de desalineamiento de [-50,50] muestras, M=10	27
2.12. Observamos el efecto de la latencia respecto los máximos	28
2.13. Diagrama de bloques método de correlación	28
2.14. Diagrama de bloques del método de Woody. El procedimiento iterativo finaliza al cumplir (2.52).	31
2.15. El error respecto la estimación del desalineamiento y el retardo real evoluciona de la siguiente forma en media: 1ª iteración 8'93, 2ª iteración 5'64 y 3ª iteración 4'42 muestras en cada caso. Vemos destacado en negro en las figuras 2.15a y 2.15d el promedio de las M realizaciones.	32
3.1. Observamos el error en 4 figuras para los 4 rangos de desalineamiento en muestras disponibles: $\theta = [-25, 25]$ 3.1a, $\theta = [-50, 50]$ 3.1b, $\theta = [-75, 75]$ 3.1c y $\theta = [-100, 100]$ 3.1d. En cada gráfica tenemos para cada M diferente un barrido en SNR.	35
3.2. Observamos los boxplot para un rango de desalineamiento θ de ± 25 muestras: $M = 25$ 3.2a, $M = 50$ 3.2b, $M = 100$ 3.2c. En cada gráfica tenemos para un M concreto un barrido en SNR.	36

3.3.	Observamos los boxplot para un rango de desalineamiento θ de ± 50 muestras: $M = 25$ 3.3a, $M = 50$ 3.3b, $M = 100$ 3.3c. En cada gráfica tenemos para un M concreto un barrido en SNR. . .	37
3.4.	Observamos los boxplot para un rango de desalineamiento θ de ± 100 muestras: $M = 25$ 3.4a, $M = 50$ 3.4b, $M = 100$ 3.4c. En cada gráfica tenemos para un M concreto un barrido en SNR. . .	38
3.5.	Observamos la distorsión del autovector $\mathbf{1}$ desnormalizado para diferentes θ con M y SNR fijas: $\theta = \pm 25$ 3.5a, $\theta = \pm 50$ 3.5b, $\theta = \pm 75$ 3.5c y $\theta = \pm 100$ 3.5d. Vemos además el error respecto el autovector ideal de forma numérica. Este aumenta de forma proporcional con θ al igual de lo que sucedía con el error de alineamiento.	39
3.6.	Observamos para los 4 rangos de θ , ± 25 , ± 50 , ± 75 y ± 100 muestras el error de alineamiento: $M = 10$ 3.6a, $M = 25$ 3.6b, $M = 50$ 3.6c y $M = 100$ 3.6d. En cada gráfica tenemos para un M concreto un barrido en SNR.	40
3.7.	Observamos los boxplot para un rango de desalineamiento θ de ± 50 muestras: $M = 5$ 3.7a, $M = 25$ 3.7b, $M = 50$ 3.7c y $M = 100$ 3.7d. En cada gráfica tenemos para un M concreto un barrido en SNR.	41
3.8.	Observamos el sesgo del error para los 4 rangos de θ : SNR = -10 dB 3.8a, SNR = 0 dB 3.8b, SNR = 10 dB 3.8c y SNR = 20 dB 3.8d. En cada gráfica tenemos para un SNR concreto un barrido en M	42
3.9.	Observamos el error para los 4 rangos de θ ± 25 , ± 50 , ± 75 y ± 100 : $\theta = [-25, 25]$ 3.9a, $\theta = [-25, 25]$ 3.9b, $\theta = [-50, 50]$ 3.9c, $\theta = [-50, 50]$ 3.9d, $\theta = [-75, 75]$ 3.9e, $\theta = [-75, 75]$ 3.9f $\theta = [-100, 100]$ 3.9g y $\theta = [-100, 100]$ 3.9h. En cada gráfica tenemos para un M concreto un barrido en SNR.	44
3.10.	Observamos los boxplot para un rango de desalineamiento θ de ± 25 muestras: $M = 10$ 3.10a, $M = 10$ 3.10b, $M = 30$ 3.10c , $M = 30$ 3.10d, $M = 100$ 3.10e , $M = 100$ 3.10f. En cada gráfica tenemos para un M concreto un barrido en SNR.	45
3.11.	Observamos los boxplot para un rango de desalineamiento θ de ± 100 muestras: $M = 10$ 3.11a, $M = 10$ 3.11b, $M = 30$ 3.11c , $M = 30$ 3.11d, $M = 100$ 3.11e , $M = 100$ 3.11f. En cada gráfica tenemos para un M concreto un barrido en SNR.	46
3.12.	Observamos el error para 4 rangos de θ : $\theta = [-25, 25]$ 3.12a, $\theta = [-50, 50]$ 3.12b, $\theta = [-75, 75]$ 3.12c y $\theta = [-100, 100]$ 3.12d. En cada gráfica tenemos para un M concreto un barrido en SNR.	47
3.13.	Observamos los boxplot para un rango de desalineamiento θ de ± 25 muestras: $M = 25$ 3.13a, $M = 50$ 3.13b, $M = 100$ 3.13c. En cada gráfica tenemos para un M concreto un barrido en SNR. . .	48
3.14.	Observamos los boxplot para un rango de desalineamiento θ de ± 50 muestras: $M = 25$ 3.14a, $M = 50$ 3.14b, $M = 100$ 3.14c. En cada gráfica tenemos para un M concreto un barrido en SNR. . .	49
3.15.	Observamos los boxplot para un rango de desalineamiento θ de ± 100 muestras: $M = 25$ 3.15a, $M = 50$ 3.15b, $M = 100$ 3.15c. En cada gráfica tenemos para un M concreto un barrido en SNR.	50

