



Universidad Nacional de La Plata

Tesis de Grado en Geofísica

TRANSFORMADAS DE RADON MEDIANTE  
ALGORITMOS VORACES PARA EL  
PROCESAMIENTO DE DATOS SÍSMICOS

Marcelo Isidoro Szelagowski

Director: Juan I. Sabbione

Codirector: Mauricio D. Sacchi



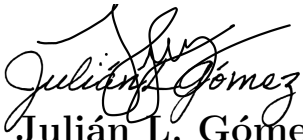
Facultad de Ciencias Astronómicas y Geofísicas

Año 2022

# Tribunal Examinador



**Dr. Juan I. Sabbione**



**Dr. Julián L. Gómez**



**Dr. Simón Lissa**

# Resumen

La transformada de Radon es una herramienta ampliamente utilizada en el procesamiento de señales sísmicas dado su vasto número de aplicaciones. En esencia, se basa en transformar el dominio de los datos, dependientes del tiempo y distancia a los receptores, en un dominio que depende, entre otras magnitudes, de las velocidades del medio por el que se propagan las ondas sísmicas. Existe una enorme diversidad de implementaciones y variaciones de la transformada de Radon. Dado que el subsuelo suele estructurarse en capas geológicas definidas, es común calcularla usando algoritmos que buscan imponer que el dato transformado sea ralo (*sparse*) en el dominio transformado. Existen dos grandes grupos de métodos para resolver este problema: los que se basan en minimizar una función de costo planteada con normas especialmente diseñadas para que la solución sea rala; y los que utilizan algoritmos voraces (*greedy*) que construyen el dominio transformado iterativamente. Esta tesis se basa en actualizar y estudiar códigos computacionales de este segundo grupo de métodos (*greedy*), que son los que han sido menos explorados. Para ello, se estudió su implementación en aplicaciones con datos sísmicos sintéticos y con datos reales de dominio público, buscando cuál de los analizados es el que mejor ajusta al dato de entrada y su eficiencia computacional.

# Agradecimientos

*A la FCAG y a todos los docentes que me formaron. A la Universidad Nacional de la Plata, y a nuestro país.*

*A mis directores Juani y a Mauricio. Cada mail, cada charla, cada instante con ellos fue sinónimo de aprendizaje. Gracias.*

*A mis viejos, a mi hermana, a mi abuela y a mi tía. A toda mi familia. Por los valores con los que me educaron y el amor de mi hogar.*

*A Angie, la encargada de sacarme una sonrisa todos los días. Gracias por tanta alegría.*

*A todos mis amigos, los que siempre están. Y particularmente a los de la facu, gracias por haber hecho de esta la mejor etapa de mi vida.*

*A mis compañeros de básquet, por tantas horas compartidas hace ya tantos años...*

*A Diego Armando, y a toda la gente del lobo.*

*A Dory.*

***Me hubiera sido imposible haberlo disfrutado más...***

# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Transformada de Radon</b>	<b>6</b>
2.1. TR lineal . . . . .	7
2.2. TR parabólica . . . . .	8
2.3. TR hiperbólica . . . . .	10
2.4. Observaciones . . . . .	13
2.5. TR por inversión . . . . .	16
<b>3. Métodos</b>	<b>18</b>
3.1. Mínimos cuadrados amortiguados (DLS) . . . . .	19
3.2. TR Hiperbólica con dominio restringido (RHRT) . . . . .	20
3.3. Métodos voraces ( <i>greedy</i> ) iterativos . . . . .	21
3.3.1. Transformada de Radon voraz (GRT) . . . . .	22
3.3.2. Búsqueda coincidente por etapas (StOMP) . . . . .	24
3.3.3. Potenciales ventajas de StOMP sobre GRT . . . . .	25
3.4. Apéndice gráfico . . . . .	25
<b>4. Aplicación sobre datos sintéticos</b>	<b>28</b>
4.1. Procedimiento, condiciones y criterios . . . . .	28
4.2. Pruebas . . . . .	30
4.2.1. Variación de relación <i>snr</i> . . . . .	30
4.2.2. Variación de la entrada . . . . .	38
4.2.3. Observaciones preliminares . . . . .	38
4.2.4. Pruebas de sensibilidad . . . . .	42
4.3. Conclusiones del capítulo . . . . .	45

<b>5. Aplicación a datos reales</b>	<b>46</b>
5.1. Procedimiento . . . . .	46
5.2. Aplicación . . . . .	47
<b>6. Conclusiones</b>	<b>56</b>
<b>A. Abreviaturas</b>	<b>58</b>
<b>B. Símbolos</b>	<b>59</b>

# Índice de figuras

2.1. Trayectorias lineales analizadas para TR lineal ( <i>izq.</i> ). Coordenadas que mapean cada evento lineal en el dominio transformado de Radon ( <i>der.</i> ). . .	8
2.2. Trayectorias hiperbólicas analizadas para TR hiperbólica ( <i>izq.</i> ). Coordenadas que mapean cada evento hiperbólico en el dominio transformado de Radon ( <i>der.</i> ). . . . .	12
2.3. <b>Efectos de borde en TR lineal.</b> <i>Izq.</i> (a): CMP de entrada, contiene un solo evento lineal EF. <i>Der.</i> (b): Dominio transformado de Radon lineal del dato de entrada. . . . .	14
2.4. <i>Fila arriba:</i> Registro sintético al que se le aplica el operador de TR adjunto y luego directo ( <i>centro</i> ), y su residuo ( <i>derecha</i> ). <i>Fila abajo:</i> Registro sintético al cual se lo transformó al dominio de Radón, se invirtieron los coeficientes y se volvió al dominio $(t - x)$ ( <i>centro</i> ), y residuo ( <i>derecha</i> ). . .	15
3.1. Criterios de selección para los métodos. En azul se muestra la señal sintética de entrada. Superpuesta a ella, en naranja, están los coeficientes que seleccionaría cada método. En el caso de OMP, GRT, y StOMP, la imagen representa la selección de una sola iteración. En la figura correspondiente a StOMP, la línea punteada representa el criterio de selección igual a $ thr \cdot \sigma $ .	26
4.1. Registro sintético 1, generado a partir de la información correspondiente a las tablas 4.1 y 4.2 sin ruido adicionado. . . . .	32
4.2. Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando un dato de entrada sin ruido adicionado. . . . .	33
4.3. Registro sintético 2, generado a partir de la información correspondiente a las tablas 4.1 y 4.2 con nivel de ruido a $snr = 1$ . . . . .	34
4.4. Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando datos con relación señal-ruido de 1. . . . .	35

4.5.	Registro sintético 3, generado a partir de la información correspondiente a las tablas 4.1 y 4.2 con nivel de ruido $snr = 0.5$ . . . . .	36
4.6.	Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando el registro con relación señal-ruido de 0.5. . . . .	37
4.7.	Registro sintético 4, generado a partir de la información correspondiente a las tablas 4.1 y 4.6 con ruido correspondiente a $snr = 1$ . . . . .	39
4.8.	Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando el dato sintético 4 relación señal-ruido de 1. . . . .	40
4.9.	<b>Sensibilidad del error de ajuste normalizado.</b> A izquierda se muestra el efecto de variar el porcentaje de coeficientes utilizados en RHRT. El eje vertical muestra el $e$ normalizado. En los paneles central y derecho se observan las variaciones en colores para distintos valores de número de iteraciones ( $N_{it}$ ) y valor umbral ( $thr$ ) de los algoritmos iterativos GRT y StOMP. . . . .	43
4.10.	<b>Sensibilidad de la relación señal-ruido.</b> A izquierda se muestra el efecto de variar el porcentaje de coeficientes utilizados en RHRT. El eje vertical muestra el valor del parámetro $snr$ obtenido. En los paneles central y derecho se observan las variaciones en colores para distintos valores de número de iteraciones ( $N_{it}$ ) y valor umbral ( $thr$ ) de los algoritmos iterativos GRT y StOMP. . . . .	43
4.11.	<b>Sensibilidad de la cantidad de coeficientes elegidos.</b> En la imagen izquierda se ve el efecto sobre GRT y a la derecha StOMP. Se observan las variaciones en colores para distintos valores de $N_{it}$ y $thr$ de los algoritmos iterativos. . . . .	44
5.1.	CMP de entrada del Golfo de México para el análisis de datos reales. . . .	48
5.2.	Coordenadas seleccionadas por StOMP en cada iteración. <i>Fila de arriba:</i> Caso de StOMP 1. <i>Fila de abajo:</i> Caso de StOMP 2, cada 2 iteraciones. . .	50
5.3.	Coordenadas seleccionadas por StOMP y RHRT en el dominio de Radon. En celeste se muestran las regiones consideradas para atenuar múltiples. Entre líneas punteadas rojas los tiempos en que se analizará el dominio ( $t - x$ ) en detalle . . . . .	51
5.4.	Dato de entrada con NMO aplicado. Foco entre 3.4 y 6.5 segundos. . . . .	52



5.5. Para la ventana de tiempo entre 3.4 y 6.5 segundos, se representa la atenuación de múltiples mediante StOMP 1 (*arriba izq.*), RHRT 1 (*arriba der.*), StOMP 2 (*abajo izq.*) y RHRT 2 (*abajo der.*). . . . . 53

5.6. Análisis de semblanza como control de calidad. Dato de entrada (*izq.*), dato con múltiples atenuadas mediante StOMP (*centro*) y mediante RHRT (*der.*) 54

# Índice de tablas

4.1. Parámetros utilizados para los registros sintéticos de entrada. . . . .	29
4.2. Parámetros de las tres hipérbolas de reflexión para el registro sintético. . .	31
4.3. Relación señal-ruido, error de ajuste ( <i>misfit</i> ), cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 1. . . . .	31
4.4. Relación señal-ruido, error de ajuste ( <i>misfit</i> ) normalizado, cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 2. . . . .	32
4.5. Relación señal-ruido, error de ajuste ( <i>misfit</i> ) normalizado, cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 3. . . . .	34
4.6. Parámetros de las cinco hipérbolas de reflexión para el registro sintético 4.	38
4.7. Relación señal-ruido, error de ajuste ( <i>misfit</i> ) normalizado, cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 4. . . . .	38
4.8. Parámetros de entrada utilizados en cada método para todos los ejemplos.	41
5.1. Parámetros de entrada, cantidad de coeficientes utilizados, y tiempo de cómputo para las pruebas StOMP 1 y StOMP 2 RHRT 1, StOMP 2, y RHRT 2. . . . .	49

# Capítulo 1

## Introducción

La Geofísica es la ciencia que estudia a la Tierra y sus fenómenos desde un punto de vista físico-matemático. Se puede separar en una serie de distintas disciplinas que se han ido formando dentro de ella (Udias, 1997). Estas son:

- Geodesia y Gravimetría
- Sismología
- Geomagnetismo y Aeronomía
- Oceanografía física
- Meteorología
- Geodinámica
- Geoquímica y Volcanología
- Planetología
- Prospección Geofísica

En el marco del último ítem, la **prospección geofísica** es la encargada del estudio del subsuelo terrestre mediante métodos físicos, los cuales se dividen principalmete en los **métodos potenciales** (gravimétrico y magnético), el **método eléctrico**, y el **método sísmico**. Este trabajo de Tesis se enmarca completamente en el **método sísmico**.

El observable del método son las series de tiempo (trazas) de ondas generadas por una fuente artificial y registrados en superficie, y tiene como uno de sus objetivos conocer la distribución y estructuras geológicas del subsuelo. Una de las ventajas que presenta es que permite obtener información a grandes profundidades (algunos kilómetros de profundidad)

---

con una resolución aceptable de unas pocas decenas de metros. Los parámetros que busca determinar son: el espesor de las capas y la velocidad con la que se propagan las ondas en cada una de ellas. La principal aplicación se da en exploración de hidrocarburos.

En líneas generales, podríamos dividir al método en etapas. En la primera etapa, es necesario **adquirir** la información. Por lo tanto, se deben diseñar los tendidos sísmicos y sus respectivos parámetros, prestando mayor atención a la profundidad de interés a explorar. Luego, el **procesamiento** ordena dicha información, busca mitigar lo más posible el ruido (ya sea coherente o aleatorio) que ésta pueda tener, realiza correcciones de ganancia, aplica correcciones de *normal moveout* (NMO), lleva a cabo las sumas correspondientes entre las trazas sísmicas obtenidas (*stack*), entre otros procedimientos. Lo anteriormente mencionado tiene una gran relevancia, ya que si el procesamiento es apropiado, el dato procesado resulta representativo del subsuelo. Posteriormente, el dato debe ser **migrado e invertido**. Entre otras cosas, la migración busca colapsar la energía difractada y, además, estimar la profundidad de las capas. La inversión del dato busca estimar propiedades físicas del subsuelo, tales como velocidad y densidad. Algunos ejemplos de inversión son: *amplitude variation with offset* (AVO) y *full wave inversion* (FWI). Por último, en la etapa de **interpretación** se analiza y estudia al dato procesado para obtener conclusiones geológicas, geofísicas, y petrofísicas, entre otras.

Suponiendo un medio de capas planas y paralelas se puede demostrar, utilizando la ley de Snell, que los sucesivos arribos de los rayos reflejados en la interfaz entre dos capas pueden ser bien aproximados por una hipérbola, denominada **hipérbola de reflexión**. Para alcanzar este resultado se realiza un análisis geométrico y además se aproxima la trayectoria del rayo, considerando que se propaga por el subsuelo con una única velocidad media. Esta última, es proporcional a las velocidades características y espesores de las capas suprayacentes.

El mejor dato obtenido es aquel que solamente contiene información correspondiente a reflexiones de distintas interfaces. Sin embargo, la realidad no es así a causa de una enormidad de sucesos, como por ejemplo la presencia de *ground roll*, la existencia de puntos difractores, o el arribo de ondas múltiples. El último mencionado se explica a continuación. Sabemos que la energía de una onda sísmica es proporcional al producto de los coeficientes de reflexión y refracción de las distintas interfaces. Entonces, en el caso que existan coeficientes de reflexión grandes, se registran hipérbolas generadas por las ondas que reflejan una sola vez en la interfaz, denominadas **primarias**, y aquellas hipérbolas asociadas a rayos reflejados en dos o más oportunidades en una misma interfaz, llama-

---

das **múltiples**. Dichas múltiples enmascaran la señal y forman parte del ruido coherente del registro, pudiendo llevar a una errónea interpretación del subsuelo (Velis, 2022). Un escenario donde se observa comunmente esta situación es en sísmica marina. Allí, los coeficientes de reflexión correspondientes a las interfases aire-agua y agua-fondo marino, presentan amplitudes considerablemente más grandes con respecto a los coeficientes asociados a cambios en la sedimentación. Esto se asocia, principalmente, al fuerte contraste de densidad. Una característica muy importante de estas reflexiones es su baja velocidad asociada, debida a que las múltiples se propagan en las primeras capas del medio. Como consecuencia, en un registro sísmico se observan con mayor curvatura que los arribos primarios. Incluso, esta situación se enfatiza ya que arriban a tiempos profundos, donde la velocidad de las reflexiones primarias es considerablemente mayor, por lo que es evidente su presencia en registros de fuente común (*shot gather*) o de punto medio común (*common mid-point gather*, CMP), por ejemplo.

Esta problemática nos da el puntapié para investigar sobre distintas técnicas de procesamiento que nos permitan hacer diversos análisis de velocidad. Luego, a partir de dicho análisis, se pueden modelar las hipérbolas de reflexión, y así eliminar las múltiples. Ese modelado lo haremos mediante la **transformada de Radon** (TR) hiperbólica, que será explicada en el Capítulo 2. Dicha transformada representa una de las herramientas más flexibles para realizar análisis de velocidades (Thorson y Claerbout, 1985). Entre las diversas aplicaciones que la TR tiene, además de la eliminación de múltiples marinas (Foster y Mosher, 1992), se encuentran la reconstrucción de datos (Kabir y Verschuur, 1995), y la atenuación de ruido en datos microsísmicos (Sabbione et al., 2015), entre otras. Si bien históricamente la TR parabólica es más utilizada por los bajos tiempos computacionales, la TR hiperbólica permite trabajar con el dato crudo, sin la necesidad de corregirlo por NMO y las consecuencias que eso conlleva. Además, ajusta mejor las hipérbolas de reflexión a *offsets* lejanos. Es por ello, y en función de su gran aplicación en el procesamiento sísmico y su practicidad, que la TR sigue siendo intensamente estudiada en busca de nuevas aplicaciones y mejoras (Gholami y Farshad, 2019; Ma et al., 2019; Lin y Sacchi, 2020; Li y Sacchi, 2021). La búsqueda del modelo frecuentemente se hace a partir de minimizar una función de costo que depende del modelo transformado y que busca honrar los datos observados (Thorson y Claerbout, 1985). Es común proponer que el dominio transformado de Radon se representa de forma rala (*sparse*), concepto a desarrollar en el Capítulo 2. Surgen diversos enfoques para imponer en el dominio esta característica. Un camino posible es el de utilizar un término extra en la función de costo, para así

---

imponer la característica de escasez de elementos no nulos sobre el modelo. Se utilizan normas de tipo Cauchy (Sacchi y Ulrych, 1995) o L1 (Lu, 2013), entre otros. Por ejemplo, en caso de utilizar norma L1, se suele recurrir a algoritmos iterativos de optimización lineal, como *Fast iterative shrinkage-thresholding algorithm* (Beck y Teboulle, 2009). Este es un problema bien conocido, y tiene la ventaja de que ofrece condiciones de convergencia probadas. Sin embargo, una de las principales problemáticas que la TR hiperbólica enfrenta es su alto costo computacional para realizar inversiones, a causa de la enorme cantidad de coeficientes que requiere para reconstruir el dato. Entonces, surge otro enfoque de solución: algoritmos que buscan los mejores candidatos de soporte para modelar el espacio, seleccionando solamente aquellos que sean representativos. Dicho de otro modo, su objetivo será disminuir la cantidad de coordenadas y así lograr algoritmos eficaces. La principal ventaja de este tipo de algoritmos es que suelen ser muy rápidos y sencillos de aplicar. Se denominan voraces (*greedy*), y se los puede clasificar entre los que llevan a cabo la selección iterativamente y los que no. Los iterativos son conocidos desde hace tiempo por sus implementaciones clásicas del tipo *Orthogonal matching pursuit* (Tropp y Gilbert, 2007), aunque no brindan las mismas garantías de convergencia que los de optimización lineal basados en normas de tipo Cauchy o L1. Por otro lado, métodos voraces no iterativos se han utilizado en diversas aplicaciones, como por ejemplo *Restricted hyperbolic radon transform*, utilizado para la eliminación de múltiples (Sabbione y Sacchi, 2017) o la separación entre señal de interés y ruido indeseado (Sabbione y Sacchi, 2016).

Considerando lo anteriormente descrito, y en vista de que los métodos voraces aún tienen diversos aspectos por explorar, en este trabajo se propone hacer una evaluación exhaustiva, aplicada y profunda para comparar entre algunos de estos algoritmos. Particularmente, lo haremos entre la TR hiperbólica con dominio restringido (RHRT) (Sabbione y Sacchi, 2016), la TR voraz (*Greedy Radon transform*, GRT) (Wang et al., 2009) y el método de búsqueda coincidente por etapas (*Stagewise orthogonal matching pursuit*, StOMP) (Donoho et al., 2012). Como control de calidad, se contrastan los resultados obtenidos con un método de inversión cuya convergencia es asegurada, denominado mínimos cuadrados amortiguados (*Damped least squares*, DLS) (Levenberg, 1944). Este método utiliza todas las coordenadas para la reconstrucción, diferenciándose así de los métodos voraces (*greedy*).

Esta tesis se organiza según los siguientes capítulos: El Capítulo 1 introduce el trabajo. En el Capítulo 2 se describen los diferentes tipos de transformadas de Radon. Las explicaciones necesarias sobre cómo aplicar los algoritmos y, principalmente, los criterios de selección que cada uno de ellos emplea, se desarrollan a lo largo del Capítulo 3. En el

---

Capítulo 4, se invierten diferentes datos sintéticos y se analizan los resultados obtenidos, para así poder compararlos según diversos criterios. Además, veremos cómo se comportan ante variaciones en sus parámetros de entrada (robustez de los métodos). Por último, en el Capítulo 6, los algoritmos que presenten mejores resultados de reconstrucción de las hipérbolas de reflexión de datos sintéticos, serán aplicados en un dato sísmico del Golfo de México, con la intención de eliminar las múltiples características de la sísmica marina. El objetivo primordial de esta instancia es verificar la funcionalidad de los métodos voraces (*greedy*) en datos reales. Se pondrán en práctica diversos controles de calidad. Finalmente, se presentan las conclusiones del trabajo en el Capítulo 6.

# Capítulo 2

## Transformada de Radon

El objetivo de cualquier transformada es llevar los datos desde un dominio hacia otro que dependa de otras variables, procurando que el dominio transformado permita separar la energía de las señales de interés del resto del dato. El ejemplo más conocido es el de la transformada de Fourier, en el que un dato que se encuentra en el dominio del tiempo [s], en el dominio transformado dependerá de la frecuencia [Hz]. Para el caso de la transformada de Radon (TR), o simplemente “transformada Radon”, el objetivo será el mismo: transformar los datos a diferentes dominios. El dominio de Radon transformado depende principalmente de la velocidad de propagación de las reflexiones  $v$  y de su tiempo de ida y vuelta  $\tau$ .

