

4 Simulaciones

En este capítulo se realizará un estudio comparativo a nivel de software del potencial y la capacidad de los dos algoritmos de reconstrucción estudiados, Matching Pursuit y Orthogonal Matching Pursuit, en la reconstrucción de señales sparse en el dominio del tiempo. Para ello se realizará la reconstrucción de la misma señal para ambos algoritmos en el dominio del tiempo, y a continuación se generarán las curvas de error cuadrático medio y de probabilidad de reconstrucción en función de las medidas (proyecciones aleatorias) y del nivel de escasez (k -sparse), también se calculará el factor de sub-muestreo característico de cada método de reconstrucción.

4.1 Reconstrucción de señales sparse en el dominio del tiempo.

A modo de comparar el error generado por los dos algoritmos en la reconstrucción, se utilizará la señal que se presenta en la figura 17:

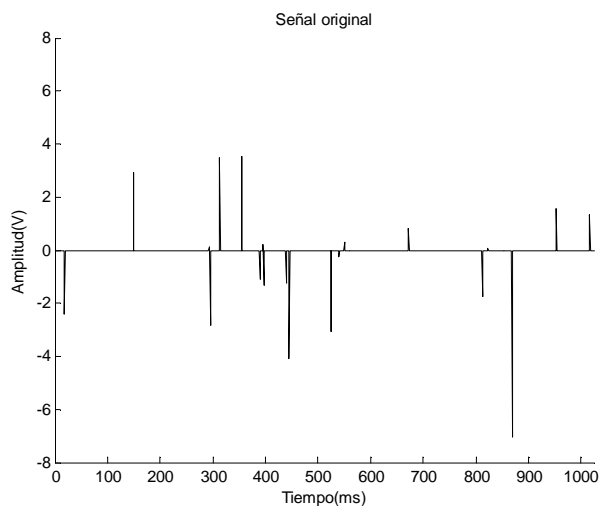


Fig17. Señal sparse en el dominio del tiempo de longitud $n = 1024$ y escasez $k = 20$, empleada en la reconstrucción usando los algoritmos MP y OMP.

Esta señal se reconstruirá utilizando los dos métodos de reconstrucción presentados en el apartado anterior y posteriormente se calculará el error medio cuadrático que existe entre la señal original y la señal reconstruida.

4.1.1 Reconstrucción usando MP

Sea x la señal sparse en el dominio del tiempo representada en la figura 17 con un tamaño de 1024 muestras donde sólo $k = 20$ son distintas de cero e y el vector de medición con 300 proyecciones aleatorias de la señal; la figura 18 muestra ambas señales. Al aplicar el algoritmo de reconstrucción MP con la señal y las proyecciones anteriores se obtienen las señales mostradas en las figuras 18(c) y 18(d), obsérvese que con una proyección aleatoria de 300 medidas la reconstrucción se puede considerar exitosa con una diferencia entre la señal original y la reconstruida en el orden de 10^{-8} .

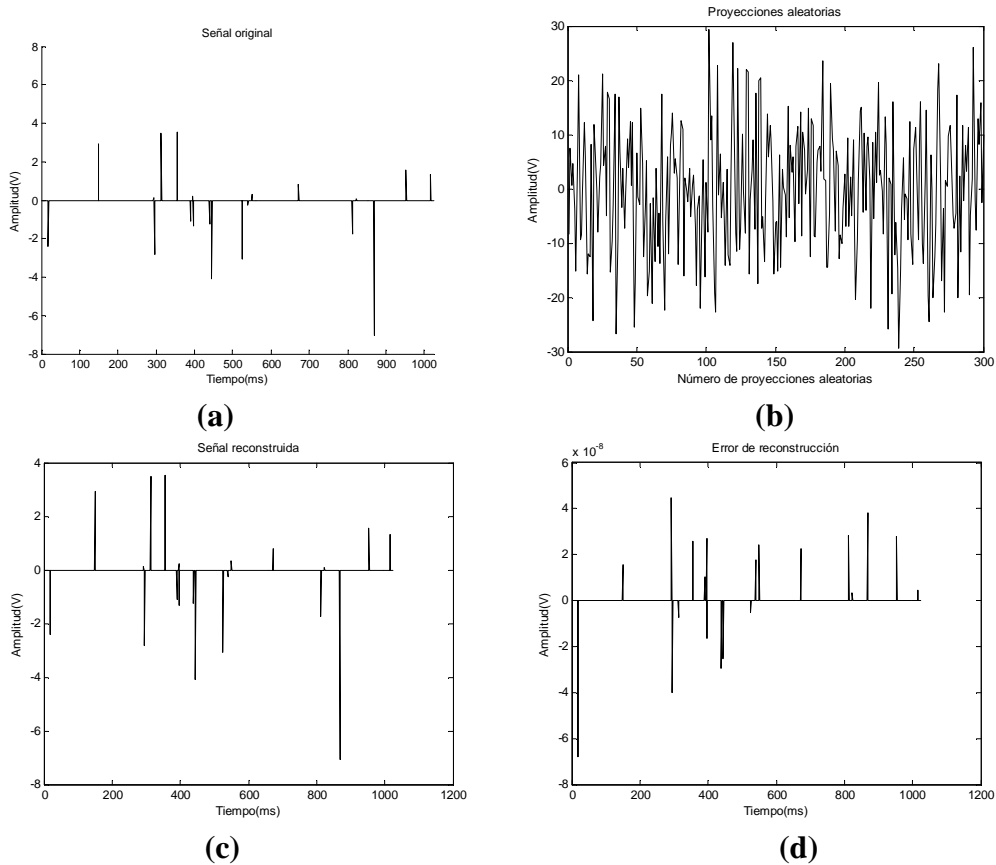


Fig18. Reconstrucción de una señal sparse en el dominio del tiempo empleando el algoritmo MP con $k = 20$ y 300 proyecciones aleatorias. **(a)** Señal original. **(b)** Proyecciones aleatorias. **(c)** Señal reconstruida. **(d)** Error de reconstrucción.

Si se calcula el error cuadrático medio definido del siguiente modo:

$$(54) \quad E_{mc} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}}$$

dónde e viene dado por la diferencia entre la señal original y la señal reconstruida. Se obtiene $E_{mc} = 1,5242 \times 10^{-14}$, como se puede ver si se compara con la amplitud original de la señal, el error medio cuadrático obtenido es prácticamente nulo ya que está diez órdenes de magnitud por debajo.

Sin embargo, si el número de proyecciones se reduce, por ejemplo en tres veces, posiblemente el número de mediciones no sea suficiente para la reconstrucción de la señal. La figura 19 muestra los resultados para tal situación, es decir, tomando tan solo 100 proyecciones.

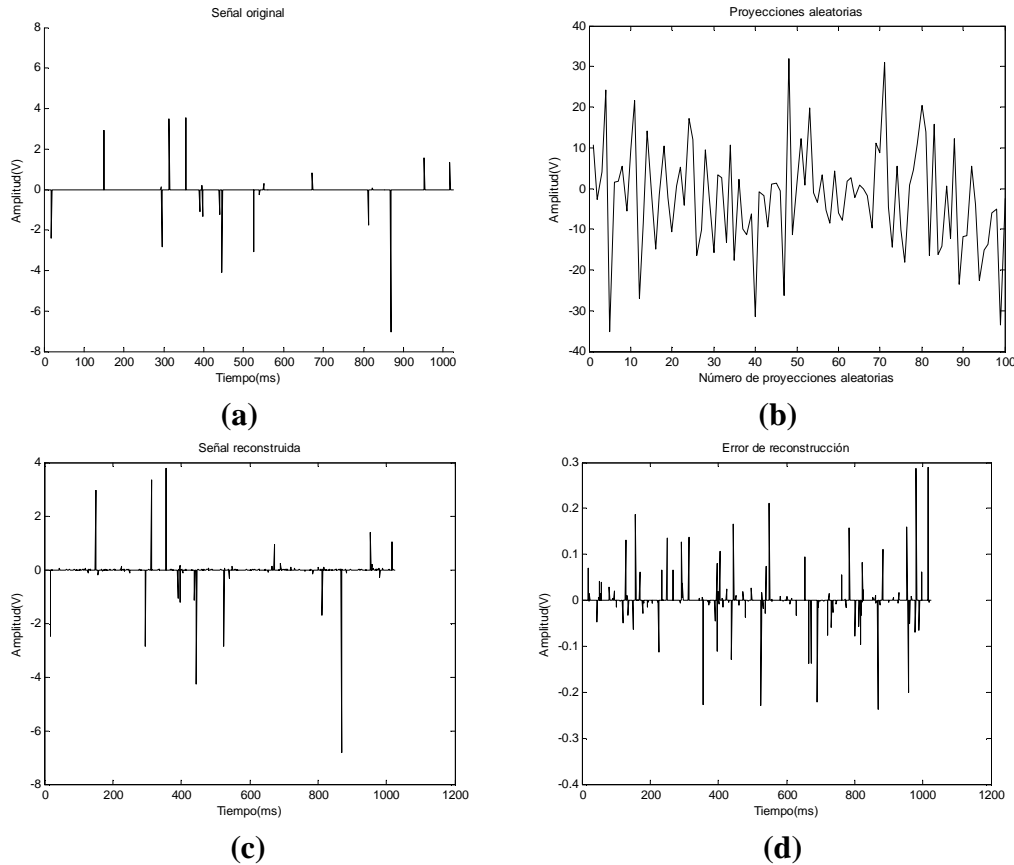


Fig19. Reconstrucción de una señal sparse en el dominio del tiempo empleando el algoritmo MP con $k = 20$ y 100 proyecciones aleatorias. **(a)** Señal original. **(b)** Proyecciones aleatorias. **(c)** Señal reconstruida. **(d)** Error de reconstrucción.