3.16. Observamos el boxplot del error para 2 rangos de $\theta \pm 50$ y ± 100 de diferentes métodos de alineamiento (Woody y autovalores): $\theta = [-100, 100]$ y $M = 25$ 3.16a, $\theta = [-100, 100]$ y $M = 25$ 3.16b, $\theta = [-100, 100]$ y $M = 10$ 3.16c, $\theta = [-100, 100]$ y $M = 10$ 3.16d $\theta = [-50, 50]$ y $M = 25$ 3.16e y $\theta = [-50, 50]$ y $M = 25$ 3.16f. En cada gráfica tenemos para un M concreto un barrido en SNR. . .	51
4.1. Observamos el error en 4 figuras para 2 rangos de desalineamiento: $\theta = [-25, 25]$ 4.1a, $\theta = [-25, 25]$ 4.1b, $\theta = [-50, 50]$ 4.1c y $\theta = [-50, 50]$ 4.1d. En cada gráfica tenemos para cada M diferente un barrido en SNR.	55

Índice de cuadros

3.1. Vemos el error en muestras, donde fijamos $M = 30$ y $\text{SNR} = 10$ dB.	39
3.2. M.Autovalores vs M.Woody, Retardos= $[-25,25]$ muestras	53
3.3. M.Autovalores vs M.Woody, Retardos= $[-50,50]$ muestras	53

Apéndice A

Estimador de máxima verosimilitud

Definimos a continuación el procedimiento seguido por este estimador centrándonos en nuestra situación. Nos encontramos ante un estimador de parámetros de estimación clásica, es decir, nuestros parámetros a estimar no se tratan de variables aleatorias.

Concretamente se busca el valor de estos parámetros que desconocemos pero se comportan de forma determinista, en nuestro caso el desalineamiento, θ .

En nuestra situación partimos de M observaciones o realizaciones independientes entre si. Por ello podremos describir la función de densidad de probabilidad conjunta de la siguiente forma,

$$p(\mathbf{x}|\theta) = f(x_1|\theta)f(x_2|\theta) \cdots f(x_M|\theta). \quad (\text{A.1})$$

El siguiente paso será buscar el conjunto de θ que maximiza la función anterior. Es decir que conjunto de parámetros hace más verosímil la ocurrencia del suceso o con que valor del parámetro se hace máxima la función de densidad conjunta.

$$\hat{\theta} = \arg \max_{\theta} p(\mathbf{x}|\theta). \quad (\text{A.2})$$

Apéndice B

Entorno alineamiento

Cabe destacar que algunos de los estimadores de retardo que implementamos en este TFG se basan en la correlación. A continuación recurriremos a la literatura [8] para poner de manifiesto que esta función es una solución óptima para el problema que encontramos en nuestro entorno de alineamiento.