En sismica, también es conocida como “transformada tau-p”, o “transformada slant-stack”. Entendemos a un registro sísmico como la superposición de diferentes eventos con la energía concentrada a lo largo de líneas (rectas, parábolas o hipérbolas) (Beylkin, 1987). Entonces, la TR transforma los elementos de un dominio al otro sumando a través de cierto tipo de curva (lo que da origen a la TR lineal, parabólica, hiperbólica) la energía que encuentre en su camino. Luego, asigna un único valor, el de la suma, a la coordenada que represente dicha curva en el dominio de Radon. En la siguiente sección, se verán ejemplos gráficos claros de algunas TR.

De forma general, podemos decir que la TR facilita la separación entre la señal y las distintas formas de ruido. Entre las tantas aplicaciones que tiene la TR, en el procesamiento sísmico se destacan la eliminación de múltiples (Hampson, 1986; Foster y Mosher, 1992), la interpolación de datos (Trad et al., 2002), la restauración de ciertas distancias fuente-receptor (*offset*) faltantes (Kabir y Verschuur, 1995) y la atenuación de ruido en datos microsísmicos (Sabbione et al., 2015).



## 2.1. TR lineal

Supongamos un evento sísmico con pendiente lineal en el dominio de los datos. En cuyo caso, el tiempo del evento está dado por:

$$t = \tau_0 + p_0 x , \quad (2.1)$$

siendo  $t$  el tiempo,  $x$  la variable,  $\tau_0$  la ordenada al origen y  $p_0$  la pendiente. Ésto es válido en una ventana en  $(t - x)$ , donde cualquier evento puede representarse con pendiente constante (lineal). Para el análisis de la TR lineal de un sismograma en el que se encuentra sumergido dicha recta, sumamos a lo largo de distintas rectas variando sus valores de  $\tau$  y  $p$ , obteniendo así la energía acumulada en cada una de ellas. De la manera más general posible, para un determinado par  $(\tau, p)$ , aplicar dicho análisis equivale a calcular la siguiente integral:

$$m_{adj}(\tau, p) = \int_{-\infty}^{\infty} d(\tau + px, x) dx , \quad (2.2)$$

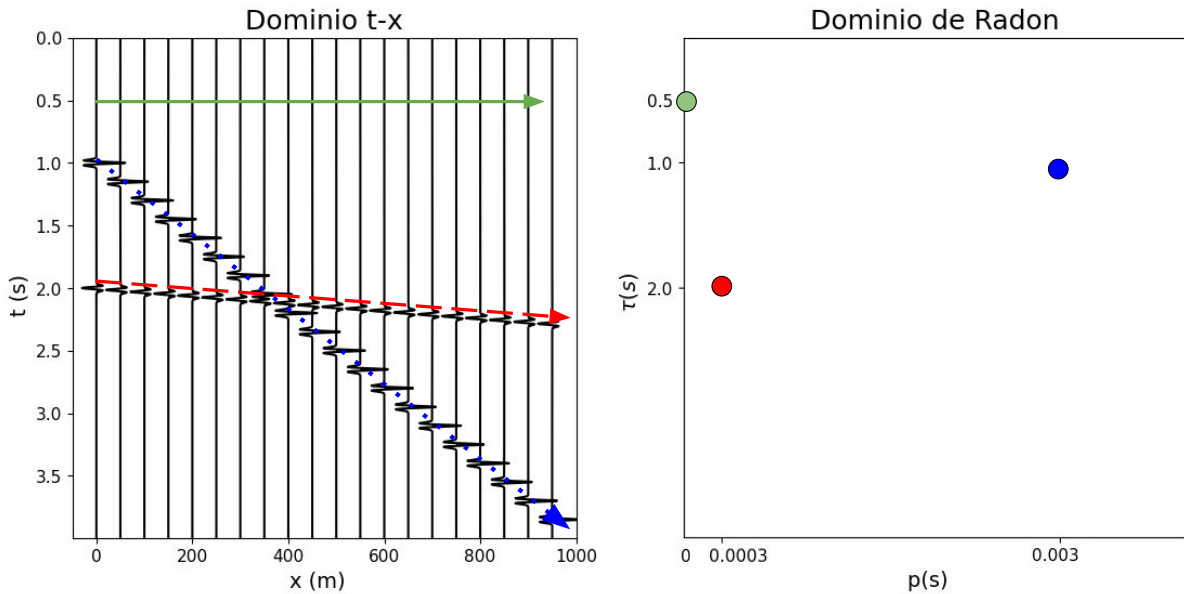
donde el subíndice *adj* indica que es el operador “adjunto”, que calcula los coeficientes de Radon a partir de los datos de amplitud  $d$ . En el caso del sismograma, teniendo en cuenta que el espaciamiento entre las trazas, y por ende la variable espacial  $x$ , es discreto y no es infinito, la ecuación anterior se puede escribir en forma discreta:

$$m_{adj}(\tau, p) = \sum_{x_{min}}^{x_{max}} d(\tau + px, x) . \quad (2.3)$$

Entonces, para  $(\tau, p) = (\tau_0, p_0)$ , esperaríamos encontrar un máximo ya que sumamos los valores sobre la recta traza a traza.

Un ejemplo sobre cómo se realiza la suma se muestra en la figura 2.1. Allí, se generó un sintético de 20 trazas con dos eventos rectos, uno con  $\tau = 1$  s,  $p = 0.003$  y amplitud = 1, el otro con  $\tau = 1.5$  s,  $p = 0.0003$  y amplitud =  $-0.5$ . Analizamos tres casos a modo de ejemplo:

- Recta continua verde:  $(\tau, p) = (0.5, 0)$  suma 0, ya que en ningún momento adiciona alguna parte de energía de algún evento, por lo que  $m_{adj}(0.5, 0) = 0$ .
- Recta punteada azul:  $(\tau, p) = (1, 0.003)$  suma 20, ya que por cada muestra a lo largo de las 20 trazas adiciona una unidad de amplitud, entonces  $m_{adj}(1, 0.003) = 20$ .



**Figura 2.1:** Trayectorias lineales analizadas para TR lineal (*izq.*). Coordenadas que mapean cada evento lineal en el dominio transformado de Radon (*der.*).

- Recta segmentada roja:  $(\tau, p) = (1.5, 0.0003)$  la suma es -10, teniendo en cuenta que por cada traza adicional  $-0.5$  unidades de amplitud, obteniendo:  $m_{adj}(1.5, 0.0003) = -10$ .

La transformada de Radon directa se define como aquella que genera los datos a partir del modelo. En su forma discreta, para el caso lineal, está dada por:

$$d(t, x) = \sum_{V_{min}}^{V_{max}} m(\tau = t - px, v) \quad (2.4)$$

donde destacamos que las ecuaciones 2.3 y 2.4 no representan un par transformado inverso, pues la TR no es ortogonal.

## 2.2. TR parabólica

Si bien esta TR no la usaremos a lo largo del trabajo, es interesante describirla brevemente ya que históricamente ha sido más utilizada que la TR hiperbólica para eliminar múltiples. Esto se debe a que la TR parabólica puede transformarse al dominio de las frecuencias y operar allí de forma más rápida. A lo largo de la sección veremos por qué podemos hacerlo. El procedimiento para eliminar dichas múltiples implica aproximar a las hipérbolas de reflexión como parábolas. Para poder imponer esa condición, es necesario realizar correcciones por NMO, para luego trabajar con la curvatura residual de las

múltiples. La desventaja que presenta, es que a la hora de aplicar el NMO ocurre estiramiento (*stretching*) en las trazas, por lo que se debe enmudecer (*mute*) parte del dato dañado. Por lo tanto, no se trabaja con el dato completo. Otro inconveniente es que la TR parabólica no se ajusta bien para *offsets* lejanos.

Así como la TR lineal suma a través de rectas, este operador suma a través de parábolas, intentando llevar una parábola en el dominio  $(t - x)$  a un punto en el dominio  $(\tau - q)$ . De todos modos, tengamos en cuenta que en el caso sísmico, aproximar una hipérbola de reflexión con una parábola introduce errores, por lo que esta transformada vale principalmente para distancias fuente-receptor cercanas. Otra manera de escapar a esto, es realizando las correcciones de NMO, en forma tal que se horizontalicen los arribos primarios y modelar con parábolas las múltiples, que tras la corrección tendrán menor curvatura (Hampson, 1986).

La TR parabólica está dada por:

$$m_{adj}(\tau, q) = \int_{-\infty}^{\infty} d(t = \tau + qx^2, x) dx , \quad (2.5)$$

siendo  $m$  el coeficiente de Radon,  $x$  la distancia fuente-receptor y  $q$  el parámetro de la parábola.

Tomando valores discretos  $q_j, j = 1, \dots, N_j$  y  $x_k, k = 1, \dots, N_k$ , el operador adjunto resulta:

$$m_{adj}(\tau, q_j) = \sum_{k=1}^{N_k} d(t = \tau + q_j x_k^2, x_k) . \quad (2.6)$$

Análogamente, la transformada de Radon parabólica directa para el caso discreto está dada por:

$$d(t, x_k) = \sum_{j=1}^{N_j} m(\tau = t - q_j x_k^2, q_j) \quad (2.7)$$

Tomando la transformada de Fourier respecto al tiempo a la ecuación 2.5, se obtiene (Hampson, 1986):

$$M_{adj}(\omega, q_j) = \sum_{k=1}^N D(\omega, x_k) e^{i\omega x_k^2 q_j^2} . \quad (2.8)$$

Esta expresión se deriva a partir de que para el caso parabólico  $t = \tau + q_j x_k^2$ , entonces  $\Delta t = q_j x_k^2$ . Luego, notemos que a la ecuación anterior la podemos escribir matricialmente como:

$$\mathbf{M}_{adj}(\omega) = \mathbf{L}^{\mathbf{H}}(\omega)\mathbf{D}(\omega) , \quad (2.9)$$

donde  $\mathbf{D}(\omega)$  es un vector que contiene los elementos  $d(\omega, k)$ ,  $\mathbf{M}_{adj}(\omega)$  es un vector formado por los coeficientes  $M_{adj}(\omega, q_j)$ , y el operador adjunto se expresa:

$$\mathbf{L}^{\mathbf{H}}(\omega) = \begin{pmatrix} e^{i\omega x_1^2 q_1^2} & e^{i\omega x_1^2 q_2^2} & \dots & e^{i\omega x_1^2 q_M^2} \\ e^{i\omega x_2^2 q_1^2} & e^{i\omega x_2^2 q_2^2} & \dots & e^{i\omega x_2^2 q_M^2} \\ \vdots & \vdots & \ddots & \vdots \\ e^{i\omega x_N^2 q_1^2} & e^{i\omega x_N^2 q_2^2} & \dots & e^{i\omega x_N^2 q_M^2} \end{pmatrix} . \quad (2.10)$$

La principal conclusión es que podemos obtener los coeficientes del dominio de Radon para la TR parabólica a partir de multiplicar el operador  $\mathbf{L}^{\mathbf{H}}(\omega)$  y la transformada de Fourier del dato.

Para la transformada parabólica de Radón directa, se demuestra que existe y que es (Hampson, 1986):

$$\mathbf{D}(\omega) = \mathbf{L}(\omega)\mathbf{M}(\omega) , \quad (2.11)$$

donde el operador directo está dado por:

$$\mathbf{L}(\omega) = \begin{pmatrix} e^{-i\omega x_1^2 q_1^2} & e^{-i\omega x_1^2 q_2^2} & \dots & e^{-i\omega x_1^2 q_M^2} \\ e^{-i\omega x_2^2 q_1^2} & e^{-i\omega x_2^2 q_2^2} & \dots & e^{-i\omega x_2^2 q_M^2} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i\omega x_N^2 q_1^2} & e^{-i\omega x_N^2 q_2^2} & \dots & e^{-i\omega x_N^2 q_M^2} \end{pmatrix} . \quad (2.12)$$

## 2.3. TR hiperbólica

La utilización de la TR hiperbólica en sismica consiste en aproximar a las reflexiones presentes en el dato como hipérbolas. Esta condición es considerablemente más real que aproximar a las reflexiones como parábolas, tal como hace la TR parabólica. Por lo tanto, logra sortear los inconvenientes que la última presenta, ya que permite trabajar con el dato crudo, sin NMO aplicado. Es por ello que resulta interesante estudiar la optimización de sus tiempos de cómputo.

De igual manera que se explicó para las TR anteriores, en este caso el operador adjunto llevará hipérbolas que se encuentren en el dominio  $(t-x)$  a puntos en el dominio de Radon sumando a través de ellas. Una forma de calcular los tiempos de arribo para una hipérbola de reflexión es mediante la siguiente ecuación:

$$t = \sqrt{t_0^2 + \frac{x^2}{v^2}} , \quad (2.13)$$

donde

$$t_0 = \frac{2z}{v},$$

siendo  $z$  la profundidad de la interfaz en la que refleja el rayo,  $v$  la velocidad media, y  $x$  la distancia entre fuente y receptor. Por lo tanto, se observa que los parámetros que definen a una hipérbola son la velocidad  $v$ , el tiempo doble  $t_0$  y la posición del ápice. No obstante, como trabajaremos con datos de un conjunto de trazas con punto medio común (*CMP gather*) el ápice estará centrado en cero. Por lo tanto, los parámetros en el dominio de Radon serán  $(\tau, v)$ , siendo  $\tau$  el tiempo de ida y vuelta del arribo y  $v$  la velocidad media del tiempo de viaje. De todos modos, si eventualmente se necesitara trabajar con datos sísmicos que incluyan hipérbolas desplazadas del origen, se recurre a la formulación de la TR hiperbólica de ápice desplazado (Hargreaves et al., 2003).

Para el cálculo de la TR hiperbólica, elegimos un par  $(\tau, v)$  para cada coeficiente y sumamos a lo largo de la variable espacial. Entonces, la TR hiperbólica está dada por:

$$m_{adj}(\tau, v) = \int_{-\infty}^{\infty} d \left( \sqrt{t = \tau^2 + \frac{x^2}{v^2}}, x \right) dx, \quad (2.14)$$

En forma discreta:

$$m_{adj}(\tau, v) = \sum_{x_{min}}^{x_{max}} d \left( \sqrt{t = \tau^2 + \frac{x^2}{v^2}}, x \right). \quad (2.15)$$

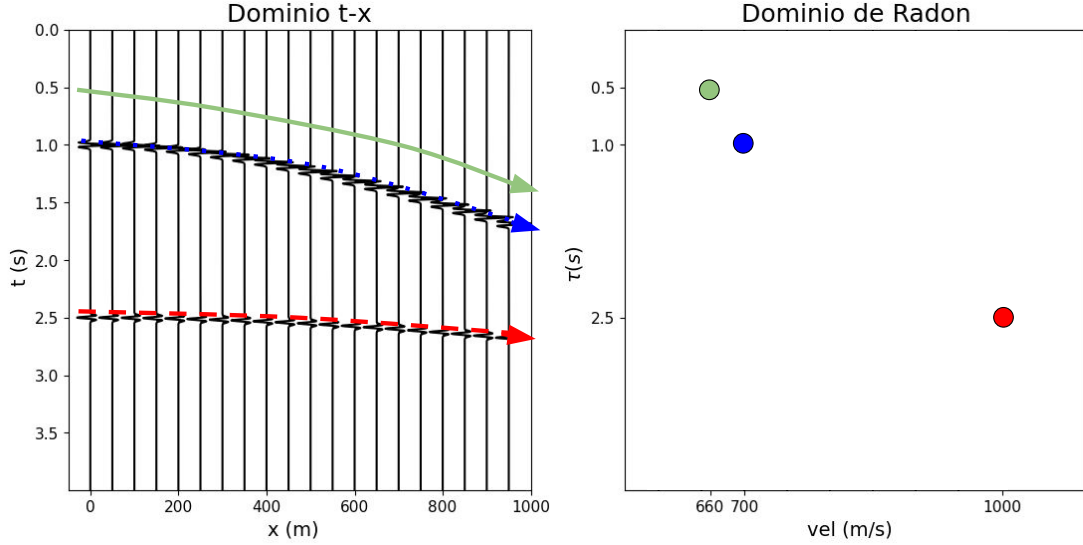
Veamos un ejemplo gráfico similar al presentado en la TR Lineal aplicado a la hiperbólica en la figura 2.2:

- Curva continua verde:  $(\tau, v) = (0.5, 660)$  suma 0, ya que en ningún momento adiciona alguna parte de energía de algún evento, por lo que  $m_{adj}(0.5, 660) = 0$ .
- Curva punteada azul:  $(\tau, v) = (1, 700)$  suma 20, ya que por cada muestra largo de las 20 trazas adicionará una unidad de amplitud, entonces  $m_{adj}(1, 700) = 20$ .
- Curva segmentada roja:  $(\tau, v) = (2.5, 1000)$  el resultado es  $-10$ , teniendo en cuenta que por cada traza suma  $-0.5$  unidades de amplitud, obteniendo:  $m_{adj}(2.5, 1000) = -10$ .

Por último, podemos definir la TR hiperbólica directa de la siguiente manera:

$$d(t, x) = \int_{-\infty}^{\infty} m \left( \tau = \sqrt{t^2 - \frac{x^2}{v^2}}, v \right) dv, \quad (2.16)$$

donde  $m(\tau, x)$  son los coeficientes de Radon que mejor ajustan al modelo y  $d(t, x)$  el dato sintetizado. Interpretativamente, podemos ver que el dato en el dominio  $(t - x)$  se irá



**Figura 2.2:** Trayectorias hiperbólicas analizadas para TR hiperbólica (*izq.*). Coordenadas que mapean cada evento hiperbólico en el dominio transformado de Radon (*der.*).

reconstruyendo considerando que para cada punto  $(t, x)$  habrá una serie de coeficientes  $m(\tau, v)$  que aporten su energía, por lo que sumados devuelven  $d(t, x)$ .

Discretizando las velocidades analizadas:

$$d(t, x) = \sum_{v_{min}}^{v_{max}} m \left( \tau = \sqrt{t^2 - \frac{x^2}{v^2}}, v \right) . \quad (2.17)$$

Una manera más concisa de definir los operadores adjunto (ec. 2.15) y directo (ec. 2.17), es mediante notación matricial. Debido a que  $\tau$  y  $v$  son discretos, a partir de vectorizar la matriz que contiene a todos los valores de  $m(\tau, v)$  se pueden representar los coeficientes de Radon con un vector  $\mathbf{m}$ . De igual modo, como  $t$  y  $x$  son discretos, los datos  $d(t, x)$  pueden escribirse como un vector  $\mathbf{d}$ :

$$\mathbf{m}_{adj} = \mathbf{L}^T \mathbf{d} , \quad (2.18)$$

$$\mathbf{d} = \mathbf{L} \mathbf{m} , \quad (2.19)$$

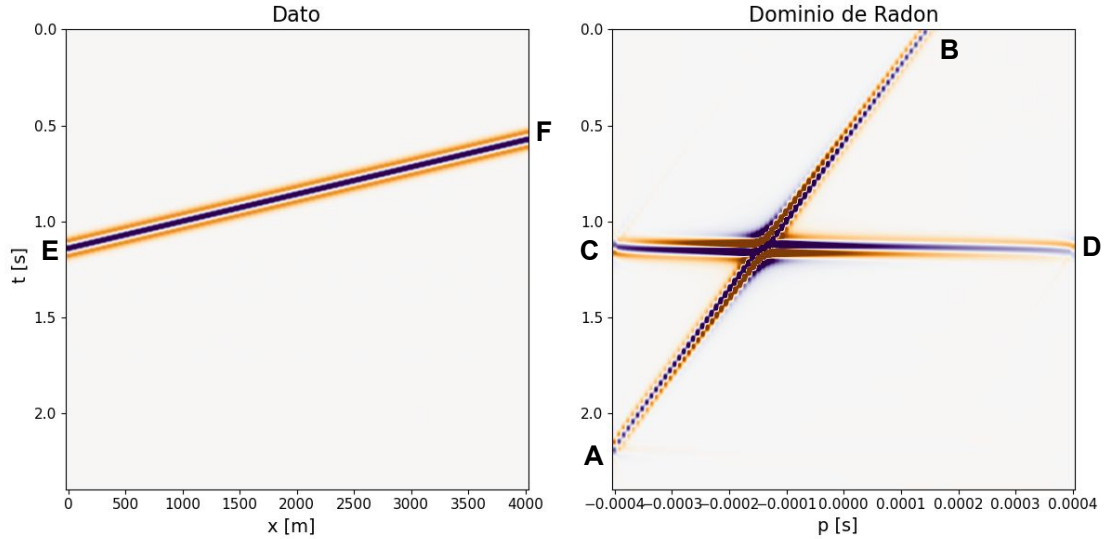
donde la matriz  $\mathbf{L}^T$  representa el operador adjunto aplicado al dato  $\mathbf{d}$ , y la matriz  $\mathbf{L}$  representa el operador directo aplicado a los coeficientes  $\mathbf{m}$ . Las ecuaciones 2.18 y 2.19 son análogas a las ecuaciones 2.15 y 2.17 respectivamente, solo que utilizan notación matricial. Dicha notación ofrece ventajas para operar con álgebra lineal. Por ejemplo, 2.18 y 2.19 facilitan el planteo de la función de costo a minimizar, cuestión que desarrollaremos más adelante. Sin embargo, en ningún momento formamos las matrices  $\mathbf{L}^T$  y  $\mathbf{L}$ , sino que son

reemplazadas por programas computacionales ejecutados al vuelo (*on the fly*). Dicho de otro modo,  $\mathbf{L}^T$  y  $\mathbf{L}$  representan los operadores de Radon adjunto y directo de forma implícita.