4.1.2 Reconstrucción usando OMP

Sea x la señal sparse en el dominio del tiempo representada en la figura 17 con un tamaño de 1024 muestras donde sólo $k = 20$ son distintas de cero e y el vector de medición con 300 proyecciones aleatorias de la señal; la figura 20 muestra ambas señales. Al aplicar el algoritmo de reconstrucción OMP con la señal y las proyecciones anteriores se obtienen las señales mostradas en las figuras 20(c) y 20(d), obsérvese que con una proyección aleatoria de 300 medidas la reconstrucción se puede considerar exitosa con una diferencia entre la señal original y la reconstruida en el orden de 10^{-15} .

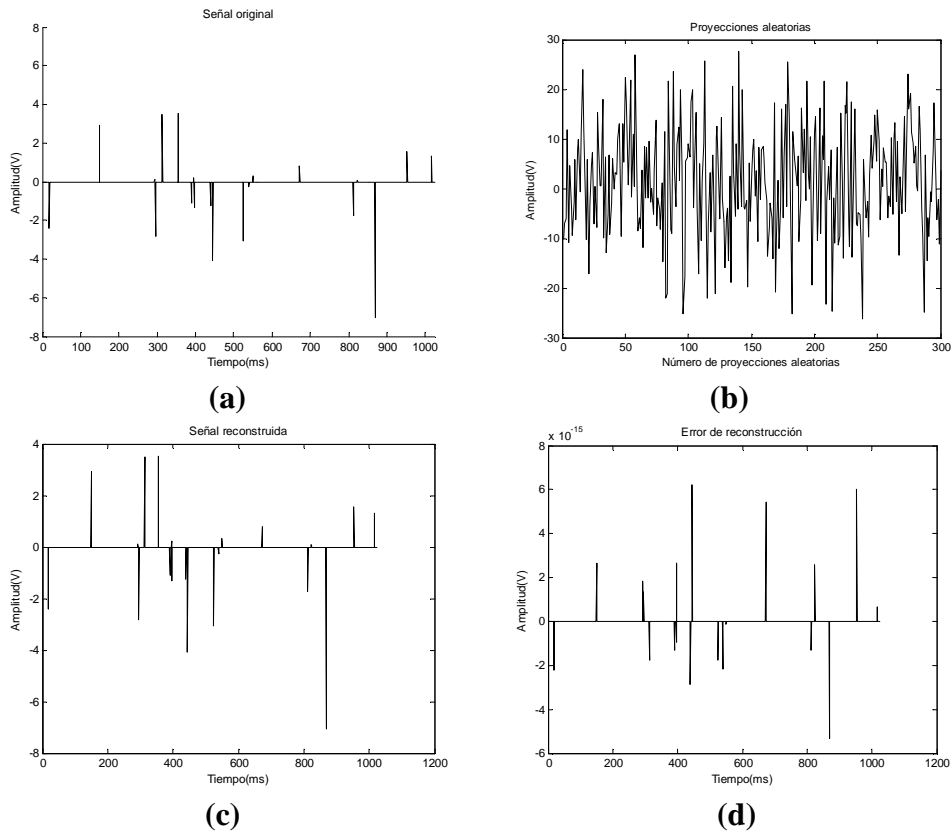


Fig20. Reconstrucción de una señal sparse en el dominio del tiempo empleando el algoritmo OMP con $k = 20$ y 300 proyecciones aleatorias. **(a)** Señal original. **(b)** Proyecciones aleatorias. **(c)** Señal reconstruida. **(d)** Error de reconstrucción.

Si se calcula el error cuadrático medio definido en la ecuación (70), se obtiene $E_{mc} = 1,1242 \times 10^{-28}$, como se puede ver si se compara con la amplitud original de la señal, el error medio cuadrático obtenido es prácticamente nulo ya que está más de diez órdenes de magnitud por debajo.

En este caso comprobamos que si el número de proyecciones se reduce en tres veces, como en el caso anterior, el error de reconstrucción obtenido es prácticamente igual al conseguido con un número de mediciones igual a 300, pero si reducimos este número de mediciones a 50, no será suficiente para la reconstrucción de la señal. La figura 21 muestra los resultados para tal situación, es decir, tomando tan solo 50 proyecciones.

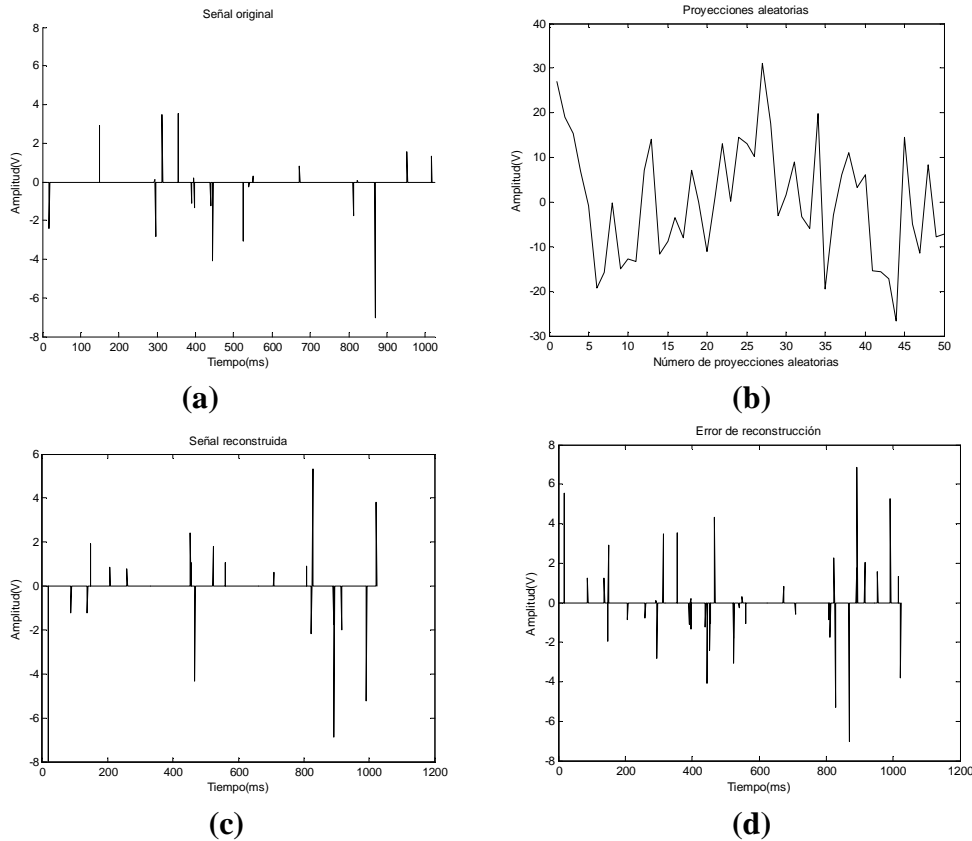


Fig21. Reconstrucción de una señal sparse en el dominio del tiempo empleando el algoritmo OMP con $k = 20$ y 50 proyecciones aleatorias. **(a)** Señal original. **(b)** Proyecciones aleatorias. **(c)** Señal reconstruida. **(d)** Error de reconstrucción.

Comparando los resultados obtenidos utilizando OMP con los que se obtuvieron en el apartado anterior con el algoritmo MP, se puede apreciar fácilmente que el error medio cuadrático generado por el OMP es catorce órdenes de magnitud menor que el generado usando MP.