Ya descrito nuestro entorno de alineamiento en el capítulo 2, pasamos de forma directa a la obtención del estimador de máxima verosimilitud.

Planteamos a continuación la forma de onda que podría tener cada una de las M realizaciones,

$$x_i(n) = \begin{cases} v_i(n), & n = 0, \dots, \theta_i - 1; \\ s(n - \theta_i) + v_i(n), & n = \theta_i, \dots, \theta_i + D - 1; \\ v_i(n), & n = \theta_i + D, \dots, N - 1, \end{cases} \quad (\text{B.1})$$

donde vemos que en un principio solo tenemos ruido, $v_i(n)$, a partir del instante θ_i comienza nuestra señal de interés y tras pasar D muestras volvemos a tener solo el ruido $v_i(n)$. Recordamos que en el capítulo 2 se comenta que la forma de onda de $s(n)$ es fija y asumiremos que inicialmente es conocida o tendremos los medios para poder estimarla. Además estará totalmente contenida en cada una de las M realizaciones, independientemente del valor de θ_i , por ello la energía de E_s también será fija,

$$E_s = \sum_{n=\theta_i}^{\theta_i+D-1} s^2(n - \theta_i) = \sum_{n=0}^{N-1} s^2(n). \quad (\text{B.2})$$

A continuación vemos el aspecto de las funciones de densidad de probabilidad a partir de las cuales llegaremos a nuestro estimador de máxima verosimilitud. Asumimos que el ruido es gaussiano, blanco (incorrelado y plano en frecuencia.) y de media nula. Las funciones de densidad de probabilidad conjunta,

$$\begin{aligned}
p(\mathbf{x}_i|\theta_i) &= \prod_{n=0}^{\theta_i-1} \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[-\frac{x_i^2(n)}{2\sigma_v^2}\right] \\
&\cdot \prod_{n=\theta_i}^{\theta_i+D-1} \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[-\frac{(x_i(n) - s(n - \theta_i))^2}{2\sigma_v^2}\right] \\
&\cdot \prod_{n=\theta_i+D}^{N-1} \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[-\frac{x_i^2(n)}{2\sigma_v^2}\right],
\end{aligned} \tag{B.3}$$

donde $\mathbf{x}_i = [x_i(0) \ x_i(1) \ \dots \ x_i(N-1)]^T$. Desarrollando el cuadrado del exponente de (B.3), podemos reescribir la función de densidad de probabilidad conjunta condicionada a θ_i ,

$$\begin{aligned}
p(\mathbf{x}_i|\theta_i) &= \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[-\frac{x_i^2(n)}{2\sigma_v^2}\right] \\
&\cdot \prod_{n=\theta_i}^{\theta_i+D-1} \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[\frac{x_i(n)s(n - \theta_i)}{2\sigma_v^2}\right] \\
&\cdot \prod_{n=\theta_i}^{\theta_i+D-1} \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[-\frac{s^2(n - \theta_i)}{2\sigma_v^2}\right],
\end{aligned} \tag{B.4}$$

donde el primer productorio no es una función que dependa de θ_i , y el tercer productorio por lo visto en (B.2) se mantiene con una valor constante también. Para simplificar más la expresión anterior aplicamos logaritmo neperiano a ambos lado de la igualdad,

$$\ln p(\mathbf{x}_i|\theta_i) = C + \frac{1}{\sigma_v^2} \sum_{n=\theta_i}^{\theta_i+D-1} x_i(n)s(n - \theta_i). \tag{B.5}$$

En C agrupamos los valores constantes independientes de θ_i . El estimador de máxima verosimilitud nos dará el valor de θ_i que maximiza el sumatorio visto en (B.5),

$$\hat{\theta}_i = \arg \max_{\theta_i} \left(\sum_{n=\theta_i}^{\theta_i+D-1} x_i(n)s(n - \theta_i) \right). \tag{B.6}$$

Apéndice C

Variabilidad de amplitud

El método de alineamiento propuesto en este TFG asume que la amplitud de las señales en cada realización es idéntica, esto lo podemos ver al inicio del capítulo 2, recurrimos de nuevo a [2]. Esto no es del todo cierto, por ejemplo los latidos del corazón pueden cambiar debido diversos factores (por el esfuerzo realizado, la respiración) de la misma forma la señal de interés puede sufrir ligeros cambios en su amplitud. Es por tanto, la variación de amplitud de la señal de interés $s(n)$, un parámetro que puede influir en nuestro método de alineamiento. Volvemos a hacer inciso que en este TFG no se llegó a hacer uso de este parámetro en simulación.

Extendemos el modelo de cada realización que planteamos en (2.1), inicio del capítulo 2

$$x_i(n) = a_i s(n - \theta_i) + v_i(n), \quad n = 0, \dots, N - 1; \quad i = 1, \dots, M. \quad (\text{C.1})$$

Donde la amplitud a_i se representa por una variable aleatoria de media $E[a_i] = m_a = 1$ y una varianza $\sigma_a^2 (<< m_a)$. Las variables a_i y θ_i se asumen que son incorreladas. El resto de los términos de (C.1) quedan descritos a lo largo del capítulo 2.

Vemos el aspecto de la matriz de autocorrelación, vemos primeramente el vector que lo forma en detalle

$$\mathbf{x}_a = a_i \mathbf{s} - a_i \theta_i \mathbf{s}' + \mathbf{v}, \quad (\text{C.2})$$

recordamos que el ruido \mathbf{v} es de media nula, y esta incorrelado con las señales \mathbf{s} y \mathbf{s}' , a su vez estas dos últimas son ortogonales entre si como comprobamos en (2.16). Vemos la matriz,

$$\mathbf{R}_x^a = E[\mathbf{x}_a \mathbf{x}_a^T] = a_i^2 \mathbf{s} \mathbf{s}^T + \theta_i^2 a_i^2 \mathbf{s}' \mathbf{s}'^T + \sigma_v^2 \mathbf{I}, \quad (\text{C.3})$$

ahora sus autovalores,

$$\lambda_i = \begin{cases} (\sigma_a^2 + 1)E_s + \sigma_v^2, & i = 1; \\ (\sigma_a^2 + 1)\sigma_\theta^2 E_{s'} + \sigma_v^2, & i = 2; \\ \sigma_v^2, & i = 3, \dots, N, \end{cases} \quad (\text{C.4})$$

la función ratio de autovalores tiene un aspecto muy similar a la vista en el capítulo 2, $\Lambda(\theta)$ (2.21) con la incorporación del término de variabilidad de amplitud σ_a^2 . Vemos la expresión a continuación,

$$\Lambda_a(\boldsymbol{\theta}) = \frac{\lambda_1(\boldsymbol{\theta})}{\sum_{i=2}^N \lambda_i(\boldsymbol{\theta})} \approx \frac{E_s}{\sigma_\theta^2 E_{s'} + \frac{(N-1)\sigma_a^2}{\sigma_a^2 + 1}}. \quad (\text{C.5})$$

Vemos que el comportamiento de la función ratio sigue maximizándose a medida que alineamos las M realizaciones de forma correcta. Sigue siendo válido entonces pese a la incorporación de la variación de amplitud.

Para finalizar este anexo encontramos en la literatura un detalle muy importante acerca de este parámetro que aporta la variabilidad de amplitud [6], dadas las posibles complicaciones de considerar este parámetro ya que pueden aparecer singularidades en la solución. Por este motivo en la versión mejorada del método de Woody así como en la versión inicial del método de Woody no se contempla la variabilidad de amplitud, por lo tanto es un parámetro que no consideraremos en el apartado 2.1.2 pero que si podríamos aplicar en el apartado 2.1.1 como ya hemos visto en este anexo.