## 2.4. Observaciones

De las transformadas y ejemplos analizados anteriormente, podemos realizar las siguientes observaciones, de gran utilidad para desarrollar este trabajo:

- Los coeficientes de Radon se diferencian según qué tanta amplitud suman. Esto se ve claramente en el análisis de las figuras 2.1 y 2.2, donde se analizan coeficientes positivos, negativos y nulos. Podemos entonces concluir que tendremos coeficientes más grandes que otros.
- Al discretizar las expresiones de la TR, como hicimos en las ecuaciones 2.14 y 2.17, se generan efectos de borde. Recordemos que cualquier TR propone integrar entre  $(-\infty, \infty)$  la variable espacial. En ese contexto, ejemplificando con la TR lineal, una recta en el dominio  $(t - x)$  se vería como un punto en el  $(\tau, p)$ . Sin embargo, al truncar los offsets, los extremos de un evento lineal en  $(t - x)$  se ven como eventos sísmicos puntuales, razón por la cual mapean rectas en  $(\tau - p)$ . Por lo tanto, se dispersa energía debido a que no estará focalizada en un solo punto (Velis, 2022). Cuanto menor sea la cantidad de muestras en la variable espacial (lo que se denomina menor apertura), mayor será la dispersión de energía, y por ende más difícil el cálculo de la transformación directa que nos devuelve al dominio de los datos. Un ejemplo se muestra en la figura 2.3, donde a la izquierda aparece un evento lineal  $EF$ , y a la derecha su transformada lineal de Radon, con ganancia aplicada. El punto  $E$  mapeará en la recta  $AB$  y el  $F$  hará lo propio en  $CD$  (Yilmaz, 2001). Efectos de borde de este estilo no son propios solamente de la TR lineal. En el caso de la TR hipérbolica, el truncamiento a distancias fuente-receptor cortas genera líneas horizontales muy marcadas en el dominio de Radon. Por otro lado, truncar a distancias largas hace aparecer una curva descendente de menor amplitud. ¿Por qué no suceden estos efectos para todos los puntos del dominio  $(\tau - x)$ ? En verdad, sí suceden, pero en el dominio de Radon se cancelan con los puntos vecinos, es por ello que solo los veremos en los extremos.
- La TR no es ortogonal. Cuando una transformada es **ortogonal**, al aplicarla sobre un dato, y luego aplicar la transformada inversa, nos devuelve exactamente el mismo

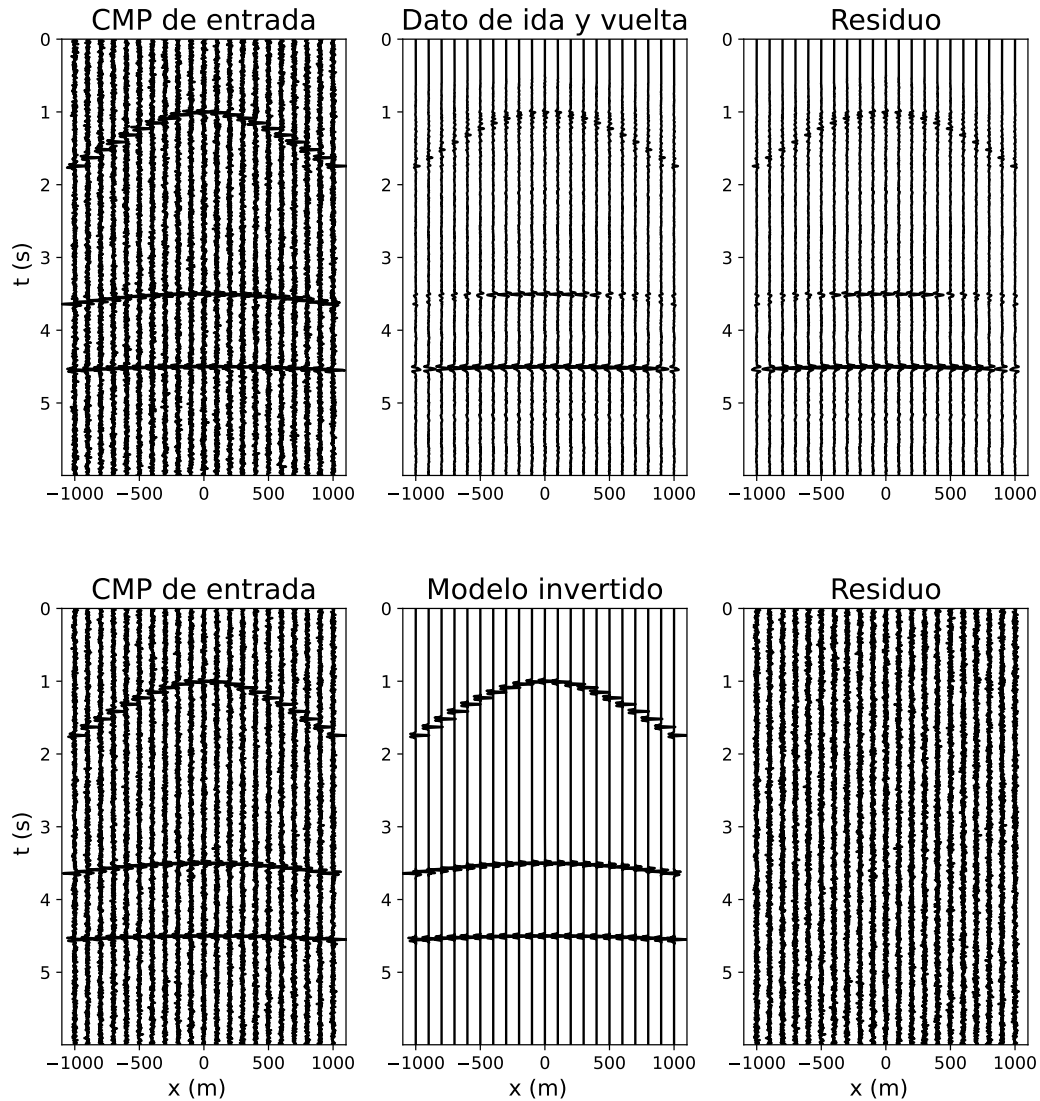


**Figura 2.3: Efectos de borde en TR lineal.** *Izq.* (a): CMP de entrada, contiene un solo evento lineal EF. *Der.* (b): Dominio transformado de Radon lineal del dato de entrada.

dato, como por ejemplo lo hace la transformada de Fourier. Por lo tanto, como en nuestro caso utilizaremos la TR, tendremos que realizar la inversión de los datos buscando cuáles son los coeficientes en el dominio de Radon que mejor ajusten. Este inconveniente lo resolveremos planteando un problema de mínimos cuadrados, utilizado para invertir los coeficientes del modelo que sintetizan los datos.

Un ejemplo sobre qué sucedería al realizar el camino de ida y vuelta sin realizar un ajuste de los coeficientes por inversión se muestra en la figura 2.4 (*arriba*). Allí, el dato de entrada es un registro de fuente común (*shot-gather*) sintético, con relación señal-ruido igual a 1 (panel izquierdo). En el panel central se muestra el resultado predicho por el modelo (el dato sintetizado). Finalmente, el residuo que se genera al calcular la diferencia entre los datos predichos por los coeficientes adjuntos y el dato de entrada se muestra en la figura derecha. Podemos ver que los coeficientes adjuntos no sirven para predecir el dato. En cambio, en la figura 2.4 (*abajo*) se puede ver cómo a partir del mismo dato de entrada logró un buen modelo gracias a la inversión de los coeficientes de Radon. El resultado es óptimo en este caso, y el residuo tan solo contiene ruido. El algoritmo utilizado en este último será explicado más adelante en este capítulo.





**Figura 2.4:** *Fila arriba:* Registro sintético al que se le aplica el operador de TR adjunto y luego directo (*centro*), y su residuo (*derecha*). *Fila abajo:* Registro sintético al cual se lo transformó al dominio de Radón, se invirtieron los coeficientes y se volvió al dominio  $(t - x)$  (*centro*), y residuo (*derecha*).

## 2.5. TR por inversión

Según mencionamos en las conclusiones de la sección anterior, no alcanza con transformar los datos de un dominio a otro, sino que debemos modelar los coeficientes, buscando que se parezcan lo más posible al dato de entrada. Para ello, podemos definir lo que se denomina **función de costo**, que tendrá como objetivo calcular las discrepancias entre el dato de entrada y la respuesta que se obtiene al calcular dicho modelo. Entonces, los parámetros óptimos serán aquellos que minimicen la función.

$$J(\mathbf{m}) = \|\mathbf{L}\mathbf{m} - \mathbf{d}\|_2^2 + \mu \Phi(\mathbf{m}) . \quad (2.20)$$

En nuestro caso,  $\mathbf{L}$  es el operador que transforma los coeficientes del dominio de Radón al dominio en que se encuentren los datos,  $\mathbf{m}$  es un vector que contiene a los coeficientes en el dominio de Radon (que a priori desconocemos, son quienes queremos encontrar),  $\mu$  se denomina parámetro de preblanqueo, y  $\Phi(\mathbf{m})$  es la matriz de pesos del modelo.

El primer término de la suma corresponde a lo que se denomina **error de ajuste** (*misfit*), que representa la norma L2 de la diferencia entre el modelo obtenido al aplicar  $\mathbf{L}$  a los coeficientes obtenidos del vector  $\mathbf{m}$  y el dato de entrada. Como el objetivo es que se parezcan lo más posible entre ambos, pedimos que la norma sea mínima.

A partir de minimizar la función de costo  $J$ , se encuentra el modelo  $\mathbf{m}$  que mejor ajusta al dato  $\mathbf{d}$ . Además, en ese proceso es posible imponerle ciertas condiciones a la solución. A continuación, se explican ciertas maneras de resolverla:

- Soluciones clásicas: Buscan minimizar la norma L2 del término de regularización. Así, la característica que le estaríamos dando al modelo obtenido, y por ende a la reflectividad, es que ésta sea ruido gaussiano. Son los métodos más estudiados históricamente, y en este trabajo utilizaremos un algoritmo de este tipo como control de calidad.
- Transformada de Radon de alta resolución (Sacchi y Ulrych, 1995): Lo que se impone en el término de regularización es la minimización de la norma L0 del modelo (o su relajación a la norma L1, dado que la 0 no tiene inversa). Realizar esta operación, implica penalizar las soluciones que tengan una mayor cantidad de elementos no nulos, ya que lo que representa la norma L0 es justamente eso. Por lo tanto, minimizarla nos llevara a una solución rala (*sparse*). No utilizaremos algoritmos de este estilo en esta Tesis.

Otra forma de obtener una solución rala, sin imponer condiciones sobre  $J$  de manera directa, está dada mediante algoritmos voraces (*greedy*) basados en la idea de *Matching Pursuit* (MP) (Mallat y Zhang, 1993), o en *Orthogonal Matching Pursuit* (OMP) (Pati et al., 1993). En OMP, el dominio invertido se restringe, o se construye iterativamente, con el objetivo de encontrar los mejores candidatos de soporte para modelar los datos, según distintos criterios. La ventaja es que suelen ser algoritmos rápidos y fáciles de aplicar. La minimización de la función de costo  $J$ , como se ha explicado anteriormente, será obtenida vía el método de mínimos cuadrados por gradientes conjugados (Hestenes y Stiefel, 1952).

# Capítulo 3

## Métodos

Tanto en sismica como en otras aplicaciones, las señales observadas contienen muestras con información de interés que suelen estar contaminadas por ruido indeseado. En sismica de reflexión, la energía ordenada en forma de hipérbolas representa la información, y lo demás se puede considerar como ruido. A partir de esto, podríamos preguntarnos si para cierto tipo de señales o en cierto dominio transformado, un pequeño número de muestras en particular contienen suficiente información como para aproximar bien a la señal de interés. Este área de investigación se denomina *Compressive Sampling* (CS) (Needell, 2009).

Haremos uso de la TR en este trabajo. La idea es encontrar las coordenadas  $(\tau, v)$  que mejor representen a las hipérbolas de reflexión y así poder aislarlas. Este análisis sería muy difícil de hacer en el dominio  $(t - x)$ . Recordemos que en el dominio de Radon tan solo tendríamos que aislar “puntos”. De este modo, mediante distintos criterios iremos buscando ese pequeño número de muestras y analizaremos una relación de compromiso entre cuál es más eficaz, cuál requiere menor cantidad de coeficientes y cuál es el que mejor ajusta al dato. ¿Por qué querríamos disminuir la cantidad de coeficientes? Debido a que cada vez que se aplican los operadores de Radon adjunto y directo, realizar dicha acción conlleva un costo computacional proporcional a la cantidad de velocidades a analizar ( $N_v$ ), al número de muestras en tiempo ( $N_t$ ) y al número de *offsets* ( $N_x$ ). Si el tamaño del dato es  $N_t \times N_x$ , y consideramos  $N_t = N_\tau$ , el costo será controlado por  $N_\tau$  y  $N_v$ , o sea el tamaño del dominio de Radon (Sabbione y Sacchi, 2016).

A lo largo de este capítulo desarrollaremos los fundamentos teóricos de los distintos métodos con los que buscaremos modelar la información del dato de entrada. Primero describiremos la inversión por mínimos cuadrados amortiguados (*Damped least squares*, DLS), que podemos definirlo como un método clásico de inversión. Lo utilizaremos en

algunas pruebas como control de calidad. Luego analizaremos tres algoritmos que buscan disminuir la cantidad de coeficientes, denominados métodos **voraces** (*greedy*), basados en la idea del *compressive sampling*. Al final del capítulo, mostramos un análisis gráfico sobre cómo es la selección de cada uno de los métodos.

### 3.1. Mínimos cuadrados amortiguados (DLS)

Consiste en minimizar la función de costo, dada por un término asociado al error de ajuste, y otro término de regularización, que le da el nombre de “amortiguados” (ec. 2.20) (*Damped least squares*) (Levenberg, 1944). Se utilizarán todos los coeficientes del dominio de Radon para realizar la inversión, lo que lo vuelve un método costoso, aunque tenemos asegurada su convergencia.

Veamos entonces el procedimiento esquemático a realizar:

- **Entrada:** Dato  $\mathbf{d}$ .

- **Salida:** Todas coordenadas  $(\tau, v)$  y el vector  $\tilde{\mathbf{m}}$  que minimiza la función de costo.

1. Transformamos el dato de entrada al dominio de Radon, obteniendo lo que se denominan coeficientes adjuntos ( $\mathbf{m}_{adj}$ ).
2. Proponemos la matriz de pesos  $\Phi(\mathbf{m})$ , definida en la ecuación 2.20, tal que esté dada por la inversa de las amplitudes de los coeficientes adjuntos. Así, si para un cierto par  $(\tau, p)$  la suma es grande, la inversa será pequeña, por lo que la función de costo no penalizará ese coeficiente, y viceversa. Es necesario aclarar que se le sumará un pequeño valor no nulo  $\epsilon$  para estabilizar la inversión.

$$\Phi_j(m_{adj_j}) = \frac{1}{|m_{adj_j}| + \epsilon}, \quad (3.1)$$

donde  $m_{adj_j}$  es un coeficiente.

3. Obtenemos los coeficientes  $\tilde{\mathbf{m}}$  del modelo que mejor ajustan al dato a partir de la minimización de la función de costo  $J$  (ec. 2.20) vía gradientes conjugados.

Para obtener el dato sintetizado volvemos al dominio  $(t - x)$  a partir del algoritmo, utilizando la TR directa sobre los coeficientes obtenidos en el paso 3, y obtenemos el dato que predice el modelo invertido.

$$\mathbf{d}_s = \mathbf{L} \tilde{\mathbf{m}}.$$

## 3.2. TR Hiperbólica con dominio restringido (RHRT)

Sus siglas se asocian a *Restrcted hyperbolic Radon transform* (RHRT), y fue presentado por Sabbione y Sacchi (2016). Observando la ecuación 2.15, es intuitivo que los coeficientes  $m_{adj}(\tau, v)$  que coincidan con una reflexión sean grandes, y aquellos que sumen a lo largo de ruido aleatorio sean muy pequeños. De esta manera, podemos ordenar a los coeficientes de mayor a menor.

Entonces, para formar un subconjunto con los coeficientes más significativos elegimos aquellos que más energía suman:

$$\mathbb{A} = \{(\tau, v) : \frac{1}{N_x} |m_{adj}(\tau, v)| > T\} , \quad (3.2)$$

donde  $N_x$  está presente para normalizar la suma y así poder variar el valor umbral  $T$  entre cero y uno. El objetivo es encontrar aquellos pares  $(\tau, v)$  que sean más importantes, ya que para esas velocidades y esos tiempos de ida y vuelta  $\tau$ , lo más probable es que exista una reflexión importante en el dato de entrada  $d$ .

La minimización de función de costo asociada a este problema nos devolverá los coeficientes  $\tilde{\mathbf{m}}$  que mejor ajusten al dato, y vendrán dados por:

$$\tilde{\mathbf{m}} = \underset{\mathbf{m}}{\operatorname{argmin}} ( \|\mathbf{L}_{\mathbb{A}} \mathbf{m}_{\mathbb{A}} - \mathbf{d}\|_2^2 + \mu \|\mathbf{W}_{\mathbb{A}} \mathbf{m}_{\mathbb{A}}\|_2^2 ) , \quad (3.3)$$

donde  $\mathbf{L}_{\mathbb{A}}$  y  $\mathbf{L}_{\mathbb{A}}^T$  son los los operadores para encontrar los coeficientes de Radon. El subíndice  $\mathbb{A}$  nos indica que solamente utilizaremos los elementos que se encuentran dentro de ese conjunto. La matriz de pesos  $\mathbf{W}_{\mathbb{A}} = \operatorname{diag}(|\mathbf{L}_{\mathbb{A}} \mathbf{d}|^{-1})$  busca lo mismo que en el método anterior, penalizar los  $m_{adj}(\tau, v)$  pequeños y no afectar a los más grandes. No se requiere sumar un término  $\epsilon$  para darle estabilidad ya que al elegir valores de  $|m_{adj}(\tau, v)|$  grandes, no hay riesgo de dividir por cero.

El procedimiento, esquemáticamente, es el siguiente:

- **Entrada:** Dato  $\mathbf{d}$ , porcentaje de coeficientes a tener en cuenta (lo que determina el umbral  $T$ ).
- **Salida:** Coeficientes del modelo  $\tilde{\mathbf{m}}$  y sus coordenadas en  $\mathbb{A}$ .

1. Cálculo de los coeficientes adjuntos:

$$\mathbf{m}_{adj} = \mathbf{L}^T \mathbf{d} .$$

2. Restricción del dominio de Radon a partir del criterio 3.2. Obtenemos el espacio  $\mathbb{A}$ .

3. Dentro del espacio  $\mathbb{A}$ , minimizamos la función de costo como se muestra en ec. (3.3) y se obtienen los coeficientes  $\tilde{\mathbf{m}}$ .

Para obtener el dato sintetizado a partir del modelo que invierte el algoritmo, utilizando la TR directa volvemos al dominio  $(t - x)$  con los coeficientes obtenidos en el paso 3:

$$\mathbf{d}_s = \mathbf{L} \tilde{\mathbf{m}}_{\mathbb{A}} .$$

### 3.3. Métodos voraces (*greedy*) iterativos

La idea de estos algoritmos es ir desarmando el dato a partir de la selección criteriosa e iterativa de los coeficientes más representativos, y así obtener una solución rala. Utilizaremos las ideas de “*Orthogonal Matching Pursuit*” (OMP) (Pati et al., 1993; Tropp y Gilbert, 2007) e intentaremos presentar un algoritmo voraz iterativo.

Para todos estos métodos, necesitaremos un **conjunto de candidatos modelo**. Como estamos trabajando en el dominio de Radon, esos candidatos serán las coordenadas  $(\tau, v)$  y su amplitud correspondiente al aplicarle el operador inverso de Radon. Necesitamos también un **criterio de selección** para encontrar los mejores candidatos. En el caso de OMP, ese criterio de selección será quedarse con el par  $(\tau, v)$  cuya amplitud sea más grande. Luego, necesitaremos una **función de factibilidad** que nos diga si ese modelo puede contribuir realmente a la solución, o sea, el operador directo de Radon. Debemos definir una **función de costo**, para poder asignarle un valor a la solución, que en nuestro caso será la minimización de  $J$  (ec. 2.20). Por último, se define un **criterio de corte** para darle un fin a las iteraciones (Wang et al., 2009). En nuestros métodos, lo asociaremos al número de iteraciones.

Esquemáticamente:

- **Entrada:** Dato  $\mathbf{d}$ , número de iteraciones (a veces denominado “nivel ralo”, *s-sparse*). En este caso, además, será igual a la cantidad de coordenatas  $(\tau, v)$  elegidas.
- **Salida:** El conjunto de coordenadas seleccionadas  $\mathbb{A}$  y sus coeficientes  $\tilde{\mathbf{m}}$  asociados.

1. Calcular el residuo, inicializando con  $\mathbf{d}_{res_1} = \mathbf{d}$  y definir el conjunto de elementos elegidos, que inicializamos como el conjunto vacío  $\mathbb{A} = \{\emptyset\}$
2. Se repite iterativamente  $s$  veces lo siguiente

a) Cálculo de los coeficientes adjuntos

$$\mathbf{m}_{adj}^i = \mathbf{L}^T \mathbf{d}_{res_i} \quad (3.4)$$

b) Elegimos el par  $(\tau_{max}, v_{max})^i$  correspondiente al valor  $|m_{adj}(\tau, v)|$  más grande y actualizamos el subespacio  $\rightarrow \mathbb{A} = \{\mathbb{A} \cup (\tau_{max}, v_{max})^i\}$

c) Minimizamos la función de costo  $J$  en todo el subespacio de candidatos actualizado, y hallamos  $\tilde{\mathbf{m}}^i$ .

d) Si no es la última iteración, actualizamos  $\mathbf{d}_{res_{i+1}} = \mathbf{d} - \mathbf{L} \tilde{\mathbf{m}}_i$ .

Para predecir los datos a partir del algoritmo, utilizaremos el operador directo de Radon sobre el modelo invertido y obtenemos el dato sintetizado:

$$\mathbf{d}_s = \mathbf{L} \tilde{\mathbf{m}}_{\mathbb{A}}$$

Así, estaremos agregando criteriosamente en cada iteración una hipérbola nueva en el modelo a partir de los datos. El OMP nos sirvió para entender el concepto de voracidad y lo que significa que un método sea iterativo, pero tiene tiempos de convergencia muy lentos ya que, en el mejor de los casos, deberíamos hacer una iteración por cada hipérbola que se encuentre en el dato de entrada. ¿Cómo podemos solucionar este problema? Haciendo que en cada iteración ingrese una mayor cantidad de coeficientes.

### 3.3.1. Transformada de Radon voraz (GRT)

Este método fue presentado como transformada de Radon rápida y de alta resolución mediante métodos voraces de mínimos cuadrados (*greedy Radon transform*, GRT) (Wang et al., 2009).

El criterio de selección en este caso está dado por un **valor umbral** (*threshold value*,  $thr$ ) multiplicado por el máximo valor de  $|m_{adj}(\tau, v)|$ . Se seleccionan todos los pares  $(\tau, v)$  cuyos coeficientes  $|m_{adj}(\tau, v)|$  superen dicho producto. Formalmente:

$$\mathbb{A}^i = \{(\tau, v) : |m_{adj}(\tau, v)| > thr \cdot |m_{adj}(\tau, v)|_i^{max}\} . \quad (3.5)$$

Según sus autores (Wang et al., 2009), seleccionar cuidadosamente el valor umbral lleva a aumentar la resolución en el dominio de Radon. Además, los autores proponen la idea de que cuanto más alto sea ese valor umbral, más alta será la resolución. Esta última afirmación es implica cierto riesgo, ya que si aumentáramos demasiado el valor umbral, serían muy pocos los valores de  $(\tau, v)$  que cumplan con la condición (3.5). En ese



contexto, podría suceder que entre las coordenadas que no la cumplen se encuentre alguna correspondientes a información y no a ruido. Entonces, en la búsqueda de aumentar la resolución estaríamos dañando la señal.

El procedimiento que el algoritmo pone en práctica es el siguiente:

- **Entrada:** Dato  $\mathbf{d}$ , número de iteraciones ( $N_{it}$ ) y valor umbral ( $thr$ ).
- **Salida:** El conjunto de coordenadas seleccionadas  $\mathbb{A}$  y sus correspondientes coeficientes  $\tilde{\mathbf{m}}$  asociados a los valores de  $\mathbb{A}$ .

1. Inicializamos el residuo  $\mathbf{d}_{res_1} = \mathbf{d}$ , siendo  $\mathbf{d}$  el dato de entrada y definimos un vector solución del modelo, que inicializamos como  $\mathbf{m}_{sol} = \vec{0}$ .
2. Se repite iterativamente el siguiente bucle hasta que el criterio de corte, dominado por el número de iteraciones, lo indique.
  - a) Cómputo de los coeficientes adjuntos

$$\mathbf{m}_{adj}^i = \mathbf{L}^T \mathbf{d}_{res_i} . \quad (3.6)$$

- b) Elegimos las coordenadas  $(\tau, v)^i$  que cumplan con el criterio 3.5, formando un subespacio  $\mathbb{A}^i$ .
- c) Dentro del subespacio  $\mathbb{A}^i$ , minimizamos la función de costo:

$$J(\mathbf{m}^i) = \|\mathbf{d}_{res_i} - \mathbf{L} \mathbf{m}^i\|^2 ,$$

y obtenemos la solución  $\tilde{\mathbf{m}}^i$ .

- d) Actualizamos la solución del modelo  $\mathbf{m}_{sol} = \mathbf{m}_{sol} + \tilde{\mathbf{m}}^i$ .
- e) Si no es la última iteración, actualizamos  $\mathbf{d}_{res_{i+1}} = \mathbf{d}_{res_i} - \mathbf{L} \tilde{\mathbf{m}}^i$ .

Si es la última iteración, debemos tomar el vector  $\mathbf{m}_{sol}$  y eliminar todas aquellas coordenadas  $(\tau, p)$  tales que  $m_{sol}(\tau, p) = 0$ , ya que para esos valores no hubo información representativa.

Para obtener la síntesis del dato a partir del algoritmo, utilizamos el operador directo de Radon sobre el modelo final y obtenemos

$$d_s = \mathbf{L} \tilde{\mathbf{m}}_{\mathbb{A}} .$$

### 3.3.2. Búsqueda coincidente por etapas (StOMP)

Este método propone un criterio de selección más explícito en cuanto a la naturaleza de una señal. Se basa en el trabajo *Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit* (StOMP) (Donoho et al., 2012). Allí, los autores afirman que los coeficientes de un dominio (en nuestro caso, el de Radon) son una mezcla de señal verdadera y ruido gaussiano sumado al que se produce por la no ortogonalidad de la TR (*cross-talk noise*). Por lo tanto, si pudiéramos elegir apropiadamente el valor umbral, es posible sustraer la señal disminuyendo el ruido no deseado en cada iteración.

Asumiendo una distribución gaussiana de la señal, se entiende que tendremos muchos coeficientes pequeños distintos de cero distribuidos como ruido. Por lo tanto, deberíamos elegir un umbral fuerte tal que detecte los valores que sean significativamente distintos de cero. Esa restricción devolverá un conjunto de pocas coordenadas tales que:

$$\mathbb{A}^i = \{(\tau, v) : |m_{adj}(\tau, v)| > thr_i \cdot \sigma_i\} , \quad (3.7)$$

siendo  $thr_i$  el valor umbral que funcionará como escala de  $\sigma_i$ , el nivel de ruido gaussiano.

$$\sigma_i = \frac{\|\mathbf{m}_{adj}\|_2}{\sqrt{N_{\mathbf{m}_{adj}}}} .$$

El procedimiento se muestra a continuación.

- **Entrada:** Dato  $\mathbf{d}$ , número de iteraciones ( $N_{it}$ ), y valor umbral ( $thr$ ).
  - **Salida:** El conjunto de coordenadas seleccionadas  $\mathbb{A}$  y sus correspondientes coeficientes  $\tilde{\mathbf{m}}$  asociados a los valores de  $\mathbb{A}$ .
1. Inicializamos el residuo como  $\mathbf{d}_{res} = \mathbf{d}$ . Definimos el espacio donde iremos guardando cada coeficiente obtenido, inicializándolo vacío:  $\mathbb{A} = \{\emptyset\}$ .
  2. Se repite iterativamente el siguiente bucle hasta que el criterio de corte, dado por el número de iteraciones, lo indique.

a) Cálculo de los coeficientes adjuntos:

$$\mathbf{m}_{adj}^i = \mathbf{L}^T \mathbf{d}_{res_i} . \quad (3.8)$$

b) Calculamos el valor de  $\sigma_i$ .

- c) Elegimos las coordenadas  $(\tau, v)^i$  que cumplan con el criterio 3.7, actualizando las coordenadas del subespacio  $\rightarrow \mathbb{A}^i = \{\mathbb{A}^{i-1} \cup (\tau_{selec.}, v_{selec.})^i\}$ .
- d) Dentro del espacio  $\mathbb{A}^i$ , minimizamos la función de costo:

$$J(\mathbf{m}^i) = \|\mathbf{d}_{resi} - \mathbf{L} \mathbf{m}^i\|_2^2 + \|\mathbf{I}_{N_{\mathbb{A}^i}} \mathbf{m}^i\|_2^2,$$

donde  $\mathbf{I}_{N_{\mathbb{A}^i}}$  es una matriz diagonal de dimensión igual a la cantidad de coordenadas en  $\mathbb{A}$  ( $N_{\mathbb{A}^i}$ ) y se obtiene la solución  $\tilde{\mathbf{m}}^i$ .

- e) Si no es el último paso, actualizamos  $\mathbf{d}_{resi+1} = \mathbf{d} - \mathbf{L}\tilde{\mathbf{m}}^i$ .

Una vez finalizado el bucle, el método nos devolverá los valores de los coeficientes  $\tilde{m}^i$  y el espacio  $\mathbb{A}$ .

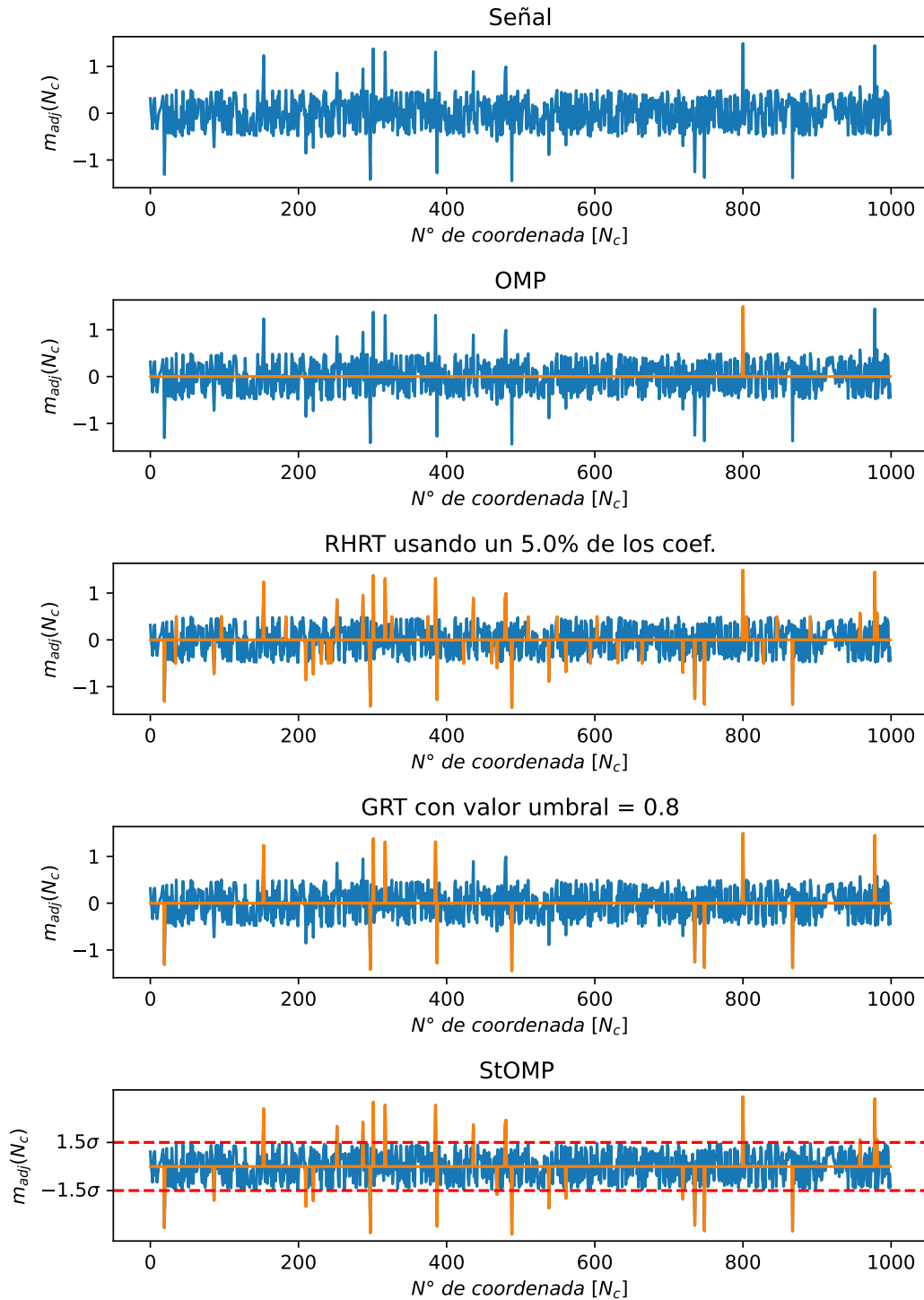
### 3.3.3. Potenciales ventajas de StOMP sobre GRT

Una breve lista de puntos destacables del el método StOMP respecto a GRT es la siguiente:

- StOMP en cada iteración ajusta a todo el paquete de coeficientes  $\mathbf{m}$  correspondientes a cada coordenada que se encuentra en el espacio  $\mathbb{A}$ , sin importar si ya había sido ajustada o no. Esto lo hace mediante el método de CG, lo que hace que le de a todos los coeficientes igual importancia. En contrapunto, GRT calcula la minimización de  $J$  solo para los elementos seleccionados, y luego ajusta mediante una suma los valores anteriormente obtenidos.
- El residuo en StOMP se calcula para todas las iteraciones partiendo del dato original. En cambio, GRT arrastra el residuo de la iteración anterior y le sustrae el modelo obtenido en cada paso.

## 3.4. Apéndice gráfico

Aquí analizamos gráficamente cómo es que cada uno de los métodos selecciona los coeficientes en cada iteración (fig. 3.1). Teniendo en cuenta que estos algoritmos funcionan para cualquier tipo de señal (no necesariamente sísmica, y sin necesidad de introducir el concepto de TR), propondremos una señal genérica. Ésta, se forma por mil coeficientes de distinta amplitud y posición en el eje horizontal. Algunos corresponden a información (de distribución rala), y la gran mayoría a ruido gaussiano. En naranja, se muestran los



**Figura 3.1:** Criterios de selección para los métodos. En azul se muestra la señal sintética de entrada. Superpuesta a ella, en naranja, están los coeficientes que seleccionaría cada método. En el caso de OMP, GRT, y StOMP, la imagen representa la selección de una sola iteración. En la figura correspondiente a StOMP, la línea punteada representa el criterio de selección igual a  $|thr \cdot \sigma|$ .

coeficientes seleccionados por cada uno de los métodos. Para los métodos iterativos, se representan los elegidos en cada iteración.

1. Se muestra la señal original. El método de DLS buscará reconstruir el modelo a partir de todos los coeficientes visibles.
2. OMP: En este caso, una sola iteración seleccionará el coeficiente más grande.
3. RHRT: A partir del criterio 3.2, con un valor umbral del 5%, obtendríamos este resultado. Vemos como en algunos casos, por ejemplo en posiciones cercanas a los 50 primeros coeficientes, selecciona algunos que son pequeños (ruido gaussiano) que los otros métodos no eligen.
4. GRT en una sola iteración depende mucho del valor umbral elegido. Por ejemplo, en este caso, selecciona pocos coeficientes correspondiente a los valores más grandes.
5. StOMP en una iteración, seleccionará aquellos coeficientes cuyo módulo esté por encima de la línea punteada roja, que toma el valor de un factor de escala multiplicado por la desviación estándar  $\sigma$ . El propósito es conseguir los coeficientes que excedan el ruido gaussiano.

# Capítulo 4

## Aplicación sobre datos sintéticos

En este capítulo se aplican los métodos anteriormente explicados en datos sísmicos sintéticos. En primera instancia, se analiza cómo se comporta cada algoritmo para diversos datos de entrada con diferentes características. Para verificar que es posible realizar la inversión de dichos datos de entrada, también se predicen los sintéticos en el dominio de Radon mediante DLS, dado que su convergencia está probada. Luego, para cada método *greedy*, se compara en cuánto se ve afectada la síntesis a partir de variar los parámetros de entrada propios de cada método. En caso de que alguno resulte estable ante tales estímulos, diremos que es un algoritmo **robusto**. El principal objetivo de la sección es encontrar cuál de los métodos *greedy* propuestos es más eficaz, cuál reconstruye de forma más exacta al dato de entrada, y cuál el más robusto.

### 4.1. Procedimiento, condiciones y criterios

Se analizan las inversiones obtenidas a partir de aplicar los algoritmos que implementan cada método. RHRT, GRT y StOMP fueron explicados en las secciones 3.2, 3.3.1, y 3.3.2, respectivamente. Cabe destacar que los coeficientes de Radon  $\tilde{\mathbf{m}}$  (ec. 2.17) obtenidos son distintos según cada método. Una vez aplicado el flujo, el objeto de estudio son los datos sintetizados y la diferencia entre el dato sintético de entrada y dicha síntesis, denominada **residuo**. El sintético utilizado está formado por hipérbolas de reflexión sumergidas en ruido aleatorio con diferentes amplitudes. En el caso extremo de que la predicción de las hipérbolas sea perfecta, esperaríamos ver solamente hipérbolas en la inversión y únicamente ruido aleatorio en el residuo.

Los parámetros utilizados para el dato sintético de entrada se describen en la tabla 4.1

Descripción de parámetro	Valor
Cantidad de trazas	$N_t = 21$
Distancia entre trazas	$\Delta t = 100 \text{ m}$
Frec. pico de la ondícula	$f_0 = 20 \text{ Hz}$
Intervalo de muestreo	$\Delta t = 0.004 \text{ s}$

**Tabla 4.1:** Parámetros utilizados para los registros sintéticos de entrada.

Para la generación de los datos sintéticos, se definen cierta cantidad de hipérbolas con sus respectivos parámetros: velocidad, amplitud y tiempo doble. Recordemos que al trabajar con datos de CMP, el ápice de las hipérbolas está centrado, siendo igual a cero. A este sintético lo llamaremos dato limpio ( $\mathbf{d}_{limpio}$ ) y un ejemplo se muestra en la figura 4.1. El procedimiento para obtener información inmersa en ruido aleatorio (dato ruidoso,  $\mathbf{d}_{ruidoso}$ ) implica realizar la siguiente suma:

$$\mathbf{d}_{ruidoso} = \mathbf{d}_{limpio} + \mathbf{ruido} , \quad (4.1)$$

donde

$$\mathbf{ruido} = \frac{\mathbf{r}_g}{\|\mathbf{r}_g\|_2} \cdot \frac{\|\mathbf{d}_{limpio}\|_2}{snr} , \quad (4.2)$$

siendo  $\mathbf{r}_g$  una matriz de igual dimensión que la correspondiente al dato sintético, cuyos coeficientes representan ruido gaussiano de banda limitada y normalizado por el denominador  $\|\mathbf{r}_g\|_2$ . Además, se amplifica por el cociente entre la norma del dato limpio y la relación señal-ruido ( $snr$ ) impuesta. Notar que cuanto menor sea el valor de  $snr$ , mayor amplitud tiene el ruido sumado al dato.

Idealmente, la cantidad de energía de la señal presente en el residuo debe ser nula, entendiéndose como señal a las hipérbolas de reflexión. Por ende, los algoritmos iterativos deben iterar tanto como sea necesario para alcanzar este objetivo. Bajo el mismo criterio, se busca el porcentaje de coeficientes suficiente para que RHRT reconstruya enteramente la señal.

Los criterios de comparación entre los métodos son los siguientes:

- **Error de ajuste (*misfit*,  $e$ ):** Comparamos la norma L2 al cuadrado de la diferencia entre el dato de entrada y la síntesis obtenida. Expresado matemáticamente:

$$e = \|\mathbf{d} - \mathbf{L} \mathbf{m}\|_2^2 . \quad (4.3)$$

Tomando en cuenta que el dato es sintético y conocido, entonces se puede discriminar entre qué parte es información y cuál no. Esto se puede ver explícitamente en la ecuación 4.1. Por lo tanto, es posible calcular el error de ajuste esperado, ya que si la reconstrucción resulta ser perfecta, entonces  $\mathbf{L m} = \mathbf{d}_{limpio}$ . En tal caso, el  $e$  esperado se obtiene haciendo:

$$\mathbf{e}_{esperado} = \|\mathbf{d}_{ruidoso} - \mathbf{d}_{limpio}\|_2^2 = \|\mathbf{ruido}\|_2^2. \quad (4.4)$$

Para normalizar este criterio de comparación, se realiza el cociente entre el error de ajuste obtenido y el esperado, es decir, un error relativo. Por consiguiente, en el caso de una predicción perfecta de la información, dicho cociente es igual a 1. Esta operación no se podrá realizar en el caso del dato sin ruido, ya que el valor de  $e_{esperado}$  tiende a cero, lo que hace que el cociente tienda a infinito. Para este caso se muestra solamente el error absoluto de ajuste obtenido (*misfit*,  $e$ ).

- **Relación señal-ruido (SNR):** De igual manera, al predecir únicamente la información, es posible comparar cuál es la relación señal-ruido que hay entre el dato de entrada y la síntesis. Cuando la reconstrucción aproxima de buena manera a la información, el parámetro *snr* resultante es cercano al impuesto en la ecuación 4.2 utilizado para generar el sintético. La manera de calcularlo es haciendo el cociente de las normas entre la síntesis obtenida y el dato de entrada:

$$snr = \frac{\|\mathbf{L m}\|_2}{\|\mathbf{d}\|_2}, \quad (4.5)$$

- **Cantidad de coeficientes utilizados:** Cuanto menor es la cantidad de coeficientes utilizados para alcanzar una correcta predicción de la información (según los criterios anteriormente explicados), más efectivo es el método.
- **Tiempo de cómputo:** Cuanto menor es el tiempo de cómputo requerido para alcanzar una correcta predicción de la información (según los criterios anteriormente explicados), más eficaz es el método.

## 4.2. Pruebas

### 4.2.1. Variación de relación *snr*

El dato de entrada es un registro sintético que contiene tres hipérbolas, con tiempos y velocidades que se muestran en la tabla 4.2.



Tiempo doble (s)	Vel (m/s)	Amplitud
1	700	1
3.5	1000	-1
4.5	1500	1

**Tabla 4.2:** Parámetros de las tres hipérbolas de reflexión para el registro sintético.

Método	$snr$	$Misfit$	Cant. de coef. [%]	Tiempo [s]
DLS	186.5	0.01	100	37.5
RHRT	79.4	0.04	2	2.96
GRT	17.53	0.77	0.18	7.05
StOMP	111.97	0.02	1.2	9.63
Resultado deseado	<i>Inf.</i>	0.0		

**Tabla 4.3:** Relación señal-ruido, error de ajuste ( $misfit$ ), cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 1.

**Dato sintético 1 (sin ruido):** Un método de inversión lo primero que debe verificar es su capacidad de reconstruir un dato limpio. Es por eso, que sin aplicarle ruido de ningún tipo se analizan las inversiones y los residuos obtenidos.

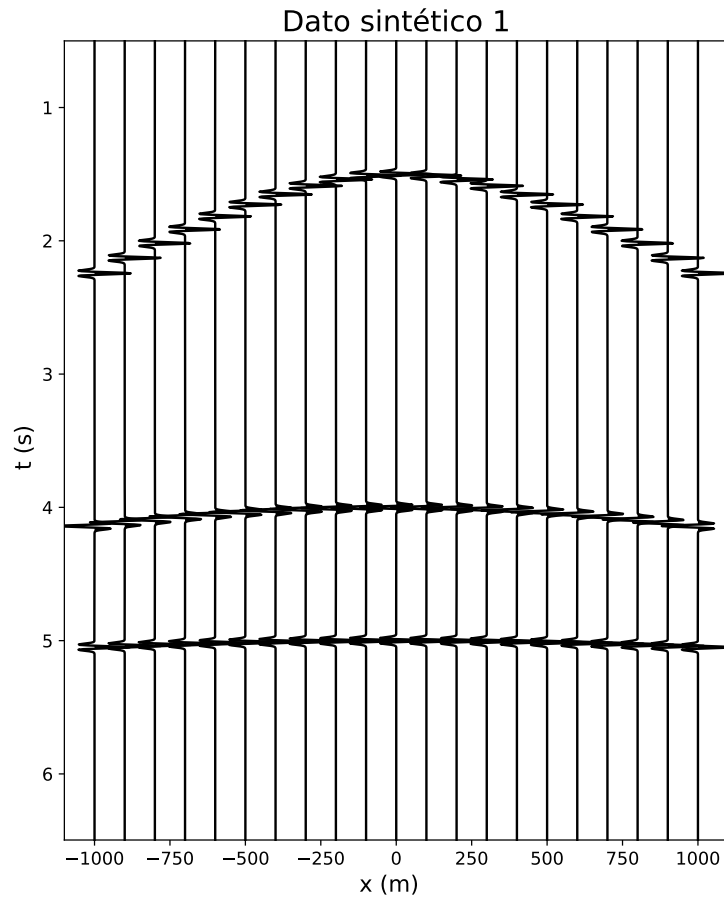
El sismograma de entrada se muestra en la figura 4.1. Los resultados gráficos para cada uno de los métodos se encuentran en la figura 4.2. La tabla 4.3 presenta los resultados numéricos para los cuatro criterios de comparación.

**Dato sintético 2 ( $snr = 1$ ):** Se busca la reconstrucción de la información a partir de un dato de entrada con relación señal-ruido media.

El dato de entrada se muestra en la figura 4.3. Los resultados gráficos para cada uno de los métodos se encuentran en la figura 4.4. La tabla 4.4 presenta los resultados numéricos para los cuatro criterios de comparación.

**Dato sintético 3 ( $snr = 0.5$ ):** Se busca la reconstrucción de la información a partir de un dato de entrada con mala relación señal-ruido.

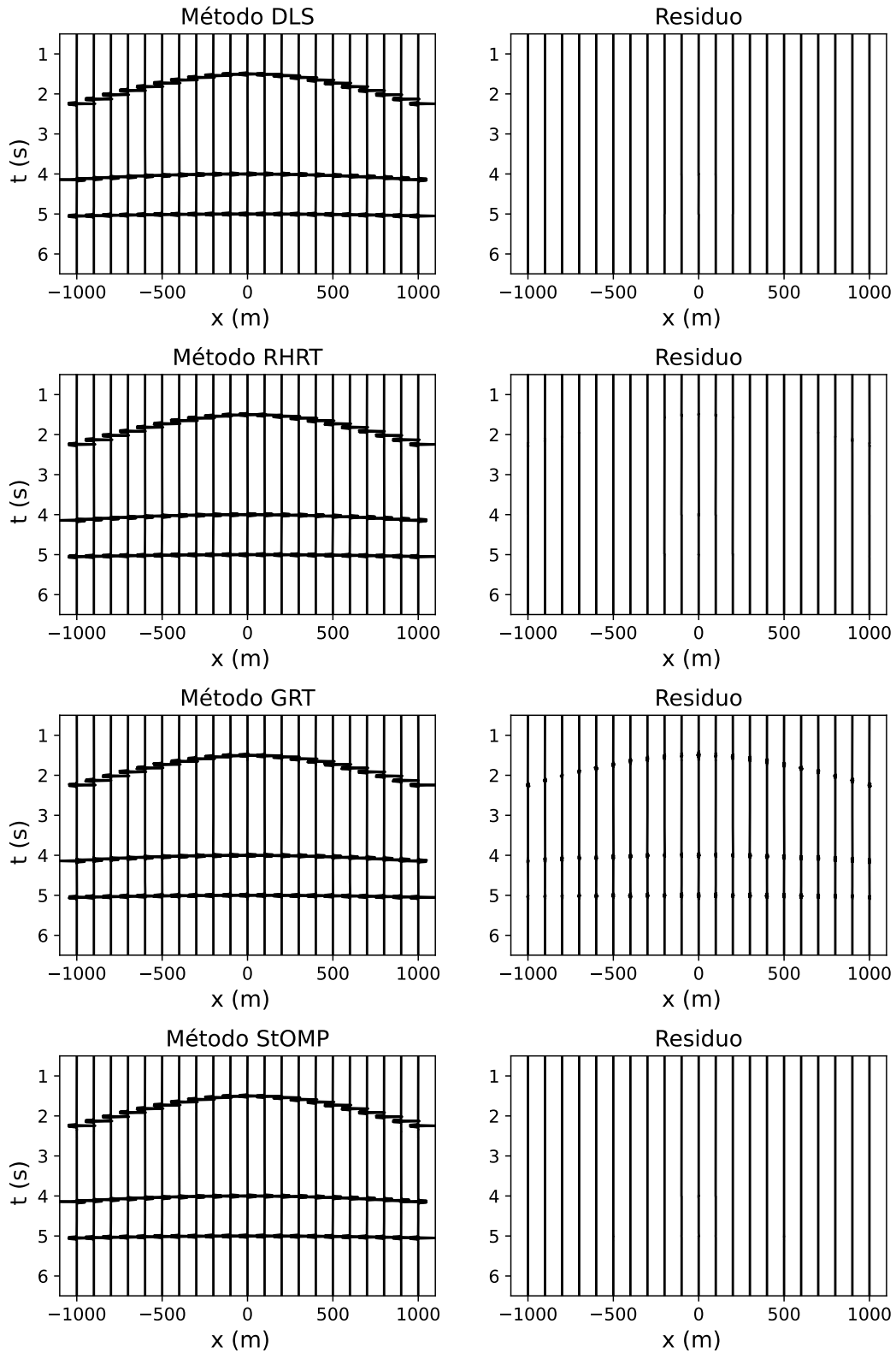
El dato de entrada se muestra en la figura 4.5. Los resultados gráficos para cada uno de los métodos se encuentran en la figura 4.6. La tabla 4.5 presenta los resultados numéricos para los cuatro criterios de comparación.



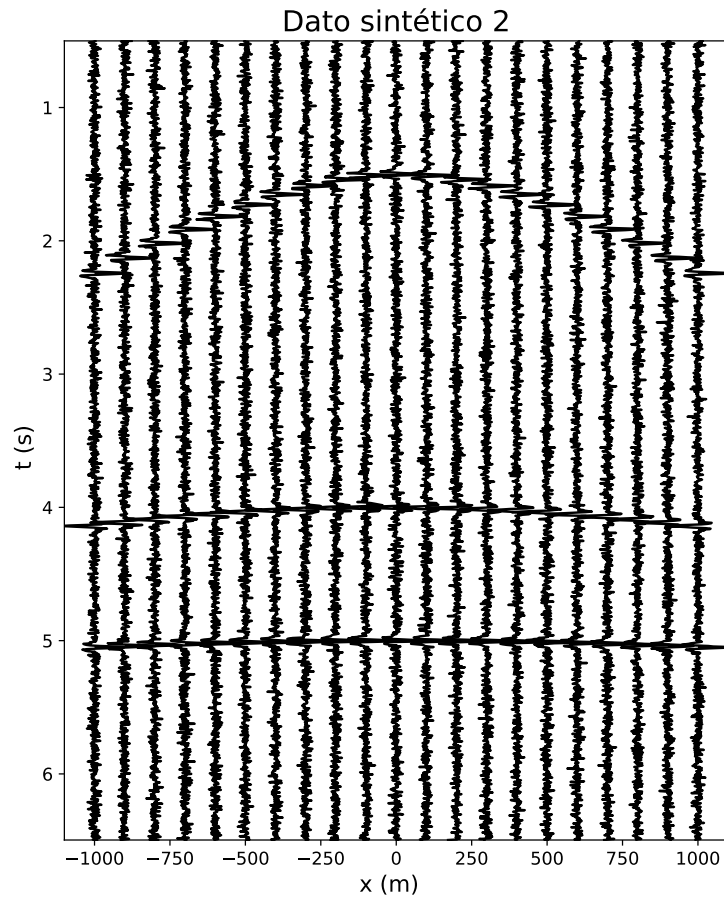
**Figura 4.1:** Registro sintético 1, generado a partir de la información correspondiente a las tablas 4.1 y 4.2 sin ruido adicionado.

Método	$snr$	$Misfit$ Norm.	Cant. de coef. [%]	Tiempo [s]
DLS	1.48	0.62	100	34.7
RHRT	1.00	1.00	0.4	0.96
GRT	1.00	1.02	0.08	2.08
StOMP	1.00	1.00	0.32	2.98
Resultado deseado	1.00	1.00		

**Tabla 4.4:** Relación señal-ruido, error de ajuste ( $misfit$ ) normalizado, cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 2.



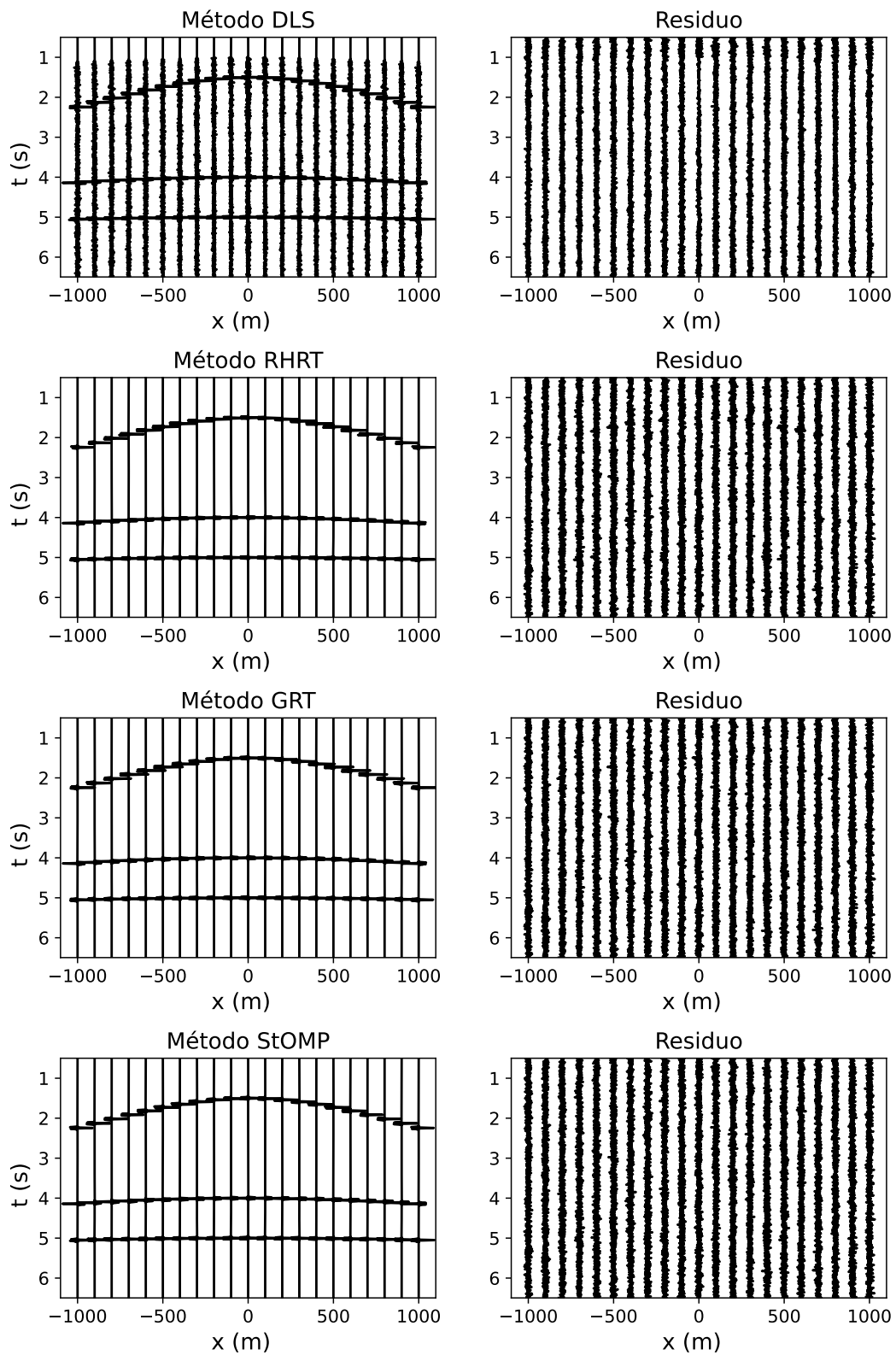
**Figura 4.2:** Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando un dato de entrada sin ruido adicionado.



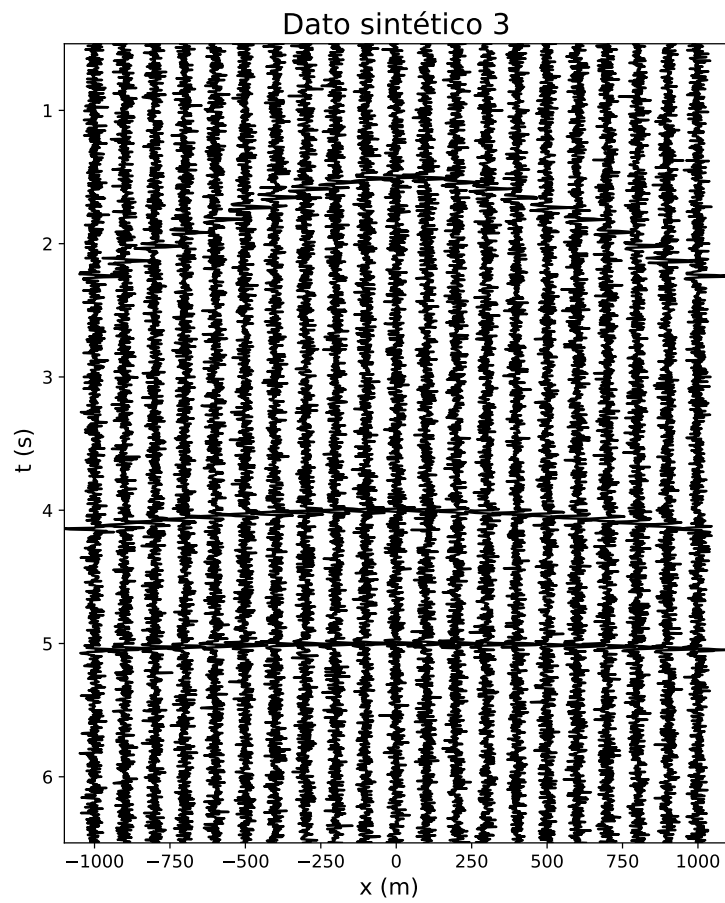
**Figura 4.3:** Registro sintético 2, generado a partir de la información correspondiente a las tablas 4.1 y 4.2 con nivel de ruido a  $snr = 1$ .

Método	$snr$	$Misfit$ Norm.	Cant. de coef. [%]	Tiempo [s]
DLS	1.04	0.59	100	35.42
RHRT	0.49	1.00	0.2	0.87
GRT	0.51	1.00	0.064	1.87
StOMP	0.49	1.00	0.21	2.77
Resultado deseado	0.5	1.00		

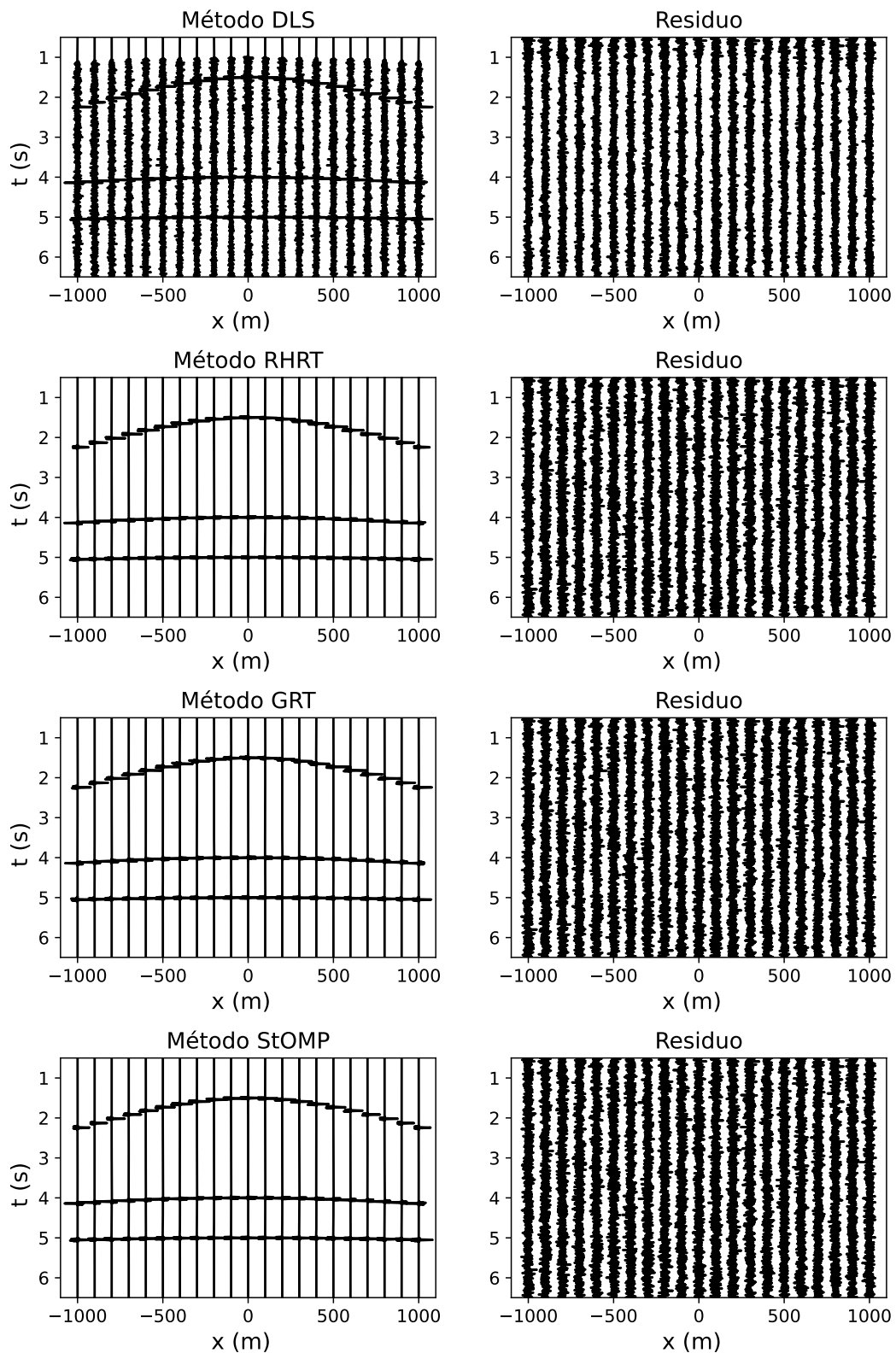
**Tabla 4.5:** Relación señal-ruido, error de ajuste ( $misfit$ ) normalizado, cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 3.



**Figura 4.4:** Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando datos con relación señal-ruido de 1.



**Figura 4.5:** Registro sintético 3, generado a partir de la información correspondiente a las tablas 4.1 y 4.2 con nivel de ruido  $snr = 0.5$ .



**Figura 4.6:** Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando el registro con relación señal-ruido de 0.5.

Tiempo doble (s)	Vel (m/s)	Amplitud
0.65	650	1
1.3	700	-1
2.4	900	1
3.0	1100	1
3.5	1300	-1

**Tabla 4.6:** Parámetros de las cinco hipérbolas de reflexión para el registro sintético 4.

Método	<i>snr</i>	<i>Misfit</i> Norm.	Cant. de coef. [%]	Tiempo [s]
DLS	2.17	0.56	100	32.66
RHRT	1.47	1.02	0.2	0.66
GRT	1.52	1.00	0.13	1.74
StOMP	1.49	1.01	0.2	2.02
Resultado deseado	1.5	1.00		

**Tabla 4.7:** Relación señal-ruido, error de ajuste (*misfit*) normalizado, cantidad de coeficientes utilizados, y tiempo de procesamiento para DLS, RHRT, GRT, y StOMP aplicados al dato sintético 4.

### 4.2.2. Variación de la entrada

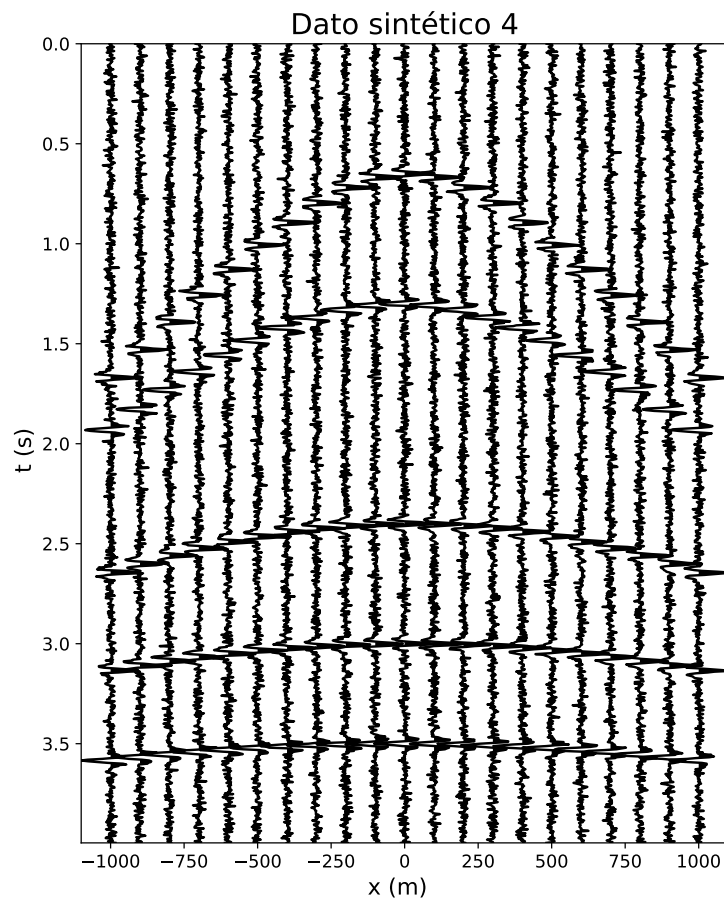
El objetivo es variar el dato de entrada incorporando una mayor cantidad de reflexiones, para luego analizar si los métodos alcanzan resultados óptimos.

El **dato sintético 4** corresponde a los parámetros de la tabla 4.1 y las reflexiones tienen los parámetros presentados en la tabla 4.6. El dato de entrada se muestra en la figura 4.7. Los resultados gráficos para cada uno de los métodos se encuentran en la figura 4.8.

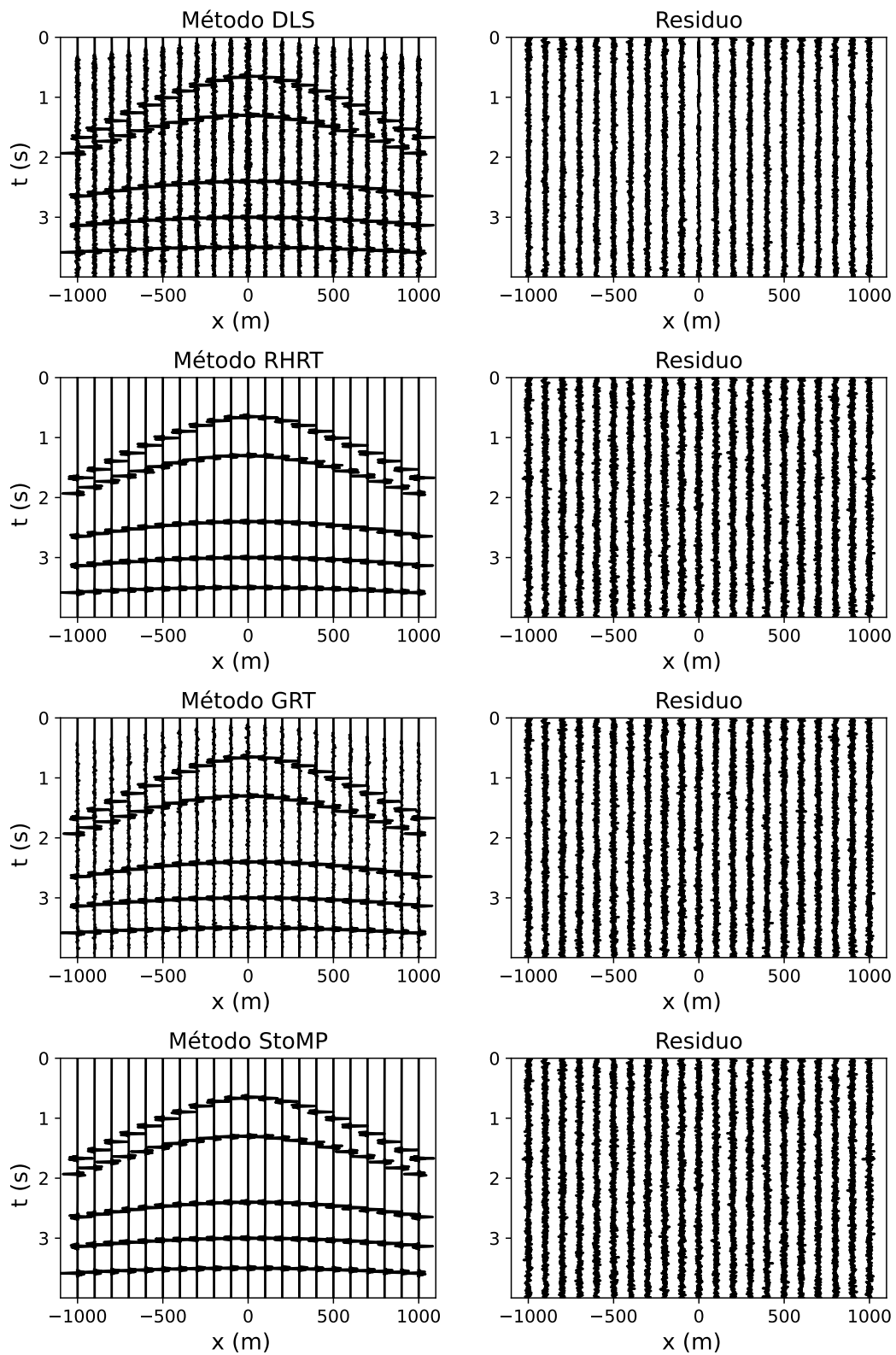
### 4.2.3. Observaciones preliminares

- La observación más importante es que el único algoritmo que no requirió que sus parámetros de entrada sean modificados para una buena reconstrucción del dato es StOMP. Incluso, para el mismo par de parámetros para los cuatro ejemplos, convergió siempre a los valores de error de ajuste (*misfit*) normalizado y *snr* esperados, haciendo mínimo el porcentaje de coeficientes utilizados. Tanto GRT como RHRT,





**Figura 4.7:** Registro sintético 4, generado a partir de la información correspondiente a las tablas 4.1 y 4.6 con ruido correspondiente a  $snr = 1$ .



**Figura 4.8:** Resultados obtenidos a partir de la inversión según DLS, RHRT, GRT, StOMP utilizando el dato sintético 4 relación señal-ruido de 1.

Dato sintético		RHRT	GRT		StOMP	
$N_{ro}$	Característica	Cant. de coef.	$N_{it}$	Valor umbral	$N_{it}$	Valor umbral
1	Sin ruido	2 %	15	0.8	4	7.5
2	$snr = 1$	0.4 %	4	0.65	4	7.5
3	$snr = 0.5$	0.2 %	4	0.8	4	7.5
4	Cinco reflexiones $snr = 1.5$	0.2 %	4	0.65	4	7.5

**Tabla 4.8:** Parámetros de entrada utilizados en cada método para todos los ejemplos.

alcanzaron excelentes resultados, pero a partir de una muy cuidadosa selección de sus parámetros, dado que necesario reajustarlos hasta obtener el resultado esperado. Si dicha acción no se llevaba a cabo, tanto los valores de la relación señal-ruido como el error de ajuste resultaban ser más distantes a los valores esperados. En la tabla 4.8 se muestran los parámetros de entrada utilizados para cada prueba.

- Para todas las pruebas realizadas, la reconstrucción del dato mediante mínimos cuadrados amortiguados (DLS) es aceptable. Sin embargo, al utilizar todos los coeficientes  $\mathbf{m}$  para la reconstrucción (ya que no es un método voraz), el algoritmo intenta reconstruir completamente la información de entrada. Esto lo hace un método lento. Asimismo, además de sintetizar las hipérbolas, también busca sintetizar ruido. Por lo tanto, el error de ajuste (*misfit*) normalizado tiende a ser menor que 1, ya que sobre-ajusta el dato. De todos modos, se espera que el término de regularización mitigue estos problemas.
- **Dato sintético 1 (sin ruido)**: Todos los métodos reconstruyeron bien el dato de entrada. Fue necesario aumentar demasiado el número de iteraciones del GRT para evitar afectar al dato cuando se calcula el residuo. Aún así, no se alcanzó el objetivo y es por esa razón que en su residuo se insinúan las hipérbolas (ver fig. 4.2). También se realizó la prueba de disminuir el valor umbral (*thr*), pero aumentó el nivel de ruido espurio.
- **Dato sintético 2 ( $snr = 1$ )**: La reconstrucción es óptima para todos los algoritmos.
- **Dato sintético 3 ( $snr = 0.5$ )**: La reconstrucción es óptima para todos los algoritmos.

- **Dato sintético 4 (cinco reflexiones)**: La reconstrucción es óptima para todos los algoritmos. No obstante, en el resultado de GRT se insinúan hipérbolas muy leves en el residuo donde no debería haberlas (por encima de la primer reflexión, por debajo de la segunda y por debajo de la tercera, ver fig. 4.8).

#### 4.2.4. Pruebas de sensibilidad

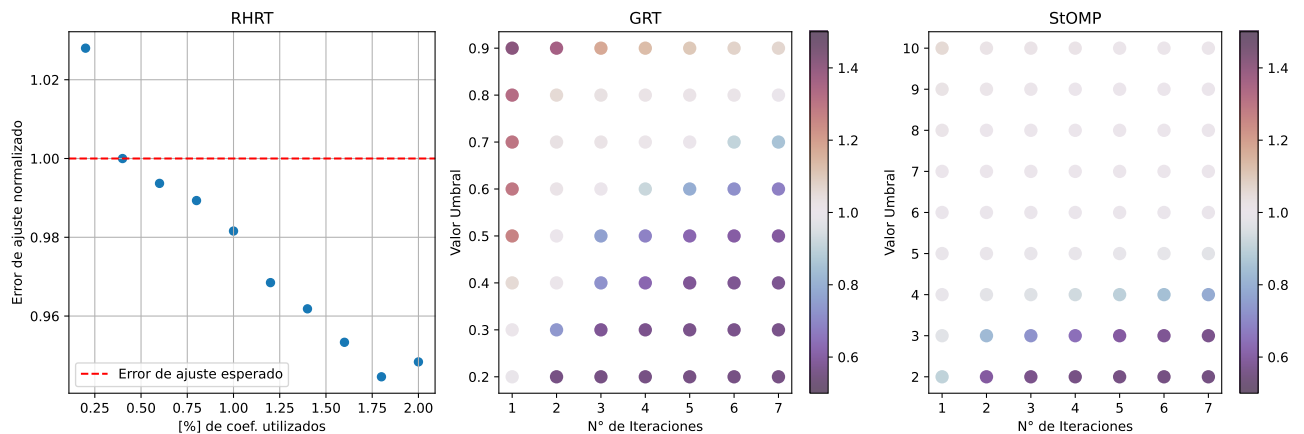
Frecuentemente, en las demostraciones prácticas de diferentes técnicas de procesamiento, los resultados exhibidos son buenos para un determinado conjunto de parámetros únicamente. No obstante, a la hora de aplicarlos variando el dato de entrada, o modificando los parámetros de los algoritmos, los métodos tienden a fallar. Es por esta razón que se decidió realizar una serie de pruebas que consisten en cambiar los valores de dichos parámetros para cada procedimiento, y así comprobar cuán sensibles son. Entendemos como **sensibilidad** a los cambios de los resultados obtenidos en función de variaciones en los parámetros de entrada. Cuanto menos sensible sea un método, se dice que es más **robusto**.

Se compara la sensibilidad de los tres métodos voraces mediante el análisis de:

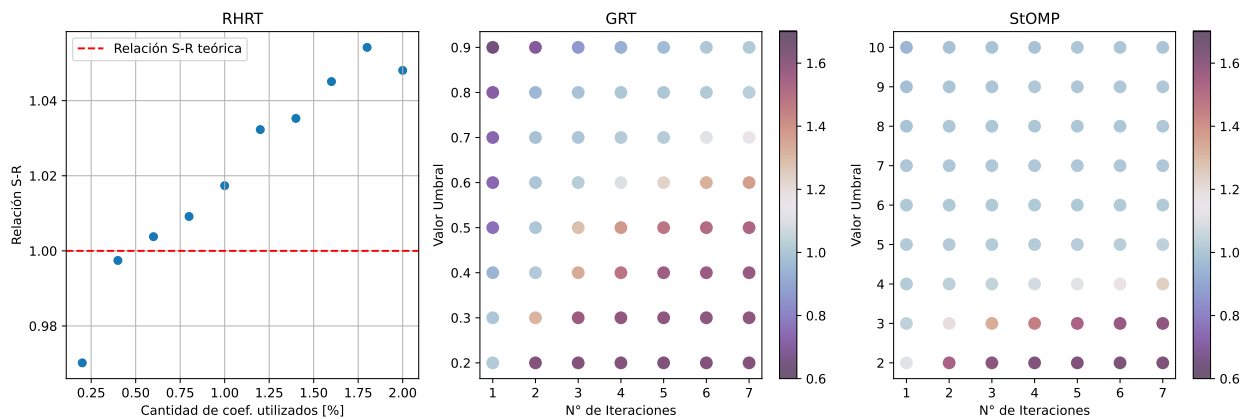
- **Error de ajuste normalizado (*misfit*)** (ver ecs. 4.3 y 4.4).
- **Relación señal-ruido (*snr*)** (ver ec. 4.5).
- **Cantidad de coeficientes utilizados** Este valor es utilizado como criterio de comparación solamente en los algoritmos voraces iterativos, ya que en el caso de RHRT es el parámetro de entrada.

El dato de entrada es un registro sintético igual al de la sección 4.2.1, al que habíamos denominado **dato sintético 2**. En este caso, la relación señal-ruido es 1.

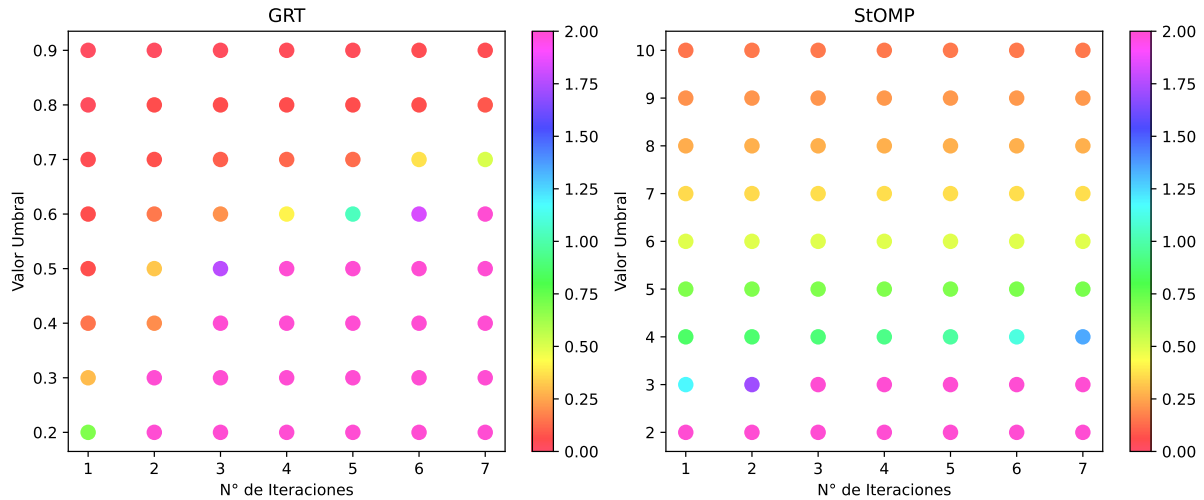
La motivación en este capítulo es la de encontrar cuál es el método más eficaz y cuál es el que mejor representa al dato de entrada a la hora de predecirlo. A partir de un análisis sobre distintos datos de entrada, observamos que los tres algoritmos voraces tienen un



**Figura 4.9: Sensibilidad del error de ajuste normalizado.** A izquierda se muestra el efecto de variar el porcentaje de coeficientes utilizados en RHRT. El eje vertical muestra el  $e$  normalizado. En los paneles central y derecho se observan las variaciones en colores para distintos valores de número de iteraciones ( $N_{it}$ ) y valor umbral ( $thr$ ) de los algoritmos iterativos GRT y StOMP.



**Figura 4.10: Sensibilidad de la relación señal-ruido.** A izquierda se muestra el efecto de variar el porcentaje de coeficientes utilizados en RHRT. El eje vertical muestra el valor del parámetro  $snr$  obtenido. En los paneles central y derecho se observan las variaciones en colores para distintos valores de número de iteraciones ( $N_{it}$ ) y valor umbral ( $thr$ ) de los algoritmos iterativos GRT y StOMP.



**Figura 4.11: Sensibilidad de la cantidad de coeficientes elegidos.** En la imagen izquierda se ve el efecto sobre GRT y a la derecha StOMP. Se observan las variaciones en colores para distintos valores de  $N_{it}$  y  $thr$  de los algoritmos iterativos.

excelente comportamiento, siempre y cuando los parámetros de cada uno de ellos sean bien elegidos. Esto nos inspiró a hacer los análisis de sensibilidad mostrados en la subsección anterior. De allí, concluimos:

- **Error de ajuste normalizado (fig. 4.9):** El método RHRT tiene un muy buen comportamiento. Variando la cantidad de coeficientes seleccionados entre 0.4 y 2 %, se logra que el error de ajuste no varíe más de un 6 % en casos extremos. Considerar 0.2 % de los coeficientes resulta poco representativo, y considerar por encima de 1.5 % hace que tienda a sintetizar ruido. En cuanto a los métodos iterativos, es evidente que para valores mínimos en los parámetros de entrada, tales como  $thr$  bajos, o tan solo una iteración, su comportamiento es pobre. La causa de este fenómeno es que son condiciones muy poco selectivas, por lo que muchos coeficientes logran cumplir el criterio de selección. Por ende, se modelan coeficientes pequeños como hipérbolas. Sin embargo, a medida que aumentan los valores de los parámetros, se observa cómo el error disminuye considerablemente. En el caso de StOMP, para valores de  $thr$  entre 4 y 10, el *misfit* normalizado se estabiliza en 1, el valor esperado. Incluso, casi no varía en función de el número de iteraciones. En cuanto a GRT, si bien la solución se acerca a los valores esperados, lo hace en una región más acotada del gráfico. Tal es así que fue muy difícil ajustar los parámetros apropiados que resuelvan de manera correcta las pruebas anteriores.
- **Relación señal-ruido ( $snr$ ) (fig. 4.10):** El comportamiento es similar al del

ítem anterior. El RHRT muestra un desempeño correcto, variando poco dentro de un rango coherente de coeficientes seleccionados. En referencia a los algoritmos iterativos, se ve cómo StOMP se estabiliza en valores muy cercanos al dato teórico ( $snr = 1$ ) para  $thr$  mayores o iguales a 4. Esto no sucede con GRT, donde resulta más difícil hallar pares de parámetros que se correspondan a dicho valor esperado.

- **Cantidad de coeficientes utilizados (fig. 4.11):** Los algoritmos iterativos toman valores muy grandes para sus mínimos de valor umbral, a punto tal de saturar la escala de colores propuesta. Esto es esperable que suceda, teniendo en cuenta que al elegir un valor umbral muy pequeño, son muchísimos los coeficientes  $\tilde{m}(\tau, v)$  que cumplen con el criterio de selección impuesto. Hecha esa salvedad, ambos algoritmos convergen a una pequeña cantidad de coeficientes, siendo StOMP el método más estable. Excluimos del análisis a RHRT, dado que la cantidad de elementos es su parámetro de entrada.

### 4.3. Conclusiones del capítulo

Los tres métodos arrojaron resultados aceptables para las pruebas con datos sintéticos impuestas, aunque StOMP supera a GRT en cuanto a estabilidad, y eso lo hace considerablemente un mejor algoritmo. Por último, no encontramos causas para afirmar que StOMP sea mejor que RHRT, y viceversa. Es por ello que en el siguiente capítulo se analizará una aplicación en datos reales utilizando ambos métodos con el objetivo de obtener conclusiones prácticas.

# Capítulo 5

## Aplicación a datos reales

Una de las principales aplicaciones de la TR es la de eliminar reflexiones múltiples (Foster y Mosher, 1992). Conceptualmente, la idea es identificar la región del dominio transformado de Radon en la que se encuentran los arribos múltiples (bajas velocidades a tiempos altos). Luego, se predicen las hipérbolas de reflexión asociadas a múltiples y se las resta al dato original. En este capítulo, se utiliza un registro de punto medio común (*common mid point gather*, CMP) correspondiente a sísmica marina del Golfo de México. El dato fue liberado para uso académico (Verschuur y Prein, 1999). Este dato se encuentra fuertemente afectado por reflexiones múltiples, a causa de los fuertes contrastes de impedancias acústicas, tanto del fondo marino como por la presencia de un domo de sal de muy alta velocidad. Por lo tanto, es crucial que en las instancias de procesamiento se elimine su presencia para evitar una interpretación errónea.

### 5.1. Procedimiento

Los pasos a seguir para atenuar las múltiples no deseadas mediante la utilización de la TR son los siguientes:

1. Selección de un dato CMP de entrada.
2. Aplicación del operador adjunto de Radon y selección de coeficientes mediante los métodos elegidos en el capítulo 4 (RHRT y StOMP). Considerando que se busca comparar ambos algoritmos, para ser objetivos se deben equiparar los parámetros de entrada. La manera de hacerlo es aplicando StOMP en primera instancia, seleccionando distintos valores para los parámetros valor umbral y número de iteraciones, hasta encontrar aquellos que mejor modelan el dato. Posteriormente, a partir de



conocer el porcentaje de coeficientes que utiliza StOMP, se ingresa dicho porcentaje como parámetro de entrada en RHRT (porcentaje de coeficientes utilizados). En vías de realizar un análisis más profundo, a lo largo del capítulo se hace la comparación de dos pares distintos de parámetros para StOMP, lo que implica también dos porcentajes de coeficientes distintos para RHRT.

3. Selección manual<sup>1</sup> de la región del dominio de Radon donde se considera que las coordenadas  $(\tau, v)$  suman amplitudes correspondientes a múltiples.
4. Se usa la TR directa sobre dicha región para obtener las múltiples sintetizadas en el dominio  $(t - x)$ .
5. Se restan las múltiples sintetizadas al dato original.

## 5.2. Aplicación

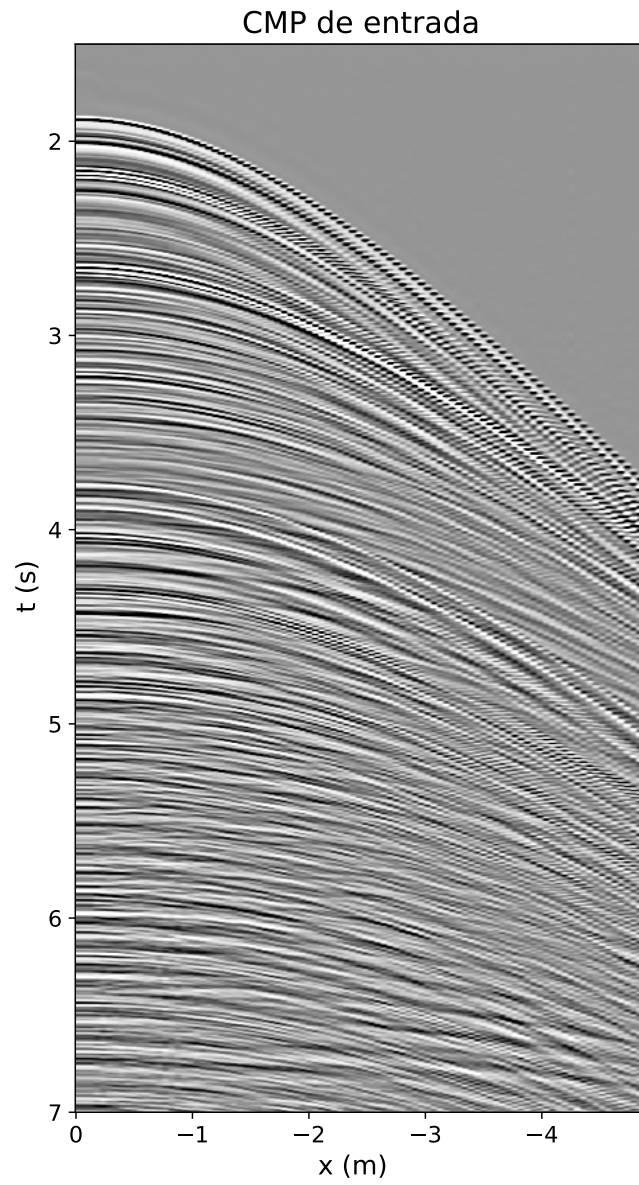
El dato de entrada se muestra en la figura 5.1. Es visible la presencia de eventos asociados a múltiples con pendientes considerablemente distintas para tiempos mayores a 3.5 segundos. Sin embargo, puede resultar complicado reconocer esta energía indeseada a primera vista. Es por ello que las múltiples en el dominio  $(t - x)$  se observan mejor si se corrige al dato por NMO. En este contexto, los eventos primarios quedan con una orientación horizontal y las múltiples conservan cierta pendiente residual, pues son sub-corregidas. En la figura 5.4 mostramos este efecto en una ventana de tiempo acotada, para que se evidencie con mayor detalle.

Según lo mencionado en el paso 2 de la sección anterior, en primera instancia se calcula la TR mediante StOMP. Considerando que el método selecciona las coordenadas iterativamente, se espera obtener distintos conjuntos de coordenadas si se modifica la cantidad de iteraciones y el valor umbral. Se realiza una prueba con una pequeña cantidad de iteraciones y valor umbral bajo, y otra con un mayor número de iteraciones y mayor valor umbral. Para identificarlos, los llamamos StOMP 1 y StOMP 2 respectivamente. Cuanto más grande es el valor umbral, menor es la cantidad de coordenadas que cumplen el criterio de selección (ec. 3.7). Por lo tanto, se necesitará de más iteraciones para alcanzar un porcentaje de coeficientes razonable.

La tabla 5.1 muestra los parámetros utilizados por cada una de las pruebas realizadas

---

<sup>1</sup>Existen criterios y maneras automáticas para hacer la selección de la región en el dominio de Radon asociada a múltiples. Son comúnmente utilizadas para una mayor cantidad de datos.



**Figura 5.1:** CMP de entrada del Golfo de México para el análisis de datos reales.

con StOMP, la cantidad de coeficientes que selecciona al finalizar las operaciones (que luego se utiliza como parámetro de entrada de RHRT en cada caso), y los respectivos tiempos de cómputo.

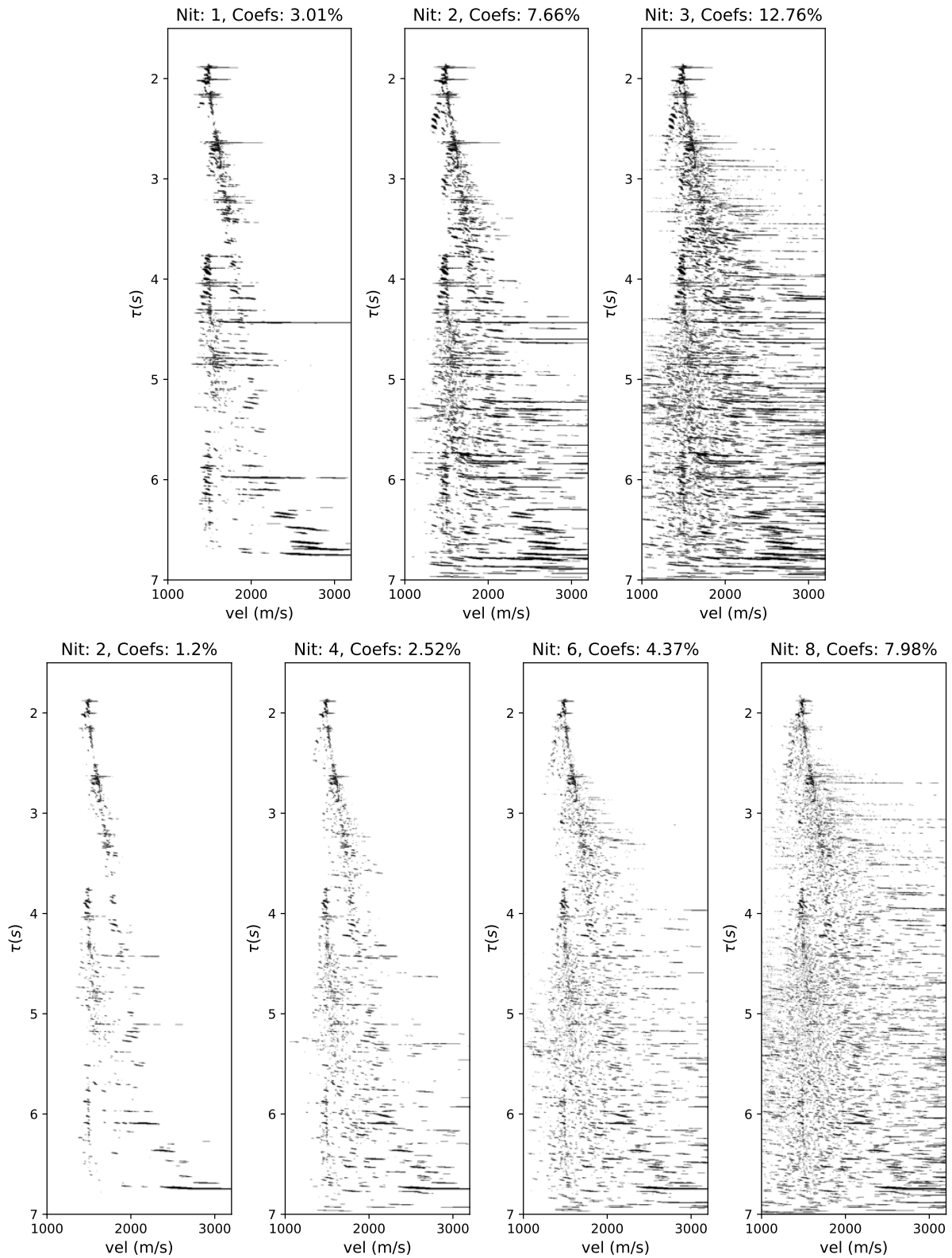
La figura 5.2 muestra donde se ubican en el espacio  $(\tau - v)$  las coordenadas seleccionadas por StOMP 1 y StOMP 2 en cada iteración. Además se indica el porcentaje acumulado de coeficientes utilizados. Este tipo de gráfico se denomina “**mapa binario**” (*bitmap*). Si se realizara un *bitmap* para un método que utiliza todas las coordenadas para realizar la inversión, como por ejemplo DLS, se vería completamente negro.

Método	$thr$	$N_{it}$	Cantidad de coefs.	Tiempo de cómputo [s]
StOMP 1	2.4	3	12.76 %	112.8
RHRT 1			12.76 %	53.6
StOMP 2	3.65	8	7.98 %	165.0
RHRT 2			7.98 %	36.6

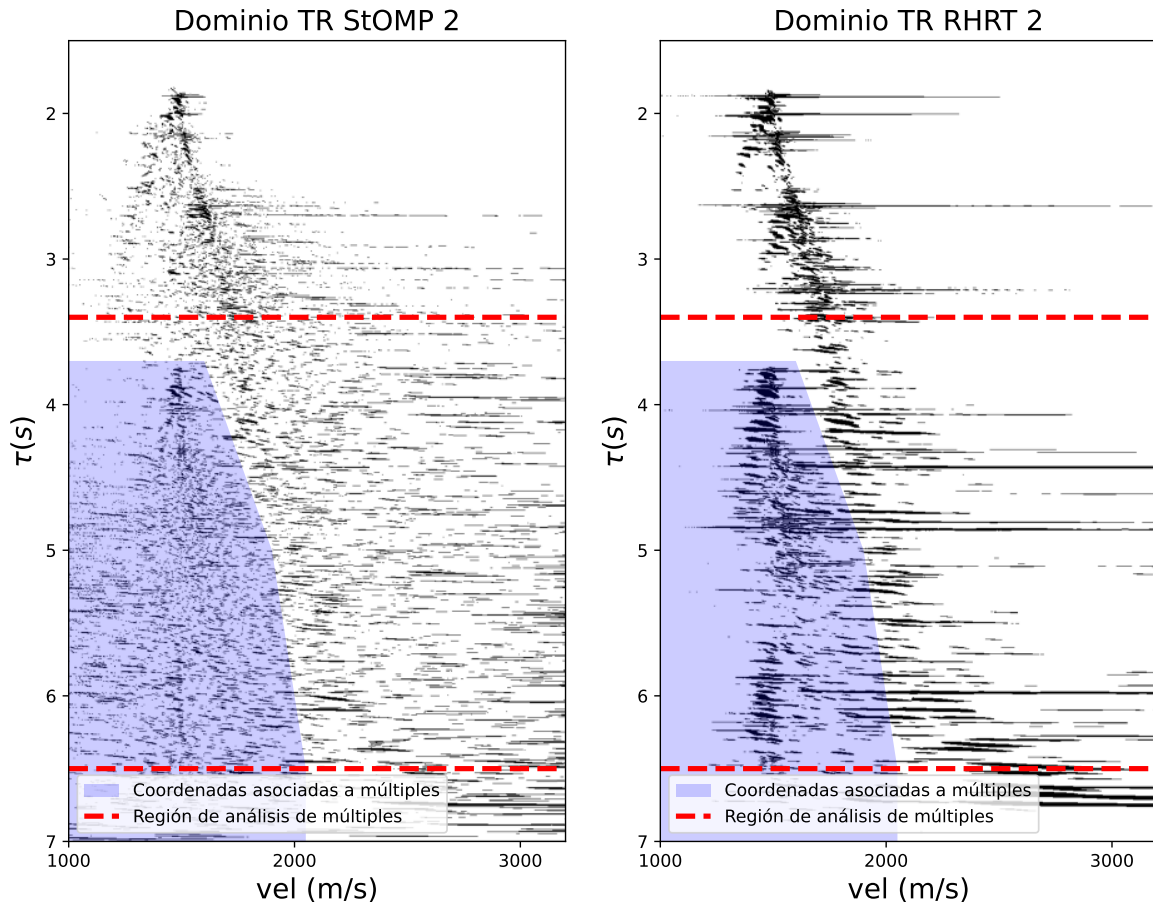
**Tabla 5.1:** Parámetros de entrada, cantidad de coeficientes utilizados, y tiempo de cómputo para las pruebas StOMP 1 y StOMP 2 RHRT 1, StOMP 2, y RHRT 2.

Las primeras conclusiones obtenidas tras la selección de coordenadas en el dominio transformado de Radon mediante StOMP son:

- Cada vez que realiza una iteración, aumenta la cantidad de coeficientes. Es esperable que esto suceda, e indica que el algoritmo funciona. Sin embargo, resulta controversial. Es deseable que el algoritmo en algún momento cese de aumentar la cantidad de coeficientes seleccionados, alcanzando cierta cota superior. En caso contrario, el porcentaje puede aumentar de manera desmedida volviéndose un método ineficaz. Al no estar probada su convergencia, StOMP es un método criticado por la comunidad científica dedicada a *Compressive Sampling* (Needell y Tropp, 2009).
- La primera iteración selecciona las coordenadas  $(\tau, v)$  más relevantes. Es visible la presencia de arribos primarios (mayor velocidad respecto a las múltiples en igual tiempo) y múltiples. Incluso, se ve la clara representación de las múltiples que solo viajan por el agua, cuya velocidad de onda primaria (p) característica es de 1500 m/s.
- El tiempo de cómputo es considerablemente menor para el caso de StOMP 1 respecto de StOMP 2, entendiéndose que esto se debe a su menor cantidad de iteraciones.



**Figura 5.2:** Coordenadas seleccionadas por StOMP en cada iteración. *Fila de arriba:* Caso de StOMP 1. *Fila de abajo:* Caso de StOMP 2, cada 2 iteraciones.



**Figura 5.3:** Coordenadas seleccionadas por StOMP y RHRT en el dominio de Radon. En celeste se muestran las regiones consideradas para atenuar múltiples. Entre líneas punteadas rojas los tiempos en que se analizará el dominio ( $t - x$ ) en detalle

- Hay una diferencia sustancial entre las coordenadas seleccionadas por StOMP 1 y 2. La realizada por el caso con mayor cantidad de iteraciones (StOMP 2) tiene menor dispersión horizontal, entendiendo que disminuye el efecto de borde explicado en el capítulo 2.

Se realiza la selección mediante RHRT para la misma cantidad de coeficientes, tal como se explicó en el paso 2 del procedimiento. Específicamente, se utilizan 12.76 % y 7.98 %. La figura 5.3 compara la selección hecha por StOMP 2 y RHRT 2. Además, en celeste se muestran las regiones que representan los arribos múltiples, y entre líneas punteadas rojas los tiempos del CMP en los que se se visualiza el detalle.

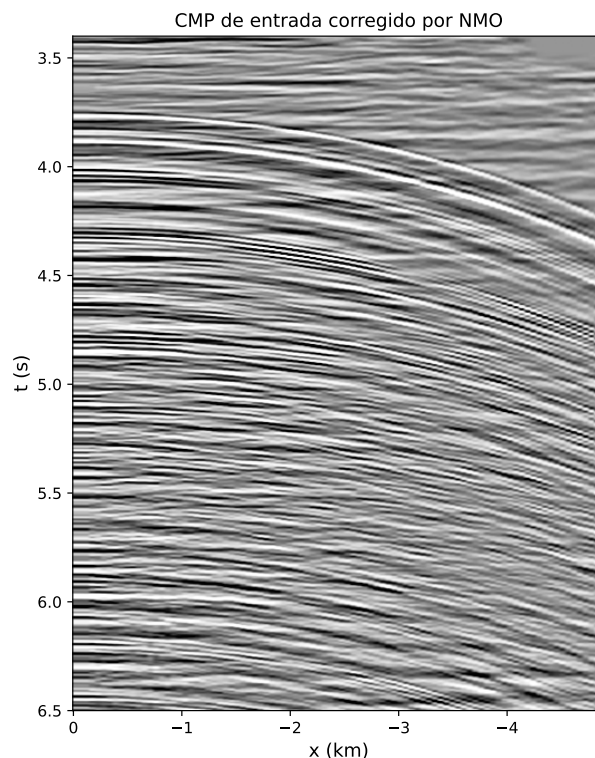
La conclusión más importante es que StOMP logra disminuir enormemente los coeficientes asociados a ruido de borde generados por discretizar de TR. Por ejemplo, a tiempos bajos cercanos a los 2 segundos, RHRT selecciona muchos coeficientes ordenados

de manera horizontal. No es así en el caso de StOMP, donde si bien hay eventos levemente alineados, son considerablemente más cortos. Dicho comportamiento se observa para distintos valores de  $\tau$ . Una conclusión al respecto, es que cuanto más pequeño sea el aumento de coeficientes seleccionados entre una iteración y la siguiente, mejor actuará el criterio de selección. Así, se cumple uno de los objetivos que tiene el método StOMP: sustraer la señal disminuyendo el ruido no deseado. Como consecuencia, se obtienen pocos coeficientes que se ubican aislados entre sí, lo que se denomina una solución rala (*sparse*), que se ajusta a la solución buscada. Esta conclusión toma mayor validez a la hora de comparar StOMP 2 contra StOMP 1 y RHRT, donde el último elige todos los coeficientes en una única acción.

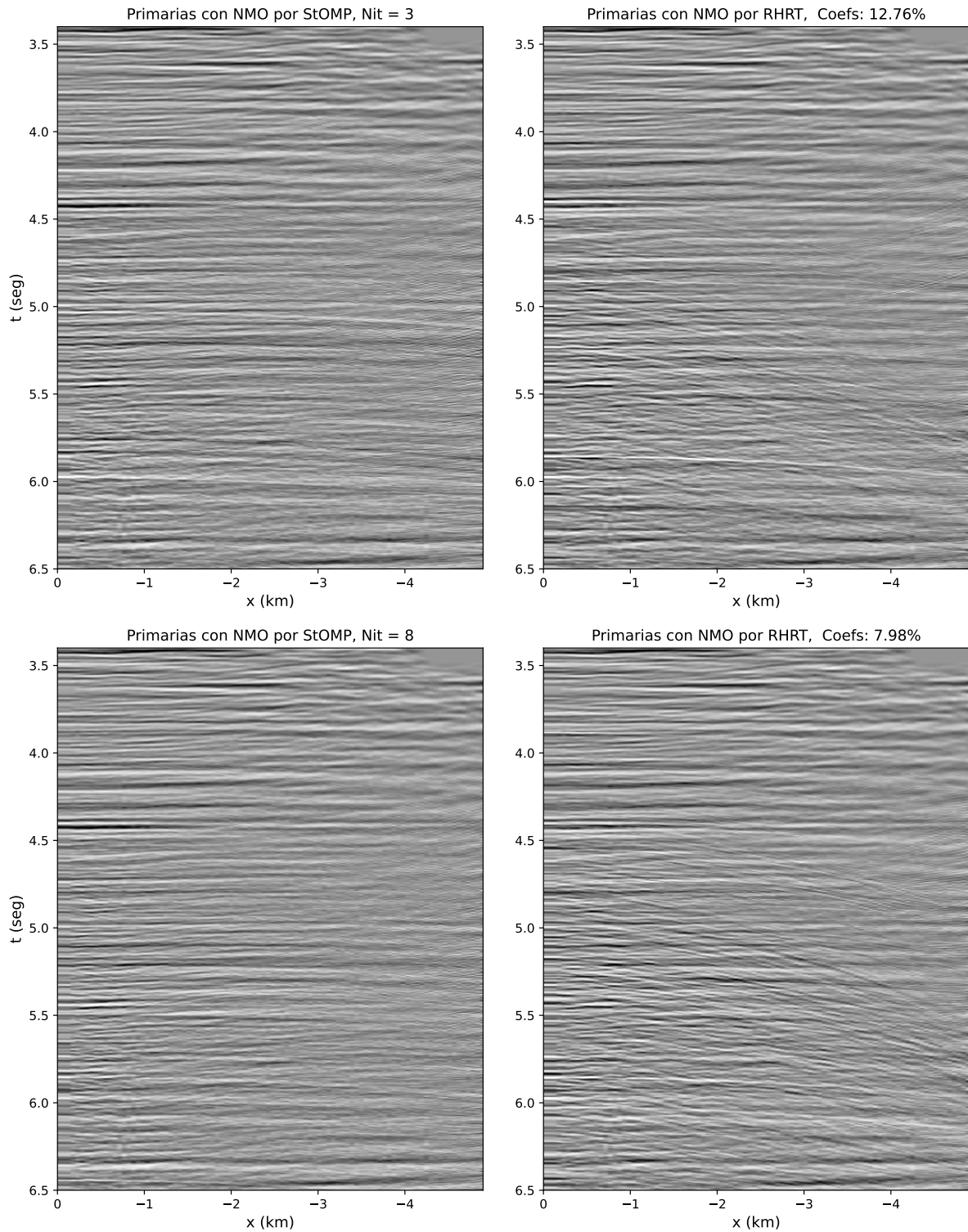
Para concluir con el procedimiento, se antitransforman las coordenadas seleccionadas dentro del área celeste para cada uno de los métodos. Luego, se resta al dato original el modelo de múltiples obtenido.

La figura 5.4 muestra la entrada con NMO aplicado. Las múltiples se presentan con mucha energía, con alta amplitud y curvatura, enmascarando las reflexiones primarias.

Las figura 5.5, representa lo que sucede al eliminar múltiples según StOMP 1 y RHRT 1 (*arriba*) y StOMP 2 y RHRT 2 (*abajo*).

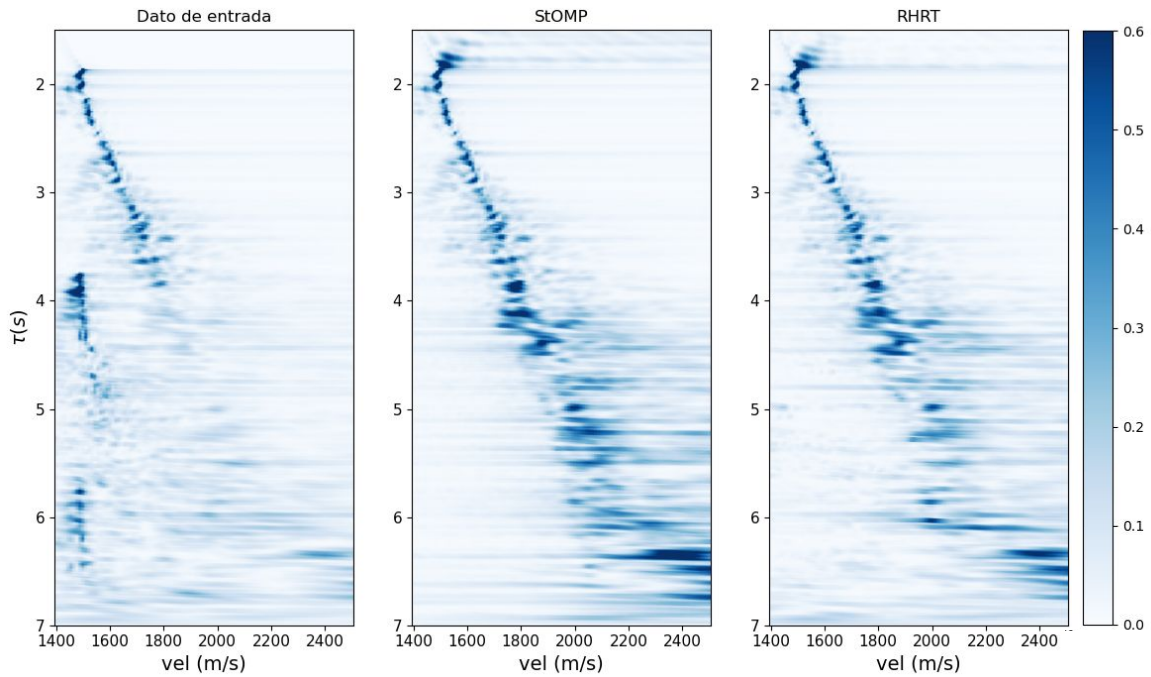


**Figura 5.4:** Dato de entrada con NMO aplicado. Foco entre 3.4 y 6.5 segundos.



**Figura 5.5:** Para la ventana de tiempo entre 3.4 y 6.5 segundos, se representa la atenuación de múltiples mediante StOMP 1 (*arriba izq.*), RHRT 1 (*arriba der.*), StOMP 2 (*abajo izq.*) y RHRT 2 (*abajo der.*).

Los resultados obtenidos son buenos, ya que los cuatro casos logran atenuar las reflexiones múltiples más importantes. Sin embargo, existen diferencias. A partir de los 4.2



**Figura 5.6:** Análisis de semblanza como control de calidad. Dato de entrada (*izq.*), dato con múltiples atenuadas mediante StOMP (*centro*) y mediante RHRT (*der.*)

segundos, y entre los 0 y 4 km de distancia fuente-receptor, utilizando RHRT se ven reflexiones residuales no horizontales que el método no logra mitigar. StOMP, en cambio, las elimina adecuadamente. En el último caso mencionado, apenas se insinúan reflexiones no horizontalizadas. Por ello, se puede concluir que la selección de coeficientes siguiendo los criterios del método StOMP ajustan mejor el dato. Además, como StOMP 2 realiza una selección más criteriosa de las coordenadas, obtiene un mejor resultado que StOMP 1, a pesar de que éste utiliza mayor cantidad de coeficientes. En detalle, se alcanza a ver la atenuación de pequeñas múltiples residuales que persisten en StOMP 1, por ejemplo entre 5.7 y 6.0 segundos, entre 3 y 5 km de distancia fuente receptor (ver figura 5.5).

Se realiza un control de calidad tanto al dato de entrada como a las salidas de los procesos. Éste consiste en hacer un análisis de velocidades a partir de la semblanza (Yilmaz, 2001). Los resultados obtenidos son los esperados y se muestran en la figura 5.6. Se observan las múltiples en la semblanza del dato de entrada para velocidades cercanas a 1500 m/s (velocidad característica del agua salada). Ambos métodos logran atenuar la información correspondiente a dichas velocidades, cumpliendo así su objetivo.

Los tiempos de cómputo son excesivamente mejores para RHRT, siendo menores a la mitad del tiempo que StOMP requiere. La causa de este fenómeno está asociada a que por cada iteración que StOMP realiza, hay involucrada una inversión por gradientes



conjugados. Por lo tanto, si bien es cierto que uno de los objetivos de los métodos voraces (*greedy*) es utilizar una menor cantidad de coeficientes para reducir tiempos, esta situación hace que los tiempos aumenten. Surge una relación de compromiso entre la buena selección de coordenadas y el aumento en el número de iteraciones, ya que ese aumento genera un incremento considerable en los tiempos de cómputo. La selección que mejor predice al dato de entrada se da para el caso de StOMP 2, obteniéndose soluciones de tipo ralas (*sparse*).

# Capítulo 6

## Conclusiones

A lo largo de todo este trabajo incorporamos y pusimos en práctica técnicas innovadoras para una etapa de lo que es el procesamiento de datos sísmicos. Partimos de la transformada de Radon, analizamos su importancia y sus características. A partir de las ventajas que ofrece a la hora de identificar y separar reflexiones en el dato sísmico, sumado a los conceptos del *compressive sampling*, nos preguntamos si era posible la reconstrucción de información sísmica mediante algoritmos voraces. Estos algoritmos, tienen la particularidad de impulsar dicha reconstrucción sin la necesidad de utilizar todos los coeficientes del dominio, en nuestro caso el de Radon. La principal discrepancia entre los algoritmos es el criterio de selección de dichos coeficientes. Se explicó el concepto de soluciones ralas y el por qué tiene sentido utilizarlas, incluso geológicamente. Dentro de los algoritmos voraces, se le dio un marco teórico a la transformada de Radon con dominio restringido (RHRT), y a aquellos que son iterativos. Particularmente, se estudió a los denominados transformada de Radon voraz (*greedy Radon transform*, GRT) y al de búsqueda coincidente por etapas (*stagewise orthogonal matching pursuit*, StOMP).

Los principales objetivos fueron determinar cuál de los métodos ajustaba mejor al dato y cuál era el más eficaz. En la búsqueda, se sometió a cada uno a distintas pruebas de datos sintéticos, siempre comparándolos con un método clásico (no voraz) de características conocidas y convergencia asegurada, el de mínimos cuadrados amortiguados (*damped least squares*, DLS). Las pruebas consistieron en variar la relación señal-ruido del dato sintético, cambiar la cantidad de reflexiones de éste, y analizar la sensibilidad de los modelos resultantes en función de los parámetros de entrada. En este sentido, observamos que el rendimiento del método GRT es inferior al de RHRT y StOMP. Esto se debe a que, si bien permite reconstruir de buena manera el dato, resulta muy difícil hallar los parámetros que lo hicieran correctamente. En contrapunto, RHRT y StOMP demostraron

---

ser robustos.

Finalmente, investigamos las diferencias entre los rendimientos de RHRT y StOMP utilizándolos en una de las principales aplicaciones que tiene la transformada de Radon: eliminar múltiples de un dato real. A raíz de los resultados, vimos que StOMP reconstruye de mejor manera el dato de entrada. Más aún, su modelo sirvió para atenuar acertadamente las reflexiones múltiples, dejando residuos casi imperceptibles, a diferencia de RHRT. Esa conclusión es una de las más importantes del trabajo. Además, notamos que cuanto menor sea el aumento de coordenadas seleccionadas entre una iteración y la siguiente, menor es el ruido a causa de los efectos de borde característicos de la transformada de Radon. Por último, teniendo en cuenta que RHRT resuelve la inversión de parámetros y modela el dato de entrada una única vez (siendo no iterativo), es considerablemente más rápido. Cuanto mayor sea el número de iteraciones de los algoritmos que las requieren, mayor será su costo computacional. De todos modos, los tiempos de cómputo de la TR hiperbólica disminuyen de gran manera utilizando cualquiera de los métodos voraces analizados en este trabajo. Resultan considerablemente más rápidos que el método de mínimos cuadrados amortiguados.

Queda pendiente para trabajos futuros la implementación de los métodos voraces (*greedy*) iterativos en diversas aplicaciones típicas de la transformada de Radon, como por ejemplo interpolación trazas sísmicas o eliminación de ruido microsísmico. Incluso, éstos se podrían aplicar en diversas transformadas de Radon (lineal, parabólica, etc.). Por otro lado, la utilización de los métodos *greedy* en transformadas diferentes a la de Radon despierta interés, principalmente en aquellas que se utilizan para reconstrucción N-dimensional. Una de ellas es la *anti-leakage Fourier transform* (ALFT) (Xu et al., 2005). Parte de su procedimiento implica seleccionar el coeficiente más grande del dominio de Fourier, y así reconstruir el dominio transformado. Entonces, es un escenario propicio para probar los criterios de selección de los métodos voraces aquí expuestos. Otra forma de interpolar trazas es mediante la *minimum-weighted norm interpolation* (MWNI) (Liu y Sacchi, 2004), la cual requiere invertir datos. En la búsqueda del aumento de la eficiencia podrían aplicarse los métodos voraces. A su vez, cabe mencionar que existen avances sobre las técnicas del *matching pursuit* en reconstrucción de trazas (Vassallo et al., 2010), de modo que los métodos *greedy* podrían utilizarse con en esta aplicación. Finalmente, en cuanto a los métodos *greedy* analizados en este Tesis, se podrían evaluar restricciones en el algoritmo StOMP para el caso que suceda un aumento brusco de coeficientes entre dos iteraciones consecutivas, y así disminuir automáticamente los efectos de borde.

# Apéndice A

## Abreviaturas

TR	Transformada de Radon
DLS	Mínimos cuadrados amortiguados
RHRT	TR hiperbólica con dominio restricto
MP	<i>Matching pursuit</i>
OMP	<i>Orthogonal matching pursuit</i>
GRT	<i>Matching pursuit</i>
StOMP	Búsqueda coincidente por etapas
NMO	<i>Normal moveout</i>
CMP	<i>Common mid-point gather</i>
AVO	<i>Amplitude variation with offset</i>
FWI	<i>Full wave inversion</i>
TF	Transformada de Fourier
L2	Norma 2 o Norma de Cauchy
CG	Gradientes conjugados
CGLS	<i>Conjugate Gradient Least-Squares</i>
ALTF	<i>Anti-Leakeage Fourier transform</i>
MWNI	<i>Minium-weighted norm interpolation</i>

# Apéndice B

## Símbolos

$x$	Distancia fuente-receptor ( <i>offset</i> )
$t$	Tiempo
$v$	Velocidad de propagación
$\tau$	Tiempo de ida y vuelta
$(t - x)$	Dominio de tiempo y espacio
$(\tau - v)$	Dominio de Radon
$(t, x)$	Coordenada en el dominio del tiempo y espacio
$d(t, x)$	Coefficiente en el dominio de los datos
$(\tau, v)$	Coordenada en el dominio de Radon
$m_{adj}(\tau, v)$	Coefficiente adjunto en el dominio de Radon hiperbólico
<b>d</b>	Vector de los datos $d(t, x)$
<b>d<sub>s</sub></b>	Síntesis del dato $d$
<b>d<sub>res</sub></b>	Residuo del dato sintetizado
<b>d<sub>limpio</sub></b>	Registro sintético sin ruido adicionado
<b>d<sub>ruidoso</sub></b>	Registro sintético con ruido adicionado
<b>ruido</b>	Ruido adicionado al dato
<b>r<sub>g</sub></b>	Matriz de ruido gaussiano
<b>m</b>	Vector de coeficientes en el dominio de Radon
<b>m<sub>adj</sub></b>	Vector de los coeficientes adjuntos $m_{adj}(\tau, v)$
<b>m̃</b>	Vector solución que minimiza la función de costo $J$
$p$	Pendiente de una recta
$q$	Curvatura
$N_g$	Cantidad de muestras de tipo $g$
$\omega$	Frecuencia angular digital

---

$M_{adj}(\omega, q_j)$	TF de un coeficiente adjunto
$D(\omega, x_k)$	TF del dato $d$
$\mathbf{M}_{adj}(\omega)$	Vector TF de los coeficientes adjuntos
$\mathbf{D}(\omega)$	Vector TF del dato
$\mathbf{L}^H$	Matriz del operador adjunto
$\mathbf{L}$	Matriz del operador directo
$J$	Función de costo
$\mu$	Parámetro de preblanqueo
$\Phi(\mathbf{m})$	Función de pesos
$\mathbf{W}$	Matriz de pesos
$e$	Error de ajuste ( <i>misfit</i> )
$\ \cdot\ _2$	Norma L2
$T$	Valor umbral (RHRT)
$thr$	Valor umbral (métodos iterativos)
$N_{it}$	Número de iteraciones
$\mathbb{A}$	Subconjunto de coordenadas en Radon
$\sigma$	Nivel de ruido gaussiano (desviación estándar)
$\mathbf{I}_N$	Matriz identidad de dimensión $N$
$snr$	Relación señal-ruido

# Bibliografía

- Beck, A., y M. Teboulle, 2009, A fast iterative shrinkage-thresholding algorithm for linear inverse problems: *SIAM journal on imaging sciences*, **2**, 183–202.
- Beylkin, G., 1987, Discrete radon transform: *IEEE transactions on acoustics, speech, and signal processing*, **35**, 162–172.
- Donoho, D. L., Y. Tsaig, I. Drori, y J.-L. Starck, 2012, Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit: *IEEE transactions on Information Theory*, **58**, 1094–1121.
- Foster, D. J., y C. C. Mosher, 1992, Suppression of multiple reflections using the radon transform: *Geophysics*, **57**, 386–395.
- Gholami, A., y M. Farshad, 2019, Fast hyperbolic radon transform using chirp-z transform: *Digital Signal Processing*, **87**, 34–42.
- Hampson, D., 1986, Inverse velocity stacking for multiple elimination, *in* *SEG Technical Program Expanded Abstracts 1986: Society of Exploration Geophysicists*, 422–424.
- Hargreaves, N., B. verWest, R. Wombell, y D. Trad, 2003, Multiple attenuation using an apex-shifted radon transform: Presented at the 2003 SEG Annual Meeting, OnePetro.
- Hestenes, M., y E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems: *Journal of Research of the National Bureau of Standards*, **49**, 409–436.
- Kabir, M. N., y D. Verschuur, 1995, Restoration of missing offsets by parabolic radon transform 1: *Geophysical Prospecting*, **43**, 347–368.
- Levenberg, K., 1944, A method for the solution of certain non-linear problems in least squares: *Quarterly of applied mathematics*, **2**, 164–168.
- Li, J., y M. D. Sacchi, 2021, An lp-space matching pursuit algorithm and its application to robust seismic data denoising via time-domain radon transforms: *Geophysics*, **86**, V171–V183.
- Lin, R., y M. D. Sacchi, 2020, Separation of simultaneous sources acquired with a high blending factor via coherence pass robust radon operators: *Geophysics*, **85**, V269–V282.
- Liu, B., y M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records:

- Geophysics, **69**, 1560–1568.
- Lu, W., 2013, An accelerated sparse time-invariant radon transform in the mixed frequency-time domain based on iterative 2d model shrinkage: *Geophysics*, **78**, V147–V155.
- Ma, Y., Y. Luo, y P. Kelamis, 2019, Superresolution stacking based on sparse radon transformsuperresolution stack: *Geophysics*, **84**, V45–V54.
- Mallat, S. G., y Z. Zhang, 1993, Matching pursuits with time-frequency dictionaries: *IEEE Transactions on signal processing*, **41**, 3397–3415.
- Needell, D., 2009, *Topics in compressed sensing*: University of California, Davis.
- Needell, D., y J. A. Tropp, 2009, Cosamp: Iterative signal recovery from incomplete and inaccurate samples: *Applied and computational harmonic analysis*, **26**, 301–321.
- Pati, Y. C., R. Rezaifar, y P. S. Krishnaprasad, 1993, Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition: *Proceedings of 27th Asilomar conference on signals, systems and computers*, IEEE, 40–44.
- Sabbione, J. I., y M. D. Sacchi, 2016, Restricted model domain time radon transforms: *Geophysics*, **81**, A17–A21.
- , 2017, Attenuating multiples with the restricted domain hyperbolic radon transform: 15th International Congress of the Brazilian Geophysical Society & EXPOGEF, Rio de Janeiro, Brazil, 31 July-3 August 2017, Brazilian Geophysical Society, 603–608.
- Sabbione, J. I., M. D. Sacchi, y D. R. Velis, 2015, Radon transform-based microseismic event detection and signal-to-noise ratio enhancement: *Journal of Applied Geophysics*, **113**, 51–63.
- Sacchi, M. D., y T. J. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: *Geophysics*, **60**, 1169–1177.
- Thorson, J. R., y J. F. Claerbout, 1985, Velocity-stack and slant-stack stochastic inversion: *Geophysics*, **50**, 2727–2741.
- Trad, D. O., T. J. Ulrych, y M. D. Sacchi, 2002, Accurate interpolation with high-resolution time-variant radon transforms: *Geophysics*, **67**, 644–656.
- Tropp, J. A., y A. C. Gilbert, 2007, Signal recovery from random measurements via orthogonal matching pursuit: *IEEE Transactions on information theory*, **53**, 4655–4666.
- Udias, A., 1997, *Fundamentos de geofísica*: Alianza Editorial.
- Vassallo, M., A. Özbek, K. Özdemir, y K. Eggenberger, 2010, Crossline wavefield reconstruction from multicomponent streamer data: Part 1—multichannel interpolation by matching pursuit (mimap) using pressure and its crossline gradient: *Geophysics*, **75**, WB53–WB67.



- Velis, D., 2022, Apuntes de clase, métodos sísmicos de prospección: Facultad de Ciencias astronómicas y geofísicas, UNLP.
- Verschuur, D., y R. Prein, 1999, Multiple removal results from delft university: The Leading Edge, **18**, 86–91.
- Wang, J., M. Ng, y M. Perz, 2009, Fast high-resolution radon transforms by greedy least-squares method, *in* SEG Technical Program Expanded Abstracts 2009: Society of Exploration Geophysicists, 3128–3132.
- Xu, S., Y. Zhang, D. Pham, y G. Lambaré, 2005, Antileakage fourier transform for seismic data regularization: Geophysics, **70**, V87–V95.
- Yilmaz, Ö., 2001, Seismic data analysis: Processing, inversion, and interpretation of seismic data: Society of exploration geophysicists.