También se puede apreciar que el número de mediciones para obtener una reconstrucción exitosa de la señal original es mucho menor en el algoritmo OMP que en el MP, lo veremos más adelante con la obtención de las curvas de probabilidad de reconstrucción y la obtención del factor de submuestreo.

4.2 Curvas de probabilidad de reconstrucción

La teoría del Compressed Sensing no es determinística sino probabilística. Como vimos en el *Teorema 4* ([12]), la probabilidad de reconstruir una señal con éxito es proporcional al número de mediciones que se le hacen a la señal original, es decir, se puede reconstruir una señal con probabilidad mayor o igual a:

$$(55) \quad p = 1 - e^{-Cm}$$

si el número de mediciones cumple con:

$$(56) \quad m \geq Ck \log\left(\frac{n}{k}\right)$$

donde C es una constante característica del método de reconstrucción, conocida como factor de submuestreo, n es la longitud de la señal, m el número de mediciones y k el nivel de escasez.

En vista del comportamiento probabilístico de la teoría en estudio, se hace evidente la necesidad de encontrar en qué circunstancias la probabilidad de reconstrucción de una señal tiende al valor deseado. Con la finalidad de resolver el problema planteado se lleva a cabo la siguiente simulación, que consiste en la elaboración de las curvas de probabilidad de reconstrucción.

4.2.1 Metodología para la elaboración de las curvas

Debido a la incertidumbre que existe en la selección del número de medidas adecuado para una reconstrucción exitosa, se realizan curvas de error medio cuadrático y probabilidad exitosa en función del número de proyecciones aleatorias y el nivel de escasez de la señal. Las curvas se realizan generando señales sparse de forma aleatoria y reconstruyendo cada una de ellas para un rango de medidas y un único valor de escasez. Se cuantifica cuántas veces ha sido reconstruida con éxito la señal, entendiendo por reconstrucción exitosa cuando el error medio cuadrático es menor a 10^{-6} .

E_{mc} representa el error cuadrático medio, donde x es la señal original y x_{est} la señal reconstruida, siendo n el número de componentes de la señal x y P el número de señales aleatorias generadas para su posterior reconstrucción:

$$(57) \quad E_{mc} = \frac{\sum_{i=1}^P ||x - x'_{est}||^2}{n}$$

La probabilidad de reconstruir una señal con éxito fue calculada como el cociente del número de reconstrucciones exitosas y el número de repeticiones:

$$(58) \quad p = \frac{\text{número de reconstrucciones exitosas}}{\text{número de repeticiones}}$$

A continuación se detallan los pasos seguidos para la obtención de las curvas de probabilidad de reconstrucción:

Nota: Para cada método de reconstrucción se utilizó la misma señal y la misma matriz de medición, con la finalidad de tener igualdad de condiciones.

Paso I: Se fija un número mínimo y uno máximo de mediciones, en este caso se toman valores en un rango entre [2,300]. Para cada número de mediciones se ejecutó el algoritmo para 100 señales generadas de manera aleatoria.

Paso II: Se ejecuta el algoritmo y se cuantifica el número de reconstrucciones exitosas, entendiendo por éxito cuando el error de reconstrucción es menor de 10^{-6} .

Paso III: Se calcula la probabilidad como el cociente del número de reconstrucciones exitosas y el número de repeticiones.

Paso IV: Se incrementa el número de mediciones en dos y se repite el proceso anterior hasta llegar al máximo número de mediciones.

Paso V: Por último se grafica la probabilidad de reconstrucción en función del número de mediciones.

Este procedimiento se repite para cada nivel de escasez que queramos representar, en nuestro caso $k = 5, 10, 20$ y 40 .

4.2.2 Curvas de probabilidad utilizando MP

Las curvas se han realizado generando 100 señales sparse de longitud $n = 256$ de forma aleatoria y reconstruyendo cada una de ellas para un rango de medidas pertenecientes al intervalo $[2,300]$ y un único valor de escasez, se repiten las curvas para $k = 5, 10, 20$ y 40 . Se cuantifica cuántas veces ha sido reconstruida con éxito la señal, entendiendo por reconstrucción exitosa cuando el error medio cuadrático es menor a 10^{-6} .

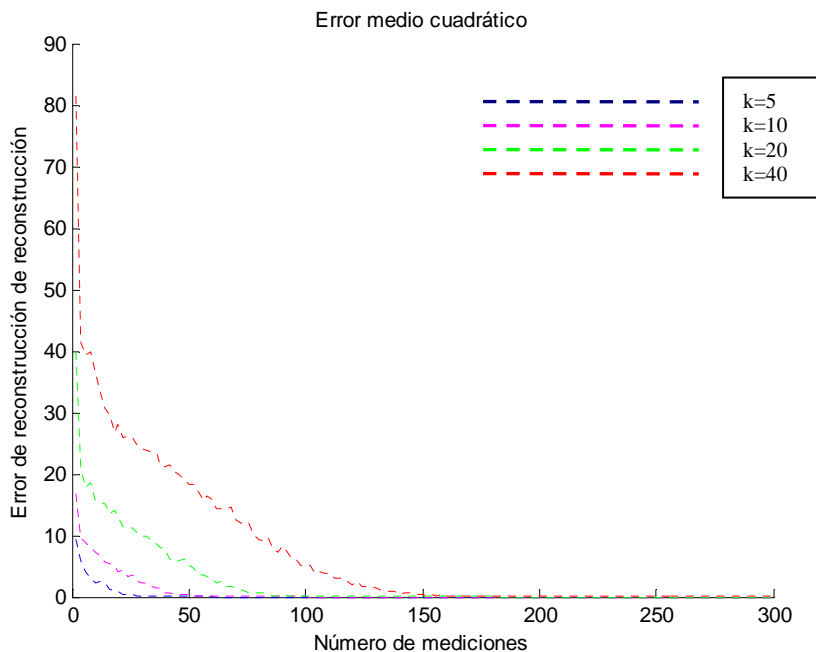


Fig22. Error cuadrático medio en función del número de mediciones para diferentes valores de escasez utilizando el algoritmo MP. Para la obtención de las curvas se reconstruyeron 100 señales aleatorias.

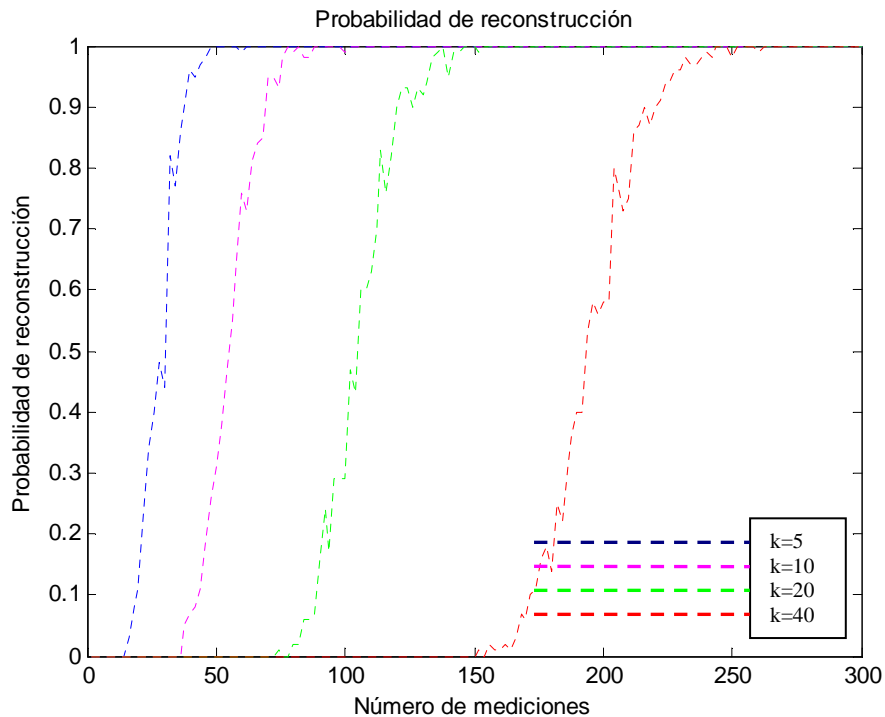


Fig23. Probabilidad de reconstrucción en función del número de mediciones para diferentes valores de escasez utilizando el algoritmo MP. Para la obtención de las curvas se reconstruyeron 100 señales aleatorias.

En las gráficas se observa que a medida que se incrementa el número de mediciones disminuye el error cuadrático medio y la probabilidad de reconstrucción tiende a 1.

La tabla siguiente refleja la cantidad de mediciones necesarias para conseguir una reconstrucción aproximadamente perfecta.

Escasez (k)	5	10	20	40
Proyecciones (Y)	75	115	185	290

Tabla 2. Proyecciones aleatorias en función de la escasez

Estos datos presentan un comportamiento aproximadamente lineal descrito por

$$(59) \quad Y = 6,06k + 52,6$$

Donde Y son las proyecciones aleatorias necesarias para una reconstrucción exitosa y k la escasez de la señal.

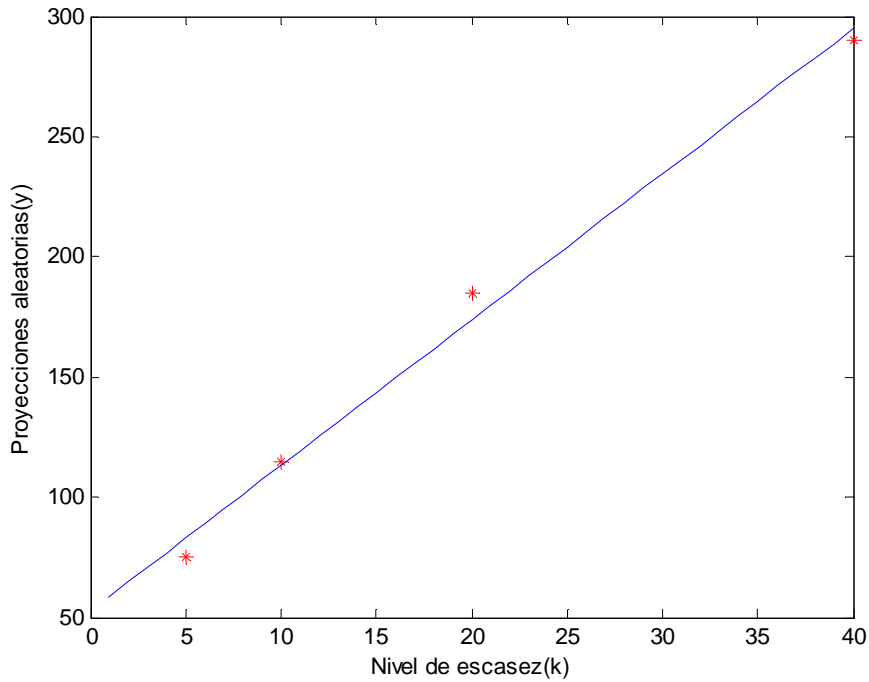


Fig24. Número de mediciones necesarias para obtener una probabilidad de reconstrucción igual a 1 en función del nivel de escasez k utilizando el algoritmo MP y recta de regresión $Y = 6,06k + 52,6$.

A continuación se obtienen las curvas de error y probabilidad empleando el método anterior pero en este caso se reconstruirán 500 señales:

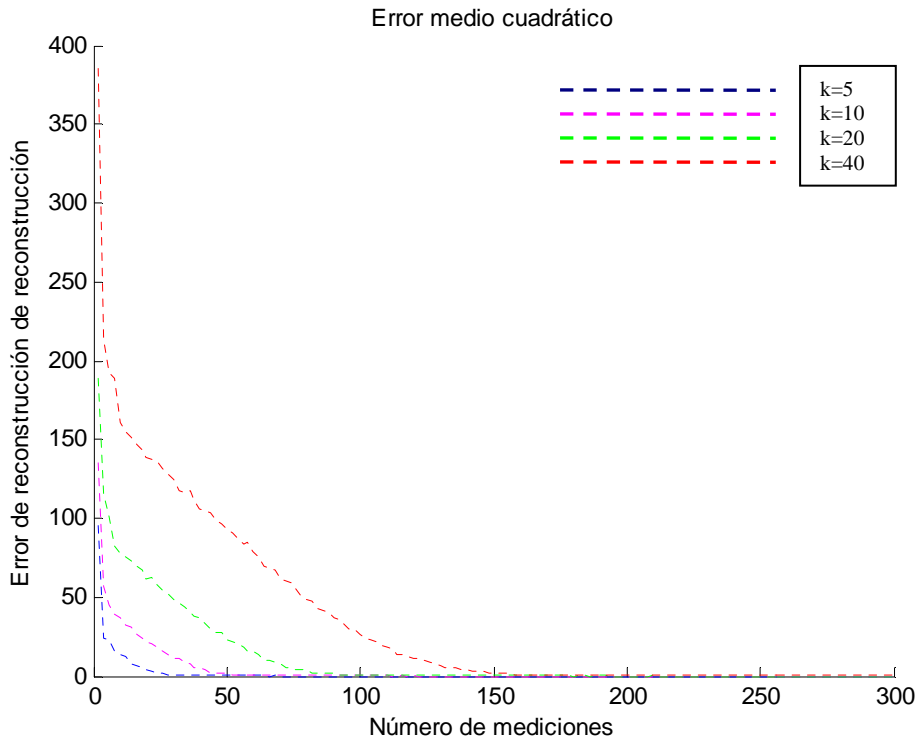


Fig25. Error cuadrático medio en función del número de mediciones para diferentes valores de escasez utilizando el algoritmo MP. Para la obtención de las curvas se reconstruyeron 500 señales aleatorias.

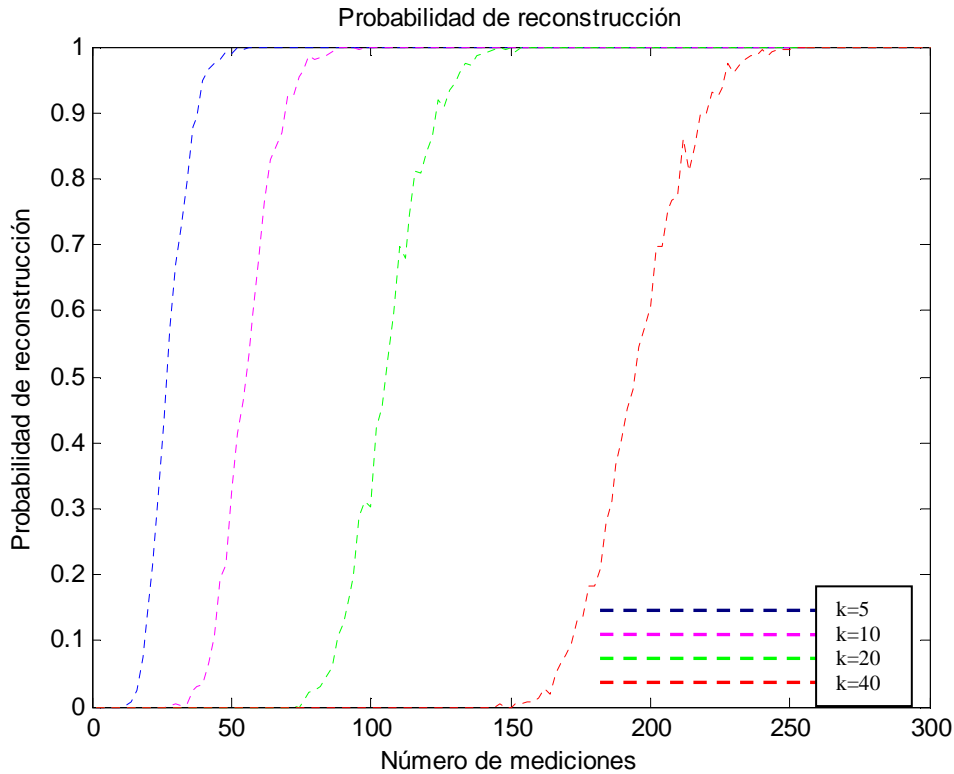


Fig26. Probabilidad de reconstrucción en función del número de mediciones para diferentes valores de escasez utilizando el algoritmo MP. Para la obtención de las curvas se reconstruyeron 500 señales aleatorias.

Incrementando el número de señales que se reconstruyen para la creación de las gráficas de error cuadrático medio y probabilidad de reconstrucción, comprobamos que los pequeños saltos que aparecen en las curvas obtenidas para 100 señales, desaparecen al aumentar el número de señales reconstruidas.

4.2.3 Curvas de probabilidad utilizando OMP

Al igual que en la reconstrucción realizada con el algoritmo MP, las curvas se han realizado generando 100 señales sparse de longitud $n = 256$ de forma aleatoria y reconstruyendo cada una de ellas para un rango de medidas pertenecientes al intervalo $[2,300]$ y un único valor de escasez, se repiten las curvas para $k = 5, 10, 20$ y 40 . Se cuantifica cuántas veces ha sido reconstruida con éxito la señal, entendiendo por reconstrucción exitosa cuando el error medio cuadrático es menor a 10^{-6} .

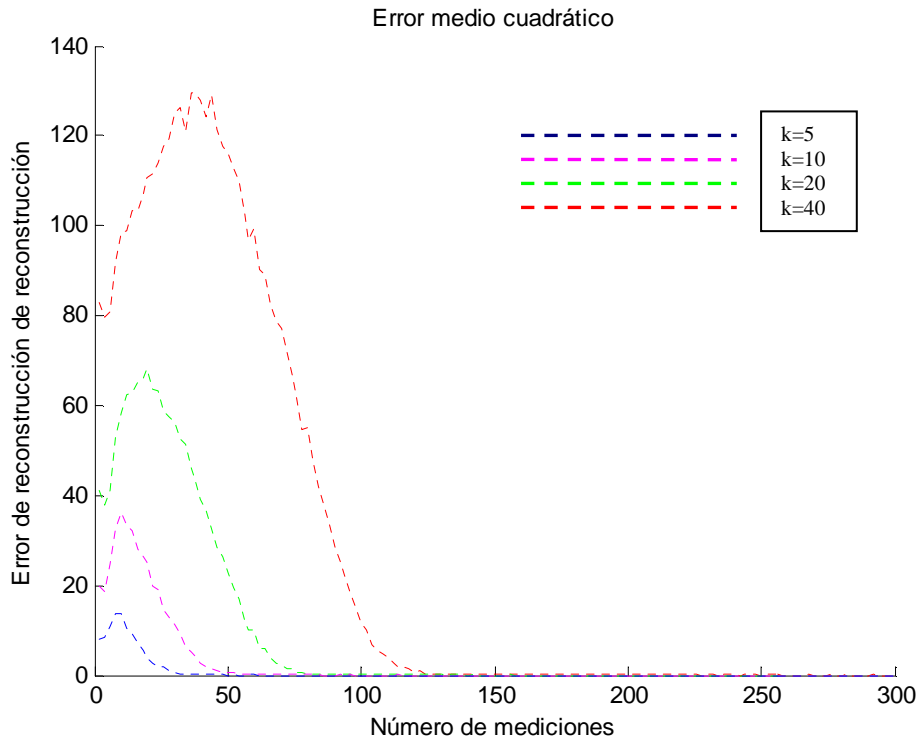


Fig27. Error cuadrático medio en función del número de mediciones para diferentes valores de escasez usando el algoritmo OMP. Para la obtención de las curvas se reconstruyeron 100 señales aleatorias.

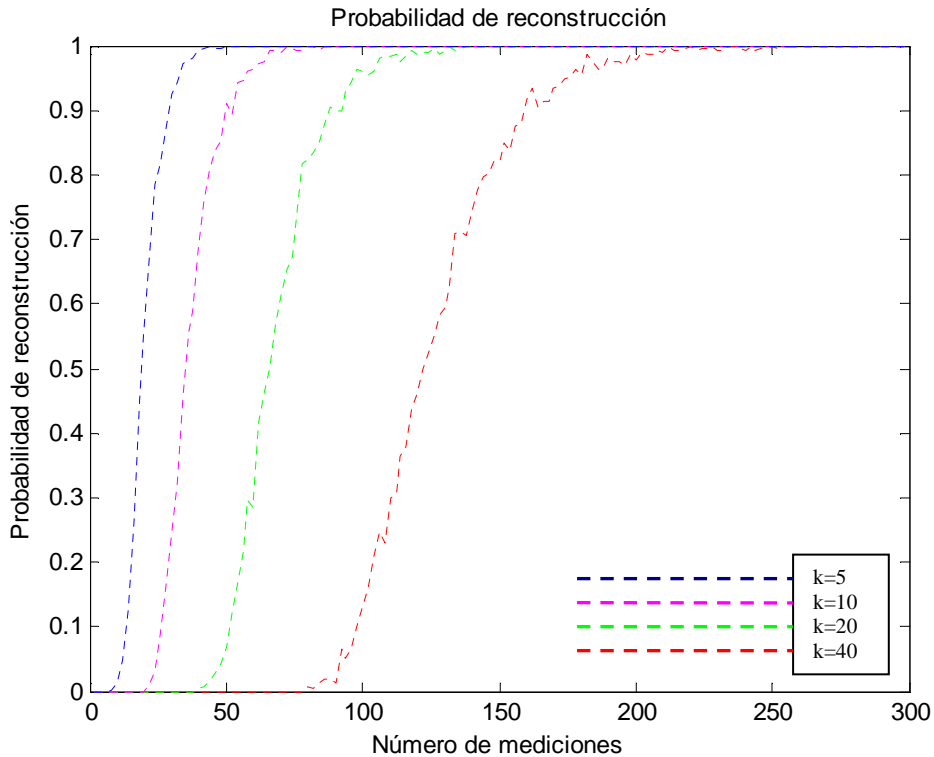


Fig28. Probabilidad de reconstrucción en función del número de mediciones para diferentes valores de escasez usando el algoritmo OMP. Para la obtención de las curvas se reconstruyeron 100 señales aleatorias.

En las gráficas se observa que a medida que se incrementa el número de mediciones disminuye el error cuadrático medio y la probabilidad de reconstrucción tiende a 1.

La tabla siguiente refleja la cantidad de mediciones necesarias para conseguir una reconstrucción aproximadamente perfecta.

Escasez (k)	5	10	20	40
Proyecciones (Y)	47	70	120	210

Tabla 1. Proyecciones aleatorias en función de la escasez

Estos datos presentan un comportamiento aproximadamente lineal descrito por

$$(60) \quad Y = 4,666k + 24,261$$

Donde Y son las proyecciones aleatorias necesarias para una reconstrucción exitosa y k la escasez de la señal.

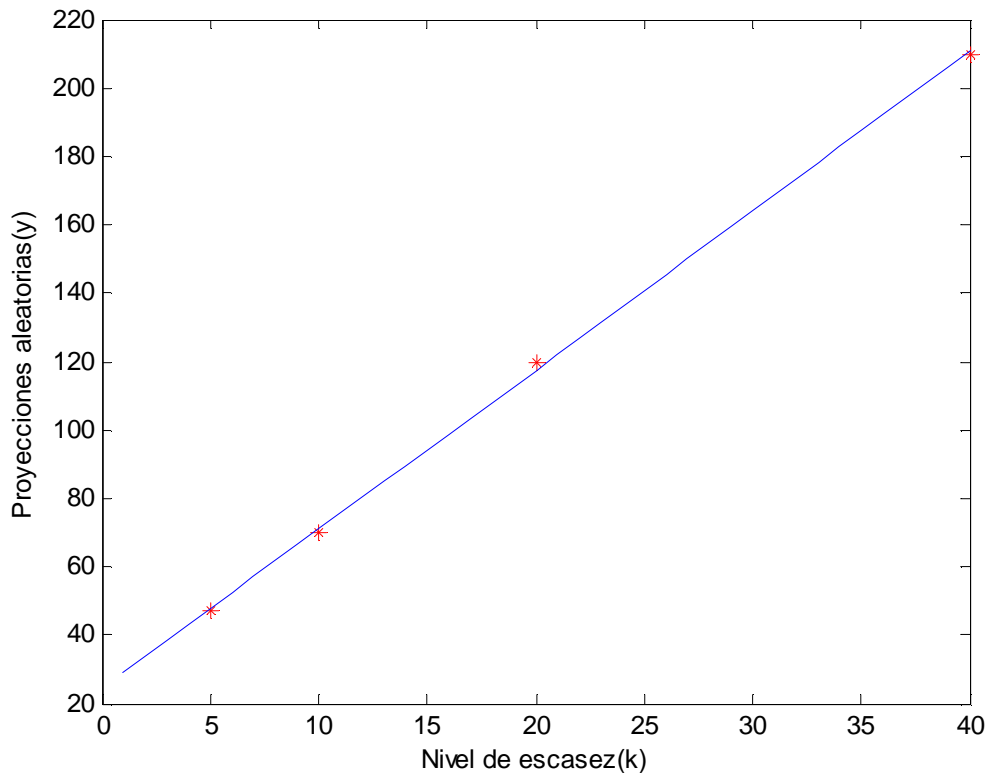


Fig29. Número de mediciones necesarias para obtener una probabilidad de reconstrucción igual a 1 en función del nivel de escasez k utilizando el algoritmo OMP y recta de regresión $Y = 2,605k + 36,65$.

A continuación se obtienen las curvas de error y probabilidad empleando el método anterior pero en este caso se reconstruirán 500 señales:

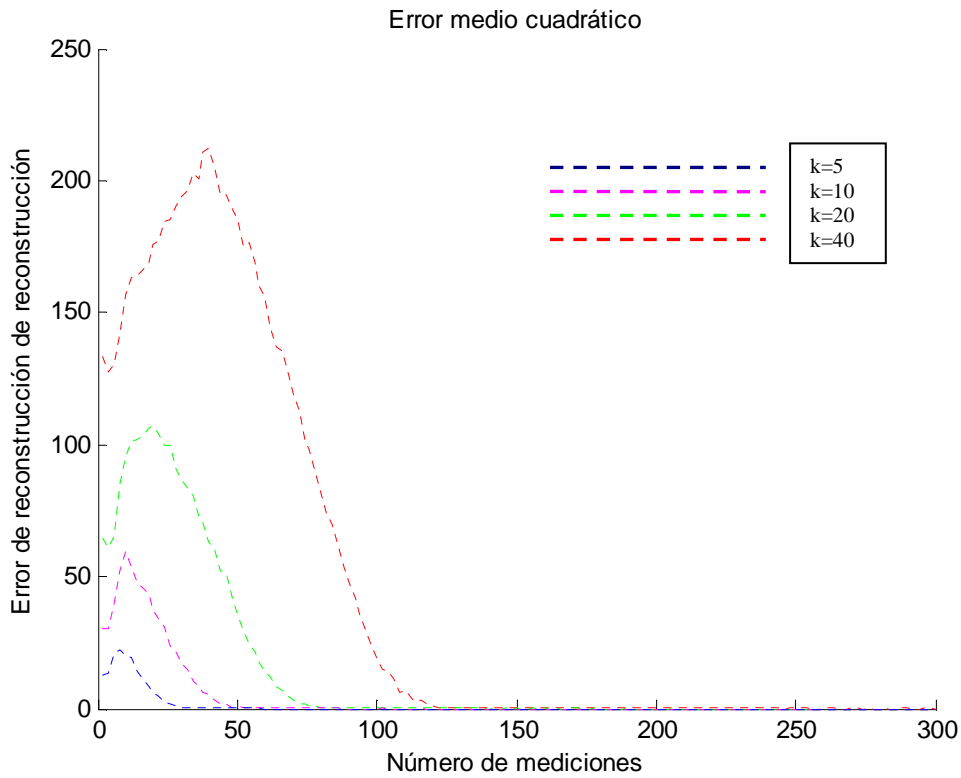


Fig30. Error cuadrático medio en función del número de mediciones para diferentes valores de escasez usando el algoritmo OMP. Para la obtención de las curvas se reconstruyeron 500 señales aleatorias.

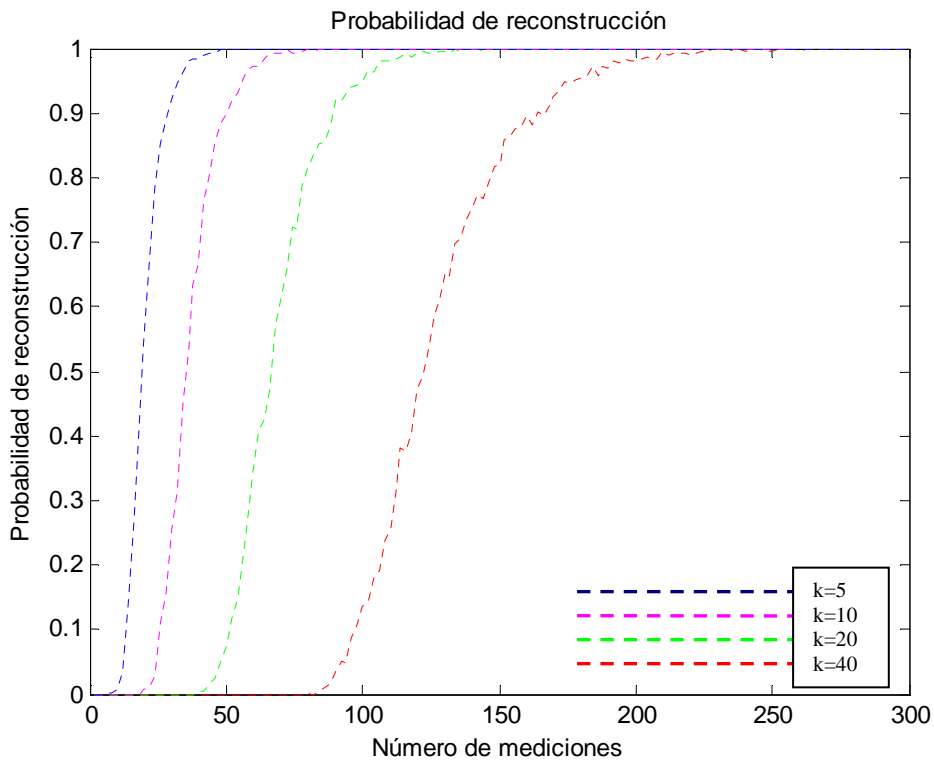


Fig31. Probabilidad de reconstrucción en función del número de mediciones para diferentes valores de escasez usando el algoritmo OMP. Para la obtención de las curvas se reconstruyeron 500 señales aleatorias.

Incrementando el número de señales que se reconstruyen para la creación de las gráficas de error cuadrático medio y probabilidad de reconstrucción, comprobamos que los pequeños saltos que aparecen en las curvas obtenidas para 100 señales desaparecen al realizar la reconstrucción de 500 señales.

4.3 Cálculo de la constante de submuestreo

Se puede garantizar que una señal será reconstruida con éxito con una probabilidad de por lo menos

$$(61) \quad p = 1 - e^{-c_2 k}$$

siempre y cuando se cumpla que

$$(62) \quad m \geq c_1 k \log\left(\frac{n}{k}\right)$$

Haciendo $c_2 \rightarrow \infty$, se obtiene que $p \rightarrow 1$, y de las figuras 20 y 25 se puede obtener empíricamente cuántas mediciones son necesarias para que cada algoritmo de reconstrucción tenga probabilidad exactamente 1. Con el número de mediciones conocido pasamos a calcular la constante c_1 característica de cada método de reconstrucción, conocido como factor de submuestreo.

Escasez (k)	5	10	20	40
Proyecciones MP	75	115	185	290
Proyecciones OMP	47	70	120	210

Tabla 2. Número de mediciones necesarias para cada algoritmo para alcanzar una probabilidad de reconstrucción igual a 1.

Entonces, la constante característica para cada método de reconstrucción vendría dada por:

$$(63) \quad c_1 \geq \frac{m}{k \log\left(\frac{n}{k}\right)}$$

Escasez (k)	5	10	20	40		
Factor de submuestreo MP	8,77	8,16	8,35	8,99	Promedio	8,56
Factor de submuestreo OMP	5,49	4,97	5,41	6,51	Promedio	5,59

Tabla 3. Comparación del valor de la constante de submuestreo para ambos algoritmos.

4.4 Reconstrucción de señales con ruido

Es bien conocido que la mayoría de los procesos, tanto de generación como de transmisión de señales, introducen cierto ruido en las señales. En este capítulo, lo que se hará será introducir ruido blanco Gaussiano aditivo a la señal original, a modo de simular condiciones reales de muestreo y transmisión de datos, ya que en la práctica nunca se obtiene una señal sin ruido.

4.4.1 Efecto del ruido blanco gaussiano al algoritmo MP

Se introduce a la señal original ruido con una potencia tal que genera una relación señal a ruido de 20dB. La relación señal a ruido está definida como:

$$(64) \quad SNR = 10 \log \left(\frac{P_s}{P_r} \right)$$

donde P_s es la potencia promedio de la señal y P_r la potencia promedio del ruido.

En la figura 32, se muestra el error que genera el algoritmo Matching Pursuit al reconstruir una señal que fue contaminada con ruido, la relación señal a ruido es de 20dB.

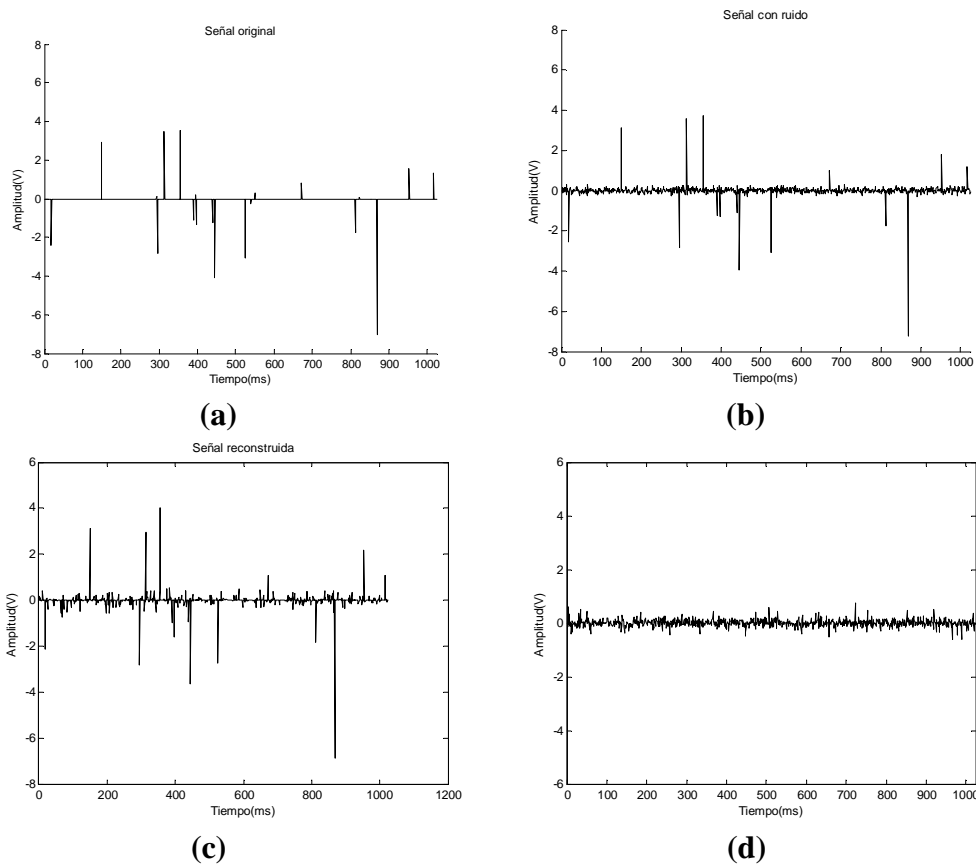


Fig32. Reconstrucción de una señal sparse en el dominio del tiempo con ruido blanco gaussiano aditivo con $SNR = 20\text{dB}$ utilizando el algoritmo MP. (a) Señal original. (b) Señal contaminada, $SNR = 20\text{dB}$. (c) Señal reconstruida a partir de MP. (d) Error de reconstrucción.

Se esperaba que la diferencia entre la señal original y la reconstruida fuese menor, esto es debido a que como la reconstrucción se hace utilizando los átomos del diccionario, tenemos que en el diccionario no hay átomos que puedan representar el ruido introducido en la señal.

Con el fin de analizar el efecto del ruido en la reconstrucción en cuanto al error cometido, se recuperó la señal anterior con un ruido blanco gaussiano aditivo de diferentes valores de potencia.

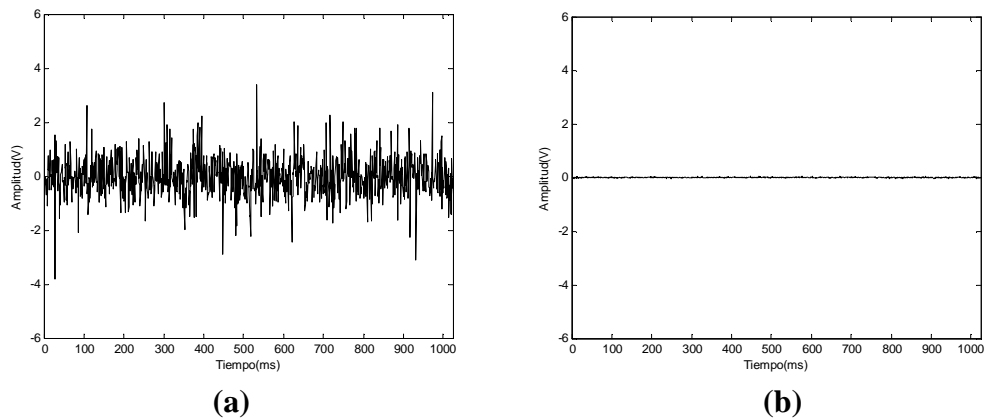


Fig33. Error de reconstrucción de una señal sparse en el tiempo con ruido blanco gaussiano aditivo de diferentes niveles de potencia, siendo el valor de la SNR(dB): **(a)** 5dB **(b)** 40dB

Se pudo observar que a medida que la potencia de ruido disminuye (es decir, aumenta la relación señal a ruido) también lo hace el error de reconstrucción, sin embargo esto es un resultado muy esperado que no introduce ninguna novedad. En la figura 34 se muestra la dependencia del error de la reconstrucción a medida que la relación señal a ruido aumenta.

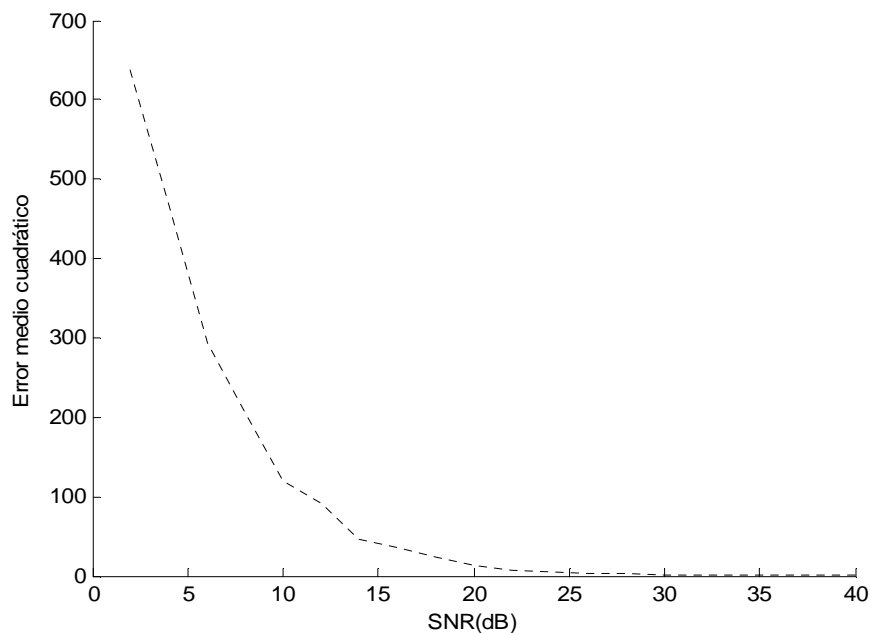


Fig34. Relación existente entre el error de reconstrucción y la SNR para el algoritmo MP.

4.4.2 Efecto del ruido blanco gaussiano al algoritmo OMP.

Se introduce a la señal original ruido con una potencia tal que genera una relación señal a ruido de 20dB. La relación señal a ruido está definida como en la ecuación (80).

En la figura 35, se muestra el error que genera el algoritmo Orthogonal Matching Pursuit al reconstruir una señal que fue contaminada con ruido, la relación señal a ruido es de 20dB.

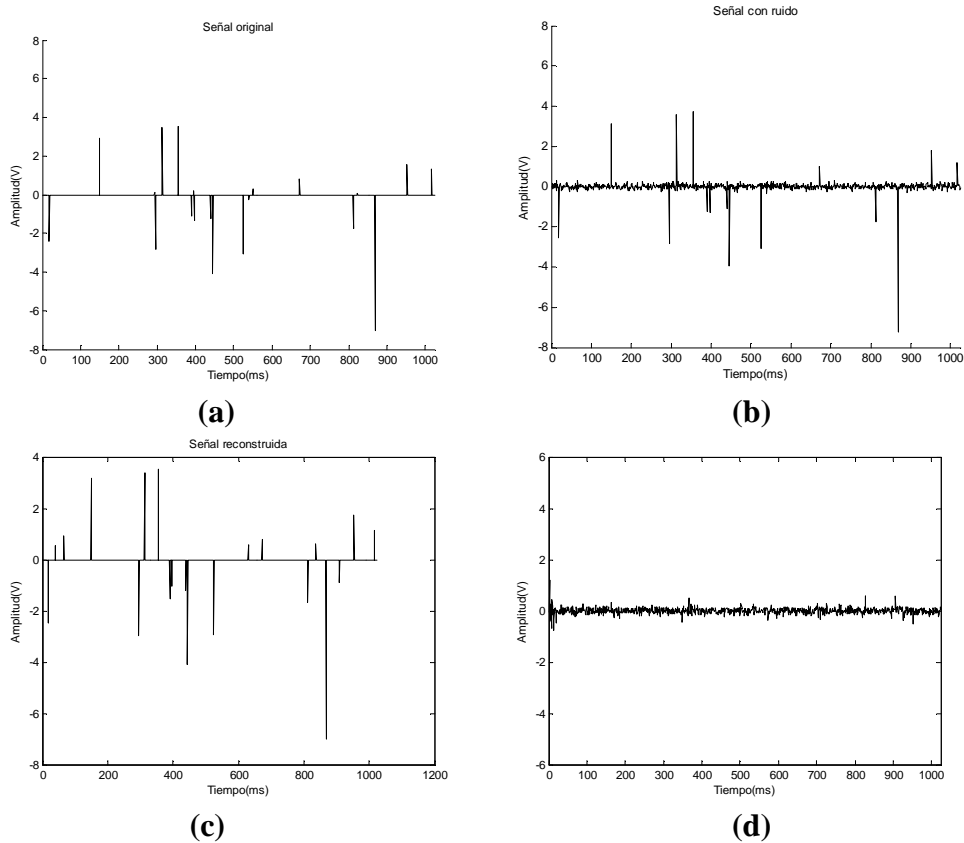


Fig35. Reconstrucción de una señal sparse en el dominio del tiempo con ruido blanco gaussiano aditivo con SNR = 20dB utilizando el algoritmo =MP. (a) Señal original.(b) Señal contaminada, SNR = 20dB. (c) Señal reconstruida a partir de OMP. (d) Error de reconstrucción.

Al igual que en el apartado anterior, al realizar la reconstrucción de la señal original contaminada con un ruido de SNR = 20dB, se esperaba que la diferencia entre la señal original y la reconstruida fuese menor, es decir, que la señal reconstruida fuese más limpia.

Con el fin de analizar el efecto del ruido en la reconstrucción en cuanto al error cometido, se recuperó la señal anterior con un ruido blanco gaussiano aditivo de diferentes valores de potencia.

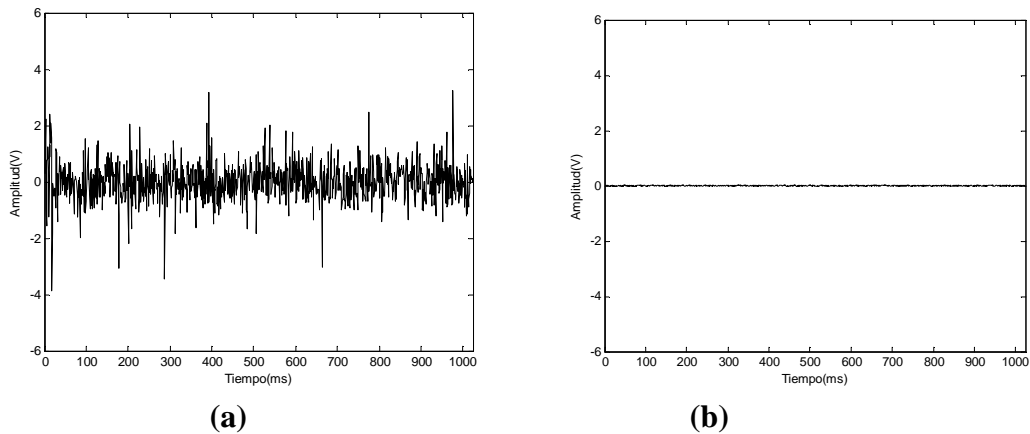


Fig36. Error de reconstrucción de una señal sparse en el tiempo con ruido blanco gaussiano aditivo de diferentes niveles de potencia, siendo el valor de la SNR(dB): **(a)** 5dB **(b)** 40dB

En la figura 34 se puede observar que a medida que la potencia de ruido disminuye (es decir, aumenta la relación señal a ruido) también lo hace el error de reconstrucción, sin embargo esto es un resultado muy esperado que no introduce ninguna novedad.

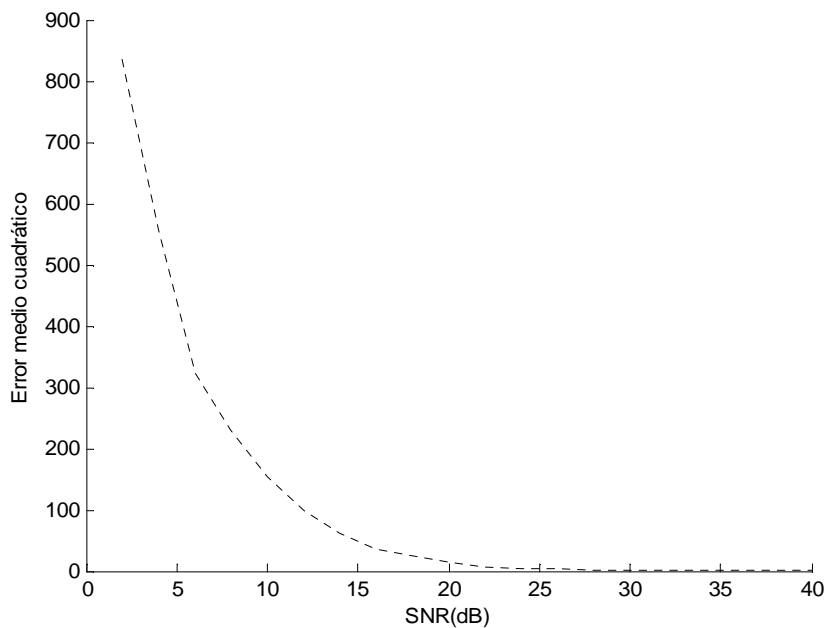


Fig37. Relación existente entre el error de reconstrucción y la SNR para el algoritmo

4.5 Conclusiones

En primer lugar se ha realizado la reconstrucción de una señal sparse empleando el algoritmo Matching Pursuit observando que con una proyección aleatoria de 300 medidas la reconstrucción se puede considerar exitosa con una diferencia entre la señal original y la reconstruida en el orden de 10^{-8} . Sin embargo, si el número de

proyecciones se reduce, por ejemplo en tres veces, posiblemente el número de mediciones no sea suficiente para la reconstrucción de la señal.

Seguidamente se ha reconstruido la misma señal sparse empleando esta vez el algoritmo Orthogonal Matching Pursuit, comprobando que con una proyección aleatoria de 300 medidas la reconstrucción se puede considerar exitosa con una diferencia entre la señal original y la reconstruida en el orden de 10^{-15} . En este caso comprobamos que si el número de proyecciones se reduce en tres veces, como en el caso anterior, el error de reconstrucción obtenido es prácticamente igual al conseguido con un número de mediciones igual a 300, pero si reducimos este número de mediciones a 50, no será suficiente para la reconstrucción de la señal.

En la siguiente tabla se compara el error cuadrático medio obtenido en la reconstrucción de la señal sparse utilizando los dos algoritmos:

Algoritmo	Error
MP	$1,5242 \times 10^{-14}$
OMP	$1,1242 \times 10^{-28}$

Tabla 4. Error cuadrático medio obtenido en la reconstrucción de una señal sparse empleando los algoritmos MP y OMP.

A continuación, se han generado las curvas de error cuadrático medio y de probabilidad de reconstrucción, en las que se puede observar que a medida que se incrementa el número de mediciones disminuye el error cuadrático medio y la probabilidad de reconstrucción tiende a 1. También comprobamos que el número de mediciones necesarias para una reconstrucción exitosa es menor empleando el algoritmo Orthogonal Matching Pursuit, concluyendo que el algoritmo OMP converge a la solución de forma más rápida. En la siguiente tabla se muestra el número de proyecciones aleatorias necesarias para una reconstrucción exitosa para distintos valores de escasez (k).

Escasez (k)	5	10	20	40
Proyecciones MP	75	115	185	290
Proyecciones OMP	47	70	120	210

Tabla 5. Número de mediciones necesarias para cada algoritmo para alcanzar una probabilidad de reconstrucción igual a 1.

En último lugar se ha realizado la reconstrucción de una señal sparse con ruido blanco gaussiano aditivo, se esperaba que la diferencia entre la señal original y la reconstruida fuese menor, esto es debido a que como la reconstrucción se hace utilizando los átomos del diccionario, tenemos que en el diccionario no hay átomos que puedan representar el ruido introducido en la señal. Se puede observar que a medida que la potencia de ruido disminuye (es decir, aumenta la SNR) también lo hace el error de reconstrucción, sin embargo esto es un resultado muy esperado que no introduce ninguna novedad.