Delft University of Technology

Closed-Loop Surface Related Multiple Estimation

Lopez Angarita, Gabriel

DOI
10.4233/uuid:99e0af9f-067e-4e67-8c81-d5beca0aeb18

Publication date
2016

Document Version
Final published version

Citation (APA)
Lopez Angarita, G. (2016). *Closed-Loop Surface Related Multiple Estimation*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:99e0af9f-067e-4e67-8c81-d5beca0aeb18

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Closed-Loop Surface Related Multiple Estimation

Closed-Loop Surface Related Multiple Estimation

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op dinsdag 15 november 2016 om 12.30 uur

door

Gabriel Armando LOPEZ

Master of Science, Physics
Los Andes University, Colombia
geboren te Bogota, Colombia

Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. ir. N. de Jong

en de copromotor:
Dr. ir. D.J. Verschuur

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. N. de Jong,	Technische Universiteit Delft
Dr. ir. D.J. Verschuur,	Technische Universiteit Delft

Onafhankelijke leden:

Dr. ir. R.G. van Borselen	Aramco Overseas Comp.
Prof. dr. J.O.A. Robertsson	ETH Zürich
Prof. dr.ir. E.C. Slob	CiTG, Technische Universiteit Delft
Prof. dr. A. Ramirez	U-Santander
Prof. dr. ir. A.J. van der Veen	EWI, Technische Universiteit Delft

ISBN 978-94-6233-465-6

Copyright ©2016, by G.A. Lopez, Laboratory of Imaging physics, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands.
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the author G.A. Lopez, Faculty of Applied Sciences, Delft University of Technology, P.O. Box 5046, 2600 GA, Delft, The Netherlands.

SUPPORT
The research for this thesis was financially supported by the Delphi consortium.

Typesetting system: LaTeX.

Printed in The Netherlands by Gildeprint.

Dedicado a mis padres Luis y Julia,
por su apoyo incondicional e infinito amor

Contents

Summary

In seismic acquisition a source wavefield is fired at or close by the surface and its response is then recorded by receivers usually located also at or close to the surface. The reflection response can be decomposed in three different types of wavefields, namely, primaries, internal multiples and surface multiples. Primaries are wavefields with only one bouncing point in the subsurface. Internal multiples are wavefields with at least one downward bouncing point in the subsurface and can be regarded as primaries in many processing schemes. Surface multiples are wavefields with at least one downward bouncing point at the surface and generally have strong amplitudes in the recordings. These strong amplitudes tend to hinder structurally important information, therefore, surface multiples are commonly regarded as noise and a large effort is then put in surface multiple prediction and subtraction.

In this thesis we present a new method for surface multiple estimation which might be regarded as an inversion-oriented version of the well-known Surface-Related Multiple Elimination (SRME) method. The proposed algorithm (known as Closed-Loop SRME, or CL-SRME) presents practical advantages over its predecessor, the most important one being the ability to easily allow built-in extensions, making it possible to use multiples, rather than to eliminate them.

In this research we will be specially interested in performing multiple separation (i.e. separating multiples from primaries) given incomplete datasets. This means that we will concentrate on extending the CL-SRME method to include data reconstruction. This is particularly useful as most of the acquisition techniques in the industry nowadays involve the recording of incomplete datasets due to coarse sampling. Coarse sampling is motivated by economic constraints.

Our interest in reconstruction will lead us to the implementation of the focal transform in the CL-SRME scheme. The focal transform is known to have reconstructive properties when applied to seismic data. This extension will allow CL-SRME to reconstruct over relatively large acquisition gaps. The information in the multiples will prove to be very useful when reconstructing the near-offset section of shallow reflection events.

The algorithm capabilities in terms of reconstruction and multiple separation is presented via synthetic and field data examples in both the 2D and 3D cases. Promising results are obtained in terms of near-offset reconstruction with an important application in shallow water scenarios. The current method is recommended for those cases. Other multiple prediction algorithms are recommended whenever the data is complete or the reconstruction effort is small.

Samenvatting

Bij seismisch onderzoek wordt een vanuit een bron een golf op of dicht bij het oppervlak afgevuurd. De weerkaatsingen van deze golf wordt door ontvangers dichtbij of op het oppervlak waargenomen. De waargenomen golven kunnen bestaan uit drie verschillende soort aankomsten, namelijk primaire reflecties, interne meervoudig weerkaatste golven en oppervlakte gerelateerde meervoudige reflecties. Primaire reflecties zijn de golven die nadat ze n keer weerkaatst zijn in de ondergrond aan het oppervlak opgemeten worden. Interne weerkaatste meervoudige reflecties zijn de golven die tussen de verschillende lagen weerkaatst zijn. Deze golven hebben n of meerdere neerwaartse weerkaatsingen in de ondergrond voordat ze bij de ontvanger terecht komen. Door de meeste verwerkingsprogramma's kunnen deze golven hetzelfde als primaire reflecties beschouwd worden. De oppervlakte gerelateerde meervoudige reflecties hebben minstens n neerwaartse weerkaatsing aan het aardoppervlak. Deze laatste categorie geven over het algemeen een hoge amplitude uitslag in de metingen. Deze hoge amplituden zijn hinderlijk bij het interpreteren van de metingen en worden over het algemeen als ruis beschouwd. Er wordt daarom veel moeite gedaan om deze golven te voorspellen en uit de resultaten te filteren.

In dit proefschrift wordt een nieuwe methode beschreven om oppervlakte gerelateerde meervoudige reflecties te voorspellen. Deze methode kan gezien worden als een inversie georinteerde versie van de bekende 'Surface-Related Multiple Elimination' (SRME) methode. Het voorgestelde algoritme, bekend als 'Closed-Loop SRME' (CL-SRME), heeft praktische voordelen ten opzichte van zijn voorgangers. Het belangrijkste voordeel is dat dit model gemakkelijk om kan gaan met uitbreidingen. Hierdoor wordt het mogelijk de meervoudige reflecties te gebruiken in plaats van ze te elimineren.

In dit onderzoek ligt het focus voornamelijk op het uitvoeren van de separatie van de

verschillende soorten reflecties (de primaire scheiden van de meervoudige reflecties), met incomplete datasets als input. We concentreren ons op het uitbreiden van de CL-SRME methode met data reconstructie. Dit komt goed van pas omdat de meeste gebruikte data sets bestaat uit data die verkregen is via spatieel vrij grof bemonsterde metingen. De metingen zijn over het algemeen grof bemonsterd omdat het te duur is om meer metingen te verrichten.

Om deze data reconstructie te bewerkstelligen wordt de zogenaamde focale transformatie aan het CL-SRME schema toegevoegd. Deze focale transformatie maakt het mogelijk de seismische data te reconstrueren. Hierdoor worden de relatieve grote gaten in de data opgevuld. Deze reconstructie van de data met kleine bron-ontvanger afstanden is vooral nuttig bij het voorspellen van de meervoudige reflecties met kleine bron-ontvanger afstanden.

De mogelijkheden van het algoritme op het gebied van reconstructie en scheiding van golven wordt gedemonstreerd voor zowel gesimuleerde als velddata in zowel 2D als 3D. Veelbelovende resultaten worden behaald voor de reconstructie van metingen met kleine bron-ontvanger afstanden, speciaal voor ondiep water scenario's. Deze methode wordt vooral voor deze scenario's aanbevolen. Voor complete of bijna complete datasets kan men beter andere (goedkopere) algoritmen gebruiken.

Sinopsis

Durante la adquisición sísmica, un campo de onda es disparado en o cerca de la superficie y su respuesta es medida en recibidores colocados también en o cerca de la superficie. La reflexión de respuesta se puede descomponer en tres tipos de campos de onda: las primarias, los múltiplos internos y los múltiplos superficiales. Las primarias son campos de onda con solo un punto de reflexión en el subsuelo. Los múltiplos internos son campos de onda con por lo menos un punto de reflexión en el subsuelo (orientado hacia abajo) y pueden ser vistos como primarias en muchos esquemas de procesamiento. Los múltiplos superficiales son campos de onda con por lo menos un punto de reflexión en la superficie y generalmente tienen amplitudes fuertes en las mediciones. Los múltiplos superficiales tienden a esconder información estructural importante, por lo tanto, son comúnmente vistos como ruido y se realiza un gran esfuerzo en su predicción y eliminación.

En esta tesis presentamos un nuevo método de estimación de múltiplos superficiales, el cual puede ser visto como una versión "orientada a la inversión" del método de Eliminación de Múltiplos Relacionados a la Superficie (o SRME por sus siglas en inglés). El algoritmo propuesto (conocido como "SRME de Circuito Cerrado", o CL-SRME por sus siglas en inglés), presenta ventajas practicas sobre su predecesor, siendo la más importante de ellas la posibilidad de admitir extensiones internas. Estas extensiones hacen posible el uso de los múltiplos, en lugar de su eliminación.

En esta investigación estaremos especialmente interesados en desarrollar separación de múltiplos (esto es, separar múltiplos de primarias) dados volúmenes de datos incompletos. Esto implica que nos concentraremos en extender el método CL-SRME para incluir reconstrucción de datos. Esto es útil ya que la mayoría de técnicas de adquisición en la industria hoy en día involucran medidas de volúmenes de datos incompletas debido a la adquisición truncada. La adquisición truncada está motivada

por ligaduras económicas.

Nuestro interés en la reconstrucción nos llevara a la implementación de la transformada focal en CL-SRME. La transformada focal es conocida por tener propiedades reconstructivas cuando se aplica a datos sísmicos. Esta extensión le permitirá al método CL-SRME la reconstrucción sobre brechas de adquisición relativamente grandes. La información en los múltiplos probará ser de gran utilidad al momento de reconstruir las secciones cercanas de los eventos de reflexión superficiales.

Las capacidades del algoritmo en términos de reconstrucción y separación de múltiplos son presentadas por medio de ejemplos sintéticos y de campo, en casos 2D y 3D. Obtenemos resultados promisorios en términos de reconstrucción de las secciones cercanas con una importante aplicación en escenarios de aguas superficiales. El método actual es recomendado para estos casos. Se recomiendan otros algoritmos de predicción de múltiplos cuando los datos están completos, o cuando el esfuerzo de reconstrucción es menor.

1

Introduction

1.1 The seismic method

In the industry, seismic exploration is used as a method to image the subsurface of the earth in order to obtain its structural details, information that will later lead to the finding or monitoring of e.g. potential hydrocarbon reservoirs to exploit commercially. The first step in seismic exploration is acquisition. Here, sources and receivers of acoustic/elastic wavefields are deployed in a grid configuration (see figure 1.1). Firing the sources generates a wavefield that propagates through the medium. When reaching material discontinuities such as velocity and/or density contrasts, the incoming wavefield gets reflected back to the surface, where it is recorded by the receivers as a set of complicated interfering events.

The key idea behind seismic exploration is that, given the proper understanding of the physics of wavefield propagation, scientists are able to translate the complicated events in the seismic recordings into a structural image of the subsurface. This process is called *seismic imaging*, and typically involves back-propagating the wavefield recorded in the receivers into the subsurface, focusing it in the points where it scattered from.

Most of today's processing and imaging algorithms are based on data that have been reflected only once in the subsurface. In practice, however, the injected wavefield may reflect/transmit many times before reaching the receivers. Reverberations due to reflections between the free surface and the first layer are typically strong in marine data. In land data, internal reverberations between the layers of the earth can also contribute noticeably to the energy of the recorded signal. Multiple reflection events need to be separated from the recorded signal if accurate imaging results are desired.

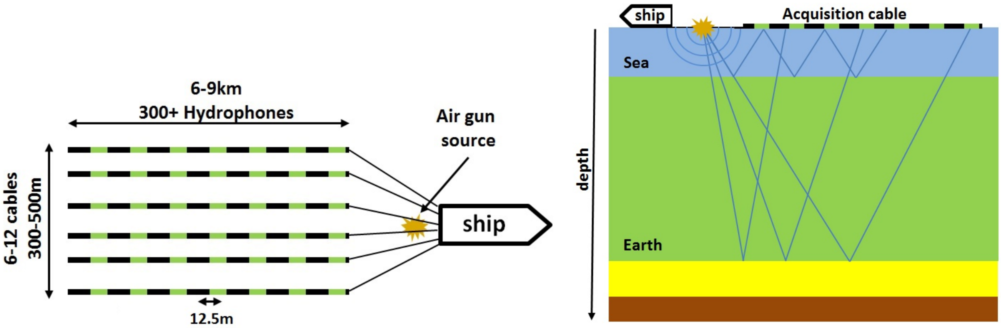

Figure 1.1: *Schematic representation of the data acquisition method for marine environments. The vessel drags a set of cables containing hydrophones, together with an airgun source. The source array generates a wavefield that is later recorded by the hydrophones. The recordings capture the refleclected wavefield due to the contrasts in the medium elastic parameters.*

1.2 Primaries and multiples

Given that the data contains events with multiple reflection points, it becomes necessary to separate the wavefields with one bouncing point (primaries) from the data with multiple reflection points (multiples).

We might distinguish two types of multiples:

- **Surface-related multiples** are those multiples with secondary reflection points at the surface. These multiples are generated from the air-water or air-earth discontinuity and would disappear if such a boundary would become transparent. Generally they present strong amplitudes, as the reflectivity of such interface is close to -1.
- **Internal multiples** are those multiples with downward reflection points within the subsurface. They are generated from the material discontinuities inside the earth and would disappear if their associated generators would become transparent. They generally present low amplitudes, unless they are generated between two strongly reflecting boundaries.

1.3 Multiple elimination in the literature

In this research we will concentrate on the separation of the surface-related multiples from the data. Internal multiple separation might be achieved with other methods

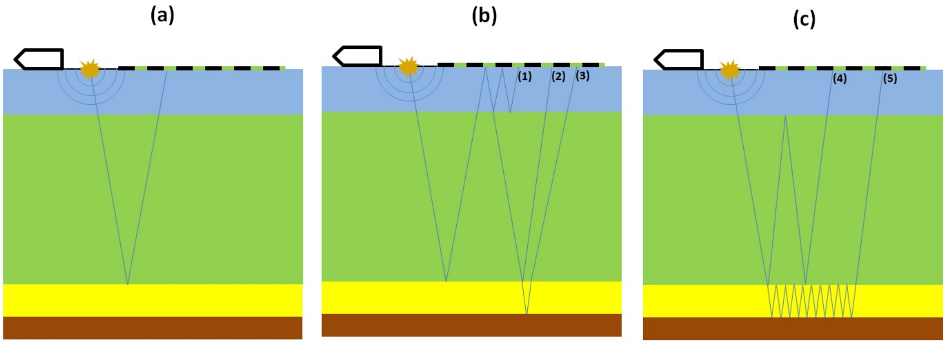

Figure 1.2: *Schematic representation of the different types of reflection events. The (a) primary event, together with some of its associated (b) surface-related multiples and (c) internal multiples. As we can see, surface-related multiples have all reflection points within the surface, while internal multiples have reflection points at the subsurface. The paths (1,2,3) have all at least one reflection point at the surface. The paths (4,5) have no reflection points at the surface.*

[Araujo et al., 1994; Coates and Weglein, 1996; Weglein et al., 1997; Verschuur and Berkhout, 2005; Ypma and Verschuur, 2013] or can be resolved during migration [Staal and Verschuur, 2013; Davydenko and Verschuur, 2014].

Surface-multiple separation method can be divided in the following categories:

■ 1.3.1 Move-out discrimination methods

These methods concentrate on exploiting the difference in the move-out behavior of the multiples with respect to the primaries after CMP gathering and NMO correction. In a simple medium, if the NMO correction is done properly, primaries are transformed into flat events and multiples are mapped into curved events. From here several options are possible. Simple stacking would eliminate part of the multiple energy, as the multiples do not add coherently to the stack. Another option is to use a transform domain to separate and later mute the undesired signals. Typical options for the transform domain are the the parabolic radon domain [Hampson, 1986] and the hyperbolic radon domain [Thorson and Claerbout, 1985]. The main draw-back of these methods is the fact that they assume parabolic or hyperbolic move-outs in the events, together with an increasing velocity profile in the subsurface.

■ 1.3.2 Predictive deconvolution methods

These method exploit the periodical nature of the multiples in order to subtract them from the input[1]. Here, a prediction-error filter is used to remove all the signals with a periodical structure [Robinson, 1954; Backus, 1959]. The procedure produces better results if applied in the linear Radon domain, given a flat-layered medium. Because it operates in a trace-by-trace fashion, the predictive deconvolution methods tend to fail in removing multiples from complex layers, where more complex wave-propagation effects might be present [Verschuur, 2006]. Because they are designed to minimize the energy of the primaries, these filters can only eliminate part of the primary energy, specially in the places where there is primary-multiple overlap.

■ 1.3.3 Wavefield extrapolation methods

Multiples can be also predicted in a model-based fashion by adding multiple round-trips in the water-layer to the input data. This process would reproduce multiple reverberations. Thus, by adding one round-trip in the first layer we can transform primaries into first-order multiples, second-order multiples into third-order multiples and so on. This technique is used for reproducing peg-leg multiples and is particularly useful in cases in which the water-bottom has a more complex structure, in which methods based on move-out discrimination and predictive deconvolution produce insufficient results. An obvious limitation of this method is that it will only produce multiples related with the bounduary or bounduaries provided by the model [see Verschuur and Berkhout, 2015]. Furthermore, care has to be taken to apply this method both to the source and receiver side of the data, in order to reproduce source-side and receiver-side reverberations [for more details, see Berryhill and Kim, 1986; Wiggins, 1988; Lokshtanov, 1999].

■ 1.3.4 Scattering series methods

These methods are based on the prediction of the multiple wavefield via a scattering series involving convolutions of the data with itself. The main representatives of this category are the inverse-scattering series method ”ISS” [Weglein et al., 1997], and the surface-related multiple elimination method ”SRME” [Verschuur and Berkhout, 1997]. Due to their physical background, these methods produce accurate multiple predictions for all orders of multiples, and proof that no subsurface information is needed for surface-related multiple prediction. The draw-back of these methods lie however in the necessity of an accurate estimation of the wavelet, which leads to extra processing steps or to additional constrains on the problem. These methods also rely on a dense acquisition, as the recorded field is used as a secondary source for the

---

[1] In simple geometries and near-offsets, multiples can be considerated as periodic, however this approximation breaks down for far offsets and more complex media.

multiple wavefield. Other methods compute multiples via SRME-type convolutions with model-based primary operators (Vasmel et al. [2016]). More of this in section 1.7.

■ 1.3.5 Inversion methods

These methods use inversion to estimate the primaries and/or multiples. The best representatives of this scheme are the EPSI method [van Groenestijn and Verschuur, 2009a], the Robust-EPSI method [Lin and Herrmann, 2013] and the CL-SRME method [Lopez and Verschuur, 2014, 2015b]. These methods estimate the multiples together with the primaries such that together they explain the input dataset. The process is based on the multiple prediction equation developed by Berkhout and Verschuur [1997] and driven by inversion theory. The main advantage of these methods rely on their extended capacities for data interpolation and the inclusion of sparsity as an extra constrain to eliminate/attenuate the dependency over the minimum energy constraint, which is known to produce suboptimal results in SRME [Guitton and Verschuur, 2004]. The main draw-backs would lie in the fact that these methods are computationally more expensive than SRME and that they require a larger degree of parameter tuning (coming from the sparsity constraints) to work properly. Other methods using inversion for multiple separation can be found in van Borselen et al. [1994]; Amundsen [2001]; Berkhout [2006].

1.4 From removing to using multiples

The idea of using multiples instead of removing them has grown a lot of acceptance in the industry in recent years. After all, multiples have shown to provide better illumination of the subsurface structures in many acquisition scenarios [Kumar et al., 2015], thus helping the migration process. By having extra travel paths through the subsurface, the multiples can also propagate though areas difficult to illuminate by a primary wavefield (such as subsalt reflectors) [Soni and Verschuur, 2015].

There is, however, still an interest in separating primaries from multiples. If imaged separately, multiples can provide less acquisition imprint when compared to the primaries [Berkhout and Verschuur, 1994; Lu et al., 2014; Davydenko and Verschuur, 2014; Berkhout et al., 2015]. Thus, a more robust image can be produced by combining the images obtained from from migrating primaries and multiples separately (figure 1.3). Techniques for including information about multiples in the migration result are currently being researched [Berkhout, 2012; Wapenaar et al., 2014; Berkhout and Verschuur, 2016].

With the later in mind, we will state the goal of this research to accurately separate the multiples from the primaries, rather than removing them.

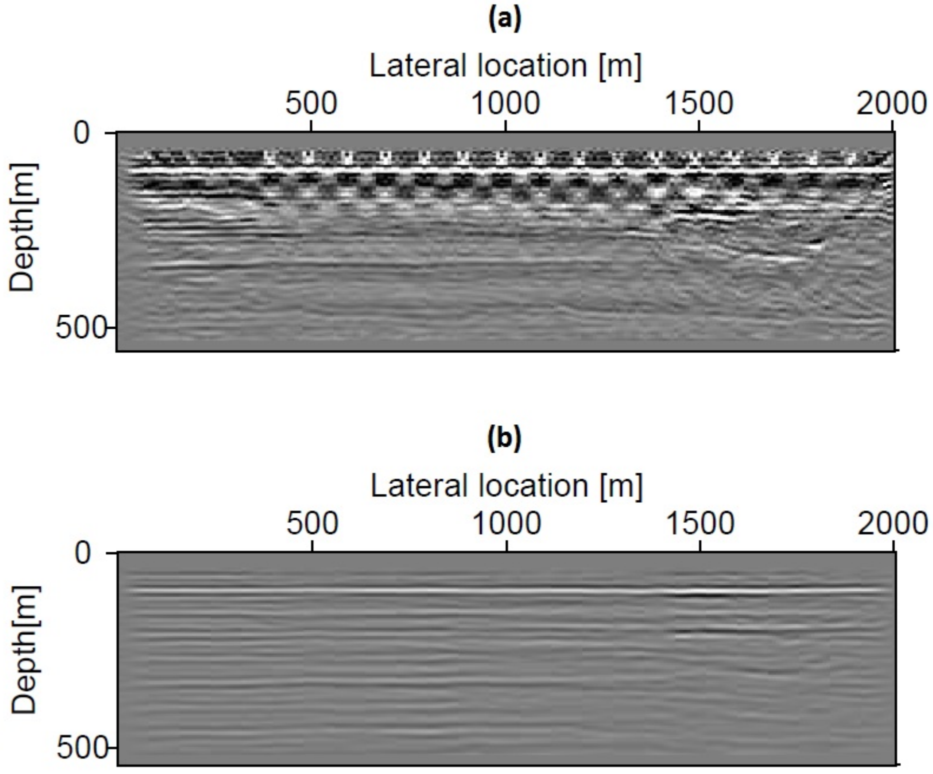

Figure 1.3: *Migration result from (a) primaries only data and (b) multiples only data. As we can see, in the shallow part multiples provide a better migration result than primaries. Courtesy of Mikhail Davidenko.*

1.5 Sampling

To obtain quality images of the subsurface we need to acquire sufficient seismic data to do the appropriated processing. In theory a wide grid fully filled with sources and receivers is required in order to achieve proper multiple separation, and consequently proper subsurface images. The problem resides, however, in the fact that in practice, due to time and budget constraints, seismic data is acquired with grids a lot sparser than those required for a proper processing (especially in the 3D case). To overcome this limitation, typically, data interpolation must be done before any attempt on multiple elimination.

This, however, presents a problem as generally the interpolation algorithms are based on mathematical/statistical properties of the seismic signal, rather than physical properties. Due to the physical inconsistency, the interpolated data tends to produce artefacts when multiple removal algorithms are used. These artefacts can manifest themselves as spurious events, small diffractions or poor primary/multiple separation.

There are two main types of sampling gaps:

[1] *Under-sampling gaps*: Is when the data is acquired in a coarse sampling grid (i.e. sparsely located sources and receivers), but a fine grid is desired for data processing. Here, the missing data typically occurs periodically in the dataset. Although some of the acquisition holes can be easily filled out with the use of reciprocity, still the larger part of locations is uncovered by both sources and receivers. We can easily assess the amount of interpolation required by finding the under-sampling ratio to the desired grid. This is given by $ratio = N_{in}/N_{out}$, with N_{in} the typical number of measured input traces and N_{out} the number of desired output traces. This quantity can also be expressed as $ratio = N_{in} : N_{out}$. If an under-sampled dataset is used in multiple prediction process, the resulting dataset will present diffraction-like artefacts related to aliasing.

[2] *Near-offset gaps*: Here the missing data is located in the near-offset part of the data. This is due to acquisition constraints, where sources and receivers cannot be brought too close together. Note that in this case, the information is consistently absent in the entire dataset and it cannot be filled with reciprocity. The near-offset traces typically contain the strongest energy as most of the event apexes fall inside it. This area is the most relevant for multiple prediction and also the most complicated to interpolate [see e.g. Dragoset and Jeričević, 1998; Verschuur, 2006].

Figures 1.5, 1.6 and 1.7 show the effect of missing data in the SRME multiple prediction process. The model used for the synthetic data generation is shown in 1.4. Figure 1.5a depicts the fully sampled data, figure 1.6a depicts a dataset with a 100m near-offset gap and figure 1.7a depicts a dataset with a 1:5 undersampling ratio. Their associated SRME-primaries are shown in figures 1.5b, 1.6b and 1.7b respectively. Several undersampling artefacts and multiples are seen in the primary estimates of the under-sampled datasets. This examples show the importance on an accurate interpolation for primary/multiple estimation.

1.6 Reconstruction in the literature

Due to economical constrains field data is often acquired in a sparse manner, making data reconstruction a necessity. There are several types of data reconstruction techniques in the literature. They can be grouped in the following categories:

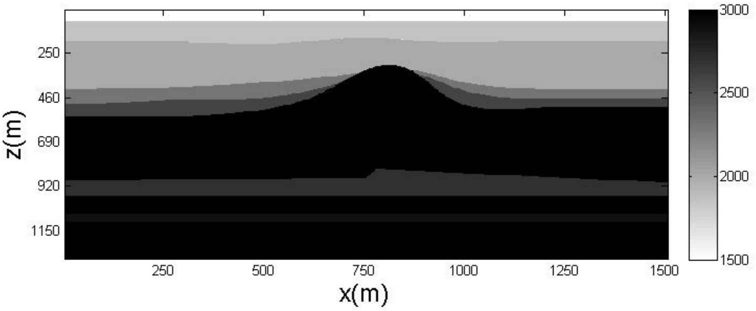

Figure 1.4: *Velocity model used for the synthetic data generation using a salt model. The first reflector is at 70m depth.*

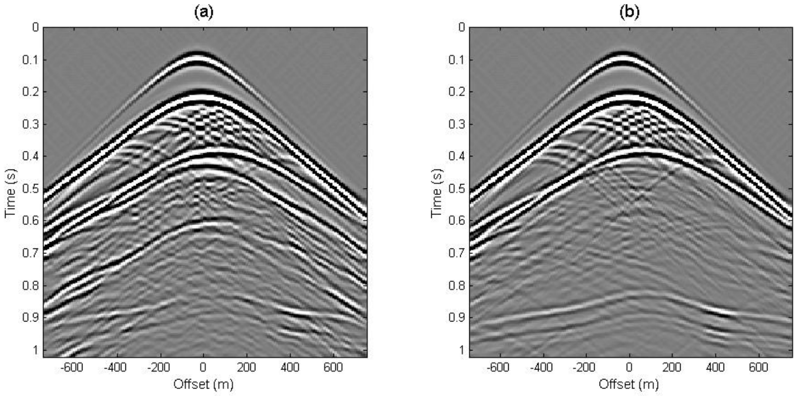

Figure 1.5: *(a) Fully sampled dataset and its corresponding (b) SRME primary estimation.*

1.6.1 Transformation-based reconstruction

These methods exploit the redundancy in the seismic data. They are based on the selection of a transform that maps the input data to a transform domain in which it can be represented more compactly. The transform is usually chosen such that it exploits some physical, statistical or mathematical structures inherent in the data. The transform parameters are typically found by inversion with a sparse regularization or by iterative thresholding [Abma and Kabir, 2006].

Several transform domains could be used for reconstruction, namely the Fourier domain [Duijndam and Schonewille, 1999], the parabolic Radon domain [Hugonnet and Canadas, 1995], the hyperbolic Radon domain [Trad, 2000], the wavelet domain

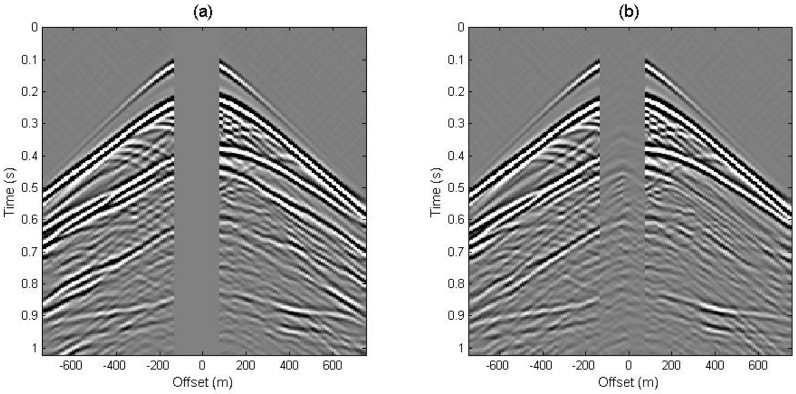

Figure 1.6: *(a) Dataset with the near-offsets removed (100m gap) and its corresponding (b) SRME primary estimation. Note that most of the multiple energy is not removed.*

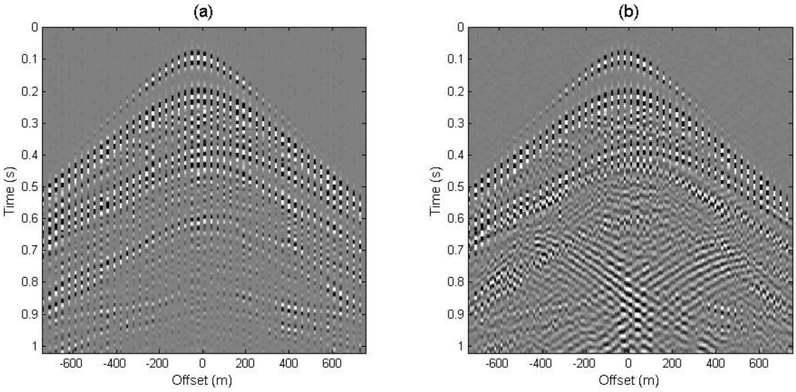

Figure 1.7: *(a) Dataset with a 1:5 under-sampling ratio and its corresponding (b) SRME primary estimation. Strong aliasing artefacts and wrong multiple prediction are observed.*

[Gan et al., 2015], the curvelet domain [Hennenfent et al., 2010], or the focal domain [Kutscha and Verschuur, 2012].

■ 1.6.2 Filter-based reconstruction

Seismic data can be reconstructed using a convolution with an interpolating filter. Probably the most well-known of these methods is the sinc interpolation based on the Whittaker-Kotelnikov-Shannon sampling theorem [Gensun, 1996]. This method can be implemented in practice by taking the signal to the $\omega - k_x$ domain, padding with zeros and then transforming back to the $x - t$ domain.

The generalized convolution method [Knutsson and Westin, 1993] uses gaussian windows as interpolating filters. Prediction-error filters in the time and frequency domain are used to interpolate data of distinctive plane waves [Spitz, 1991]. Laplacian filters [Briggs, 1994] or steering filters [Clapp et al., 2007] can also be used for interpolation. Though simple and easy to implement, these methods cannot easily overcome aliasing.

■ 1.6.3 Wavefield-operator based reconstruction

These methods are based on the Kirchhoff integral for wavefield extrapolation. One of the main advantages of using these operators is the fact that they can handle non-uniform acquisition geometries. The disadvantage would lie in the aliasing artefacts due to incomplete summations.

Several approaches are available, namely: offset continuation [Deregowski and Rocca, 1981; Bolondi et al., 1982], shot continuation [Spagnolini and Opreni, 1996], Azimuth Moveout [Biondi et al., 1998] and Inversion to Common Offset (ICO) [Chemingui and Baumstein, 2000].

■ 1.6.4 Tensor completition-based reconstruction

If organized in a tensorial form, the seismic data can be represented in a low-complexity form. The reconstruction is typically posed as an inversion problem constraint with a minimum data complexity measure under a given tensor decomposition. Matrix rank can be used as a measure of complexity [Candes and Recht, 2012]. Other methods allow different types of tensor decompositions such as High Order SVD (HOSVD) [Kreimer and Sacchi, 2011] or the Canonical Decomposition (CADECOMP) [see Kolda and Bader, 2009]. Another option is to flatten the tensors into matrices and then computing the matrix rank [Gandy et al., 2011], or to use the tensor SVD (tSVD) [Martin et al., 2013]. SVD-free methods can also be found in Lee et al. [2010] and Kumar et al. [2013].

1.7 The problem with shallow water multiples

Marine shallow-water environments provide a challenging setting to multiple separation algorithms. The reason behind this is that the primary water-bottom reflector in not always present in shallow-water data due to the near-offset gap. Without the proper primary generator operator, algorithms like SRME are known to provide inaccurate results if applied without any interpolation (figures 1.6 and 1.7). Dragoset and Jeričević [1998] demonstrated the strong effect that missing near-offsets can have in primary/multiple separation. If the sea-floor reflection is strong, water-bottom multiples can become very problematic, as they may hinder many of the details in deeper structures.

One simple strategy to overcome this limitation would be to interpolate the input data first and then run the SRME algorithm to separate the multiples. The limitation comes then from the fact that accurate near-offset interpolation remains a challenge for most of the available reconstruction methods. The problem arises as the data in the near-offsets can contain a lot of information about the angle-dependent reflectivity of the reflected events. This information is not trivially encoded in far-offset part of the data and cannot be accurately reconstructed by simple reconstruction methods. Effects like NMO wavelet-stretching can make the reconstruction effort even more challenging, as many of the previously discussed reconstruction methods are meant to be applied after an NMO correction. The challenge is larger for data from shallow water enviroments, in which the relevant information concentrates more and more in the missing part of the data.

A possible strategy to solve this problem consists in doing an initial Radon interpolation to the data, muting the post-critical reflections and then applying predictive deconvolution to eliminate the water-bottom multiples. It is possible to use predictive deconvolution to remove the water-layer reverberations, and then use SRME to remove all the remaining surface multiples. Predictive deconvolution is known to work reasonably well in the shallow-water scenarios [Verschuur, 2006]. The main drawback of these approaches is the implied simplicity of the propagation medium and the fact that the multiples are no longer periodic for offsets different than zero.

Another option is to use a sea-floor model to reproduce the first-reflector primary response. Then the necessary primary information is calculated via forward modelling of the wavefields in the model. Surface multiples are predicted after convolution with the data. This model-based scheme for multiple prediction can be seen an alternative method for fully data-driven multiple prediction, with the advantage that it can be applied to data with limited offsets. In this way the water-bottom multiples are eliminated. The former method is known as SWD (Shallow Water Demultiple) [Hung et al., 2010]. Modifications of this method have also been proposed [Hargreaves, 2006; Sun and Wang, 2010]. Being a model-based scheme, the accuracy of the predicted multiples will depend on the accuracy of the chosen subsurface model.

In this research we present an alternative approach for the proper separation of the shallow water multiples. We propose a solution in which multiple estimation is performed simultaneously with data reconstruction (such that data reconstruction helps obtaining better multiple estimates) and in which the physical primary-multiple relationship helps constraining the data interpolation process. The near-offset data is reconstructed via back-projection of the multiples into the primaries. In this way, we allow the information contained in the multiples to be used to reconstruct the missing shallow primary information. This research can be seen as an extension of the previously introduced (robust) EPSI methodology. Here, we use a similar multiple prediction machinery, but now we largely extend the interpolation capabilities of (robust) EPSI, making the new algorithm capable of: (1) interpolation of medium-large near-offset gaps, (2) interpolation of missing sources/receivers, (3) a 3D implementation.

1.8 Thesis outline

The main content of the present thesis is divided in the following chapters:

[**1**] *Chapter 2*: In this chapter the forward model for multiple prediction is explained, together with some definitions, notational conventions and some elaboration on previously introduced multiple-separation methods, namely: SRME and EPSI.

[**2**] *Chapter 3*: The Closed-Loop SRME theory is defined here. A comparison is made between CL-SRME, EPSI and SRME.

[**3**] *Chapter 4*: Here the focal transform is introduced. We start by some basic derivations to define the propagation operators and their inverses. The focal transform is later introduced as a two-sided back-propagation of the data. Some properties and uses are discussed.

[**4**] *Chapter 5*: The Focal CL-SRME method is introduced here. The basic theory is outlined together with some discussion about its characteristics.

[**5**] *Chapter 6*: This chapter contains a set of results for both 2D synthetic and 2D field datasets.

[**6**] *Chapter 7*: The 3D Focal CL-SRME method is introduced here. We devote this chapter to introduce the basic concepts needed to grasp the extension of the algorithm developed in Chapter 5 to the 3D case.

[**7**] *Chapter 8*: This chapter contains a set of results from the 3D Focal CL-SRME method.

[**8**] *Chapter 9*: We provide general conclusions to the thesis together with some discussion about its applications and suggestions for further research.

2

The forward model for seismic data

In this chapter we will study a forward model for seismic data that suits the used inversion methodology later on. First, we will concentrate on some definitions, notation and properties of some of the quantities to be used in this thesis. Then, we will describe the theoretical background behind multiple estimation and the various relationships between primaries and multiples. Finally, we will describe some common algorithms used to estimate the primary dataset namely: SRME and EPSI. The material follows closely the theory developed at Delft University for acoustic and elastic wavefield description [Berkhout, 1982; Wapenaar and Berkhout, 1989; Fokkema and van den Berg, 1993; Berkhout and Verschuur, 1997; Verschuur and Berkhout, 1997].

2.1 Matrix representation of seismic quantities

To describe our multiple estimation methodology, it is convenient to represent discrete seismic data and operators in a concise manner in data matrices.

■ 2.1.1 The data matrix: $\mathbf{P}$

2D

For our matrix notation, the reflected wavefields in a seismic experiment are assumed to be generated and recorded over a coinciding grid of sources and receivers. The measured wavefield $p(x_s, x_r, t)$ depends upon the source position x_s, the receiver position x_r and the measurement time t. If we assume an uniform acquisition, we

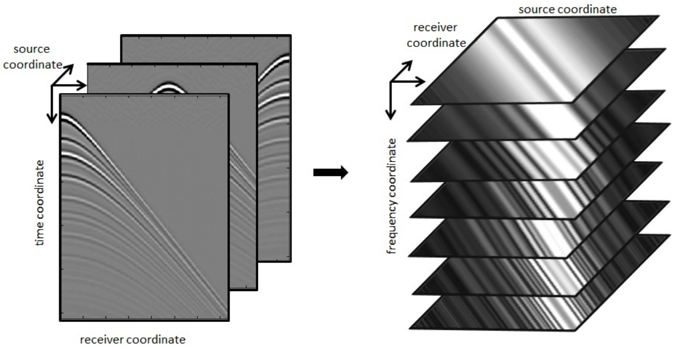

Figure 2.1: *Generation of the data matrix given the measured input data.*

have $x_s = j\Delta x$, $x_r = k\Delta x$ and $t = l\Delta t$, where Δx is the source-source or the receiver-receiver distance, Δt is the sampling ratio and $(j, k, l) \in \mathbb{Z}$ are the indices representing the position of the measurement $p(x_s, x_r, t)$ in the grid.

In practice the quantities Δx and Δt should be set before-hand in order to avoid aliasing. The coordinate origin for x_s, x_r and t is chosen arbitrarily, however, once chosen, it must remain invariant for all processing steps. Note the x_s and x_r coordinates are taken with respect to the earth and do not represent offset coordinates.

We are also assuming that the data is acquired at the surface at a common depth level $z = z_0$.

Taking the data to the frequency domain ($t \longrightarrow \omega_l \Delta\omega$) and using the defined set of coordinates, we can use the detail-hiding notation (Berkhout [1985]) to store the information from all the measurements $p(x_s, x_r, t)$ into a set of mono-chromatic matrices $\mathbf{P}$ organized as depicted in figures 2.1 and 2.2.

In this convention, every column of $\mathbf{P}$ represents the wavefield of a single source in all the receivers (for a given frequency), and every row represents the wavefield measured from all the sources in a single receiver position (for a given frequency).

Furthermore, in figure 2.2 it is shown that diagonals and anti-diagonals contain common offset or common midpoint data, respectively.

Note that, due to the orthogonally of the Fourier basis, all frequencies can be processed independently of each other.

3D

By extending the dimensionality from 2D to 3D we are actually adding two new spatial dimensions, namely $y_s = m\Delta y$ and $y_r = n\Delta y$ to our wavefield coordinates

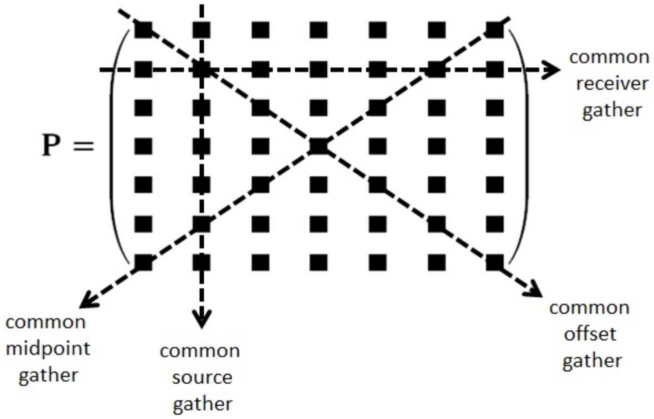

Figure 2.2: *Structure of the data matrix for one frequency component.*

(cross-line source and receiver coordinates). Now Δy is the receiver-receiver distance in the y direction. Note that Δy and Δx can differ from each other as long as both are constant [see Kinneging et al., 1989]. This extension will in principle yield a 5-dimensional entity $p(x_s, y_s, x_r, y_r, t)$.

For convenience we want to organize the data such that we can still represent it in terms of a set of 2-dimensional matrices in the frequency domain. An internal arrangement will allow us maintain the same theoretical formulas for both the 2D and the 3D cases.

To construct the desired 3D matrix organize the data in '3D shot gathers' as shown in figure 2.3. Then the coordinate system inside each sub-matrix will follow figure 2.4. Note that in this 3D matrix one column still represents a 3D shot record with all receiver positions of the 3D shot concatenated in one long vector.

Note that for practical acquisition geometries the data matrices (especially in 3D) will never be fully filled with actual data.

■ 2.1.2 The source matrix: S

The source wavefield is the total wavefield placed by the source devices in the subsurface. In the case that the source devices are located below the subsurface, this wavefield includes a source-ghost that makes the input an effective dipole source (assuming that the wavefield is measured at distances much larger than the effective dipole separation).

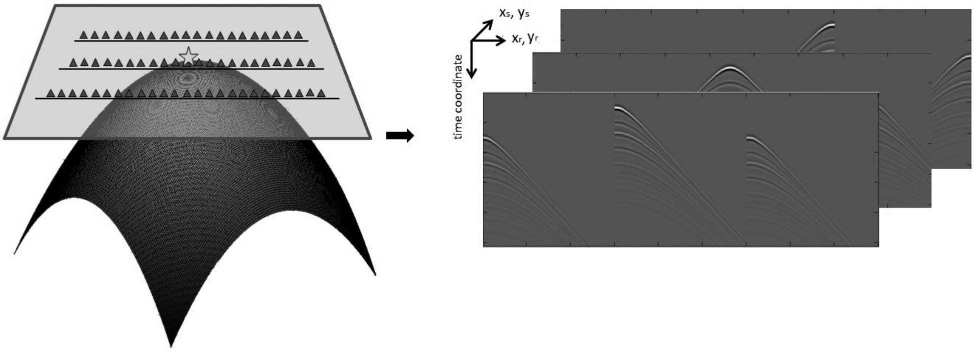

Figure 2.3: *Organization of the 3D data into a data matrix by gathering the traces from all the receiver lines of a given shot, where* (x_s,y_s) *are the source coordinates and* (x_r,y_r) *are the receiver coordinates.*

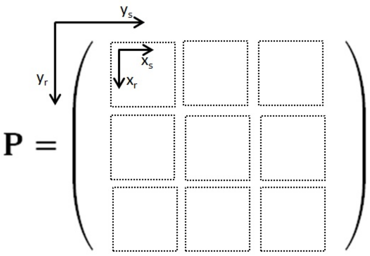

Figure 2.4: *Structure of the 3D data matrix for one frequency component, where* (x_s,y_s) *are the source coordinates and* (x_r,y_r) *are the receiver coordinates.*

Every column of the **S** matrix contains a monochromatic component of the source wavefield for a given experiment (determined by the lateral position). If **S** is a diagonal matrix then the sources have a dipolar radiation pattern [see also Verschuur, 1991]. If **S** is non-diagonal then, the off-diagonal elements in every column account for the shape of the radiation pattern (with respect to the horizontal coordinates) of that particular experiment. Fully-filled **S** matrices are typical in blended acquisitions [Berkhout et al., 2009]. If **S** is a Toeplitz matrix then all the source signatures are identical. Note also, that **S** has no need to be square (i.e. there might be less shot positions than receiver positions). Figure 2.5 depicts some the possible configurations for the source matrix.

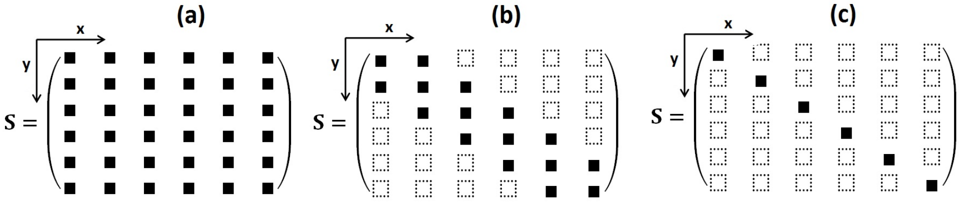

Figure 2.5: *Schematic representation of the source matrix structure for different source configurations (blank spaces are zeros, filled spaces are non-zeros). (a) Source matrix for a blended experiment, (b) source matrix for sources with a complex radiation pattern and (c) source matrix for sources with a dipolar radiation pattern. Here,* $\mathbf{x}$ *is an index enumerating the columns of the matrix (number of active sources in the experiment) and* $\mathbf{y}$ *is an index enumerating the rows of the matrix (lateral acquisition grid coordinates). The source matrix has no need to be square.*

2.1.3 The propagation matrix: $\mathbf{W}$

The $\mathbf{W} = \mathbf{W}(z_m, z_0, \omega)$ operator downward-propagates a monochromatic component of a wavefield from a z_0 depth-level to a z_m depth-level. In this description, every column of $\mathbf{W}$ contains the response of a dipolar impulsive source at a depth level z_0 measured in all the receivers at depth z_m, here $z_0 < z_m$. We will discuss the $\mathbf{W}$ operators in more detail in chapter 4.

2.1.4 The reflectivity matrix

At the surface $z = 0$*:* $\mathbf{R}^\cap$*:* This operator transforms the up-coming wavefield at the surface into a down-going wavefield. It is related with the discontinuity in the earth-air or sea-air boundary. In the marine case this reflectivity is angle-independent and the reflectivity matrix takes the form $\mathbf{R}^\cap(z_0) \approx -\mathbf{I}$ (almost all the wavefield is mirrored back to the subsurface). In the land case $\mathbf{R}^\cap(z_0)$ is more complex [see Wapenaar and Berkhout, 1989].

At a given depth-level $z = z_m$*:* $\mathbf{R}^\cup_m$, $\mathbf{R}^\cap_m$*:* These operators transform the up-coming wavefield at a depth-level z_m into a down-going wavefield ($\mathbf{R}^\cap_m$) and vice-versa ($\mathbf{R}^\cup_m$). They contain the information of the reflection properties of all lateral grid-point positions along a given depth-level. Every element on these operators depends on the material properties of the grid point in consideration. These matrices represent convolution operators, where angle-dependent reflection information is included [de Bruin et al., 1990].

2.1.5 The detector matrix: $\mathbf{D}$

This operator transforms the up-coming wavefield into the measured wavefield. In general $\mathbf{D}$ contains the information of any directivity in the receivers, together with the receiver response and the receiver acquisition geometry. Receiver-side ghosts are also included in $\mathbf{D}$.

2.1.6 The primary impulse response: $\mathbf{P}_0$

This is the up-coming wavefield obtained at the surface if we assume the water-air or the earth-air discontinuity to be absent (transparent surface). It neglects all the scattering related to the free surface. Note, however, that the internal multiple scattering in the subsurface will be present.

2.2 Notational conventions

The following conventions will be assumed in this document

- We will assume $\mathbf{D} = \mathbf{I}$. Thus, we assume receiver directivity and ghosts have been corrected for.
- We will loosely call $\mathbf{P}_0$ the "primaries" given the understanding that we actually mean "the primaries plus internal multiples".
- Estimates of theoretical quantities will be denoted by "hatted" symbols (e.g. $\hat{\mathbf{P}}_0$ will stand for "estimate of the primaries"). Un-hatted symbols will refer to theoretical quantities.
- Iteration dependent quantities will be labeled by their iteration number as a super-script, e.g. $\hat{\mathbf{P}}_0^{(0)}$ will refer to the estimate of the quantity $\mathbf{P}_0$ at the begining of the algorithm (starting value). Similary $\hat{\mathbf{P}}_0^{(n)}$ will refer to the estimate of the quantity $\mathbf{P}_0$ after n iterations. For simplicity, we will drop the iteration superscript whenever possible.
- We will omit the explicit reference to the frequency samples in the formulas (e.g. $\mathbf{P}_0 = \mathbf{P}_0(\omega)$).
- Sometimes we will also abbreviate the explicit reference to a given depth level z_m in our wavefield matrices by a using its label index m as a subindex for our matrices. Together with the ommision of the explicit frequency variation, this reduces to a simple subindexing, e.g. $\mathbf{W}(z_m, z_0, \omega) = \mathbf{W}_m$. Note that we also choose the acquisition surface to be at $z_0 = 0$.

- The data matrices $\mathbf{P}$ are assumed to be have a dipole character. A dipole response is naturally created by placing a monopole source and a pressure-free plane in close proximity to each other.
- The data matrices $\mathbf{P}$ are supposed to contain only up-going wavefields. This means that laterally propagating wavefields like direct-waves, ground-roll must be removed beforehand. Also it means that down-going wavefields like the receiver-ghost must also be removed.

2.3 Forward model for primaries and multiples

2.3.1 Primaries

Following Berkhout (1985) we can express the primary up-coming wavefield at the receiver level in terms of a convolutional propagation-reflection-propagation model also known as the WRW-model. If we omit for the time-being all the multiple-scattering phenomena and restrict ourselfs to a linear wavefield theory, we can write:

$$\mathbf{P}_0 = \sum_{m=1}^{\infty} \mathbf{W}_m^T \mathbf{R}_m^{\cup} \mathbf{W}_m \mathbf{S} = \mathbf{X}_0 \mathbf{S}, \tag{2.3.1}$$

with

$$\mathbf{X}_0 := \sum_{m=1}^{\infty} \mathbf{W}_m^T \mathbf{R}_m^{\cup} \mathbf{W}_m, \tag{2.3.2}$$

where $\mathbf{X}_0$ is the primary impulse response of the earth, $\mathbf{W}_m$ is the wavefield extrapolation operator and $\mathbf{R}_m^{\cup}$ is the reflectivity operator. Equation 2.3.1 expresses the propagation of the source wavefield to a given subsurface level z_m, the reflection in that level and the propagation of the reflected wavefield back to the surface. The responses from all the depth levels stack up to the final reflected signal.

2.3.2 Primaries and surface multiples

To include the surface multiple reflections in our forward model (eq. 2.3.2), note that once the reflected wavefield from the primaries has reached the surface, the material discontinuity at the surface will reflect back the primary information into the medium. In this way the surface itself becomes a secondary source for the down-going wavefield. This secondary wavefield will be sent back into the subsurface for an additional round-trip into the earth. Once the secondary wavefield has been reflected in the medium, the free surface will reflect again this information into the subsurface for an extra round-trip. This process continues until the energy of the multiple reflecting signals is exhausted. In this way, we expect to describe our wavefield measured at the receivers as an infinite sum of multiple contributions. The former can be expressed mathematically as [Verschuur et al., 1992; Fokkema and van den Berg, 1993]:

$$\mathbf{P} = \mathbf{X}_0\mathbf{S} + (\mathbf{X}_0\mathbf{R}^{\cap})\mathbf{X}_0\mathbf{S} + (\mathbf{X}_0\mathbf{R}^{\cap})^2\mathbf{X}_0\mathbf{S} + (\mathbf{X}_0\mathbf{R}^{\cap})^3\mathbf{X}_0\mathbf{S} + \cdots, \tag{2.3.3}$$

where every successive term represents one additional order of scattering, namely: $\mathbf{X}_0\mathbf{S}$ represent the primaries, $(\mathbf{X}_0\mathbf{R}^{\cap})\mathbf{X}_0\mathbf{S}$ are the first order multiples, $(\mathbf{X}_0\mathbf{R}^{\cap})^2\mathbf{X}_0\mathbf{S}$ are the second order multiples, etc.

The convergence of this series is guaranteed by the fact that the energy measured in $\mathbf{P}$ is finite. A formal proof of the convergence of the series can be found in Fokkema and van den Berg [1993]. Recognizing equation 2.3.3 as a geometric series and defining the *surface operator* $\mathbf{A}$ as $\mathbf{A} = \mathbf{S}^{-1}\mathbf{R}^{\cap}$ and $\mathbf{P}_0 = \mathbf{X}_0\mathbf{S}$ we can rewrite equation 2.3.3 into[1]

$$\begin{aligned}
\mathbf{P} &= \mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})\mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})^2\mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})^3\mathbf{P}_0 + \cdots \qquad (2.3.4)\\
\mathbf{P} &= (\sum_{m=0}^{\infty}(\mathbf{P}_0\mathbf{A})^m)\mathbf{P}_0\\
\mathbf{P} &= (\mathbf{I} - \mathbf{P}_0\mathbf{A})^{-1}\mathbf{P}_0\\
\mathbf{P}_0 &= \mathbf{P} - \mathbf{P}_0\mathbf{A}\mathbf{P}\\
\mathbf{P} &= \mathbf{P}_0 + \mathbf{P}_0\mathbf{A}\mathbf{P}\\
\mathbf{P} &= \mathbf{P}_0(\mathbf{I} + \mathbf{A}\mathbf{P})\\
\mathbf{P}_0 &= \mathbf{P}(\mathbf{I} + \mathbf{A}\mathbf{P})^{-1}\\
\mathbf{P}_0 &= \mathbf{P}(\sum_{m=0}^{\infty}(-\mathbf{A}\mathbf{P})^m)\\
\mathbf{P}_0 &= \mathbf{P} - \mathbf{P}(\mathbf{A}\mathbf{P}) + \mathbf{P}(\mathbf{A}\mathbf{P})^2 - \mathbf{P}(\mathbf{A}\mathbf{P})^3 + \cdots.
\end{aligned}$$

From these equations we can recognize the following set of equivalent expressions for the total and primary wavefields

- **Series**:

$$\begin{aligned}
\mathbf{P} &= \mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})\mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})^2\mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})^3\mathbf{P}_0 + \cdots \qquad (2.3.5)\\
\mathbf{P}_0 &= \mathbf{P} - \mathbf{P}(\mathbf{A}\mathbf{P}) + \mathbf{P}(\mathbf{A}\mathbf{P})^2 - \mathbf{P}(\mathbf{A}\mathbf{P})^3 + \cdots; \qquad (2.3.6)
\end{aligned}$$

---

[1] Even though the definition of $\mathbf{A}$ involves the inverse of $\mathbf{S}$, in practice $\mathbf{A}$ is never calculated like that. After some assumptions, $\mathbf{A}$ can be calculated as the optimal filter matching the measured data with the predicted multiples. This will be further explained in section 2.4.1.

- **Recursive**:

$$\mathbf{P} = \mathbf{P}_0 + \mathbf{P}_0\mathbf{A}\mathbf{P} \tag{2.3.7}$$
$$\mathbf{P}_0 = \mathbf{P} - \mathbf{P}_0\mathbf{A}\mathbf{P}; \tag{2.3.8}$$

- **Inverse space**:

$$\mathbf{P}^{-1} = \mathbf{P}_0^{-1} - \mathbf{A} \tag{2.3.9}$$
$$\mathbf{P}_0^{-1} = \mathbf{P}^{-1} + \mathbf{A}. \tag{2.3.10}$$

All these expressions are equivalent to each other, but they provide different means to reach the desired result and bring different insights into the solution. An equivalent set of expressions in the $(\mathbf{X}_0, \mathbf{S})$ parameterization is given in Appendix A.

Equations 2.3.5 and 2.3.6 provide an explicit interpretation of the scattering at the surface via a scattering series, every term accounts for one order of scattering. Here both the forward (equation 2.3.5) and the inverse (equation 2.3.6) series can be computed.

Expressions 2.3.7 and 2.3.8 provide a more practical way of accounting for the multiple scattering via single matrix products [Berkhout and Verschuur, 1997]. From equation 2.3.7 we can conclude that surface multiples $\mathbf{M}$ may be written as $\mathbf{M} = \mathbf{P}_0\mathbf{A}\mathbf{P}$. Equation 2.3.8 is the basis for the SRME method, while equation 2.3.7 is the basis for the CL-SRME method.

Equations 2.3.9 and 2.3.10 represent the solution in the so called ”inverse space” [Berkhout, 2006; Verschuur and Berkhout, 2006]. Though cumbersome to use in practice (typically $\mathbf{P}$ and $\mathbf{P}_0$ are not invertible), these set of equations provide the theoretical advantage of solving for the quantity $\mathbf{A}$ in terms of $\mathbf{P}$ and $\mathbf{P}_0$, by $\mathbf{A} = \mathbf{P}_0^{-1} - \mathbf{P}^{-1}$. However, this method of solution is complicated to achieve in practice.

Though comprehensive and intuitive, the derivation 2.3.5 is by no means detailed. A more formal derivation of these expressions, based on the principle of acoustic wavefield reciprocity, can be found in Fokkema and van den Berg [1993], which is an excellent reference for addressing these topics. The previous derivations follow closely the ones in Verschuur et al. [1992]; Verschuur and Berkhout [1997]. For a more intuitive undestanding of these expressions, the reader is referred to Dragoset and Jeričević [1998] and Dragoset et al. [2010].

2.4 Methods for multiple elimination: SRME and EPSI

Now that the basic theory behind multiple scattering is in place, we can revisit the SRME and EPSI methodologies which make use of the equation 2.3.3 to obtain estimates of the primary response of the earth.

2.4.1 SRME

Surface Related Multiple Elimination (SRME) was introduced by Verschuur et al. [1992] as a method to carry out equation 2.3.6. Later on, Berkhout and Verschuur [1997] redefined it as an iterative method of primary estimation. In principle the algorithm works by estimating the multiples and then subtracting the estimated multiples from the input data to obtain an estimate of the primaries. The method is based on the iterative version of equation 2.3.8, namely:

$$\hat{\mathbf{P}}_0^{(k+1)} = \mathbf{P} - \hat{\mathbf{P}}_0^{(k)}\hat{\mathbf{A}}^{(k+1)}\mathbf{P}, \tag{2.4.11}$$

and the constraint

$$\hat{A}^{(k+1)} = \underset{\hat{A}^{(k+1)}}{\arg\min} \sum_{\omega} ||\mathbf{P} - \hat{\mathbf{P}}_0^{(k)}\hat{A}^{(k+1)}\mathbf{P})||^2, \tag{2.4.12}$$

where k represents the iteration number. Here, $\hat{\mathbf{A}}$ is replaced by an angle-independent approximation $\hat{A}(\omega)\mathbf{I}$. Since there are more unknowns: $\hat{\mathbf{P}}_0^{(k+1)}$ and $\hat{\mathbf{A}}^{(k)}$, than knowns: $\mathbf{P}$, in equation 2.4.11 the extra constraint 2.4.12 is needed.

The process goes as follows: Starting from $\hat{\mathbf{P}}_0^{(0)} = \mathbf{P}$ and $\hat{\mathbf{A}}^{(0)} = \mathbf{0}$ we estimate the multiples $\hat{\mathbf{M}}^{(k+1)} = \hat{\mathbf{P}}_0^{(k)}\hat{\mathbf{A}}^{(k+1)}\mathbf{P}$ in each iteration. This involve two steps: (1) calculate the product $\hat{\mathbf{P}}_0^{(k)}\mathbf{P}$, and (2) estimate $\hat{\mathbf{A}}^{(k)}$. The last step is done by regarding $\hat{\mathbf{A}}^{(k)}$ as an adaptive filter and finding the optimal solution for 2.4.12 in the time domain [see Verschuur and Berkhout, 1997]. The former is equivalent to assume that the primaries have minimum energy (in the L2 norm), and holds exactly only when the primary and multiple events are non-interfering. This assumption provides satisfactory subtraction results, but does not work properly in all cases [see Nekut and Verschuur, 1998]. Guitton and Verschuur [2004] and van Groenestijn and Verschuur [2008] have shown that other minimization norms, like the L1 norm or a sparseness norm, can lead to different, and sometimes better, subtraction results.

Once a multiple estimate is available, we can use equation 2.4.11 to obtain a new $\mathbf{P}_0$ estimate. This new $\mathbf{P}_0$ estimate will allow us to calculate as well a new $\mathbf{M}$ estimates, thus generating an iterative process.

It can be shown that every iteration in this process will actually add one extra order of scattering to the scattering series solution equation 2.3.5. At the same time, every iteration will push the previous $\mathbf{A}$ estimates towards correspondingly higher orders in the scattering series, in this way the iterative process actually "cleans itself" from the influence of the non-very-accurate initial $\mathbf{A}$ estimates. The SRME primary estimate after n iterations is then:

$$\hat{\mathbf{P}}_0^{(n)} = \mathbf{P} - \mathbf{P}\hat{\mathbf{A}}^{(n)}\mathbf{P} + \mathbf{P}\hat{\mathbf{A}}^{(n-1)}\mathbf{P}\hat{\mathbf{A}}^{(n)}\mathbf{P} - \mathbf{P}\hat{\mathbf{A}}^{(n-2)}\mathbf{P}\hat{\mathbf{A}}^{(n-1)}\mathbf{P}\hat{\mathbf{A}}^{(n)}\mathbf{P} + \cdots, \quad (2.4.13)$$

where $\hat{\mathbf{A}}^{(k)}$ denotes the k-th iteration estimate of the surface operator. Note that the predominant terms of the series (the lower order ones) are always given by the latest $\mathbf{A}$ estimations, which are always more accurate.

For a detailed description of the SRME algorithm refer to Verschuur [1991]; Verschuur and Berkhout [1997]; Weglein et al. [1997]; Berkhout and Verschuur [1997]; Verschuur [2006]; Dragoset et al. [2010].

2.4.2 EPSI

To describe the Estimation of Primaries by Sparse Inversion algorithm (EPSI) [van Groenestijn and Verschuur, 2009a] we should rewrite equation 2.3.9 in the ($\mathbf{X}_0$,$\mathbf{S}$) parameterization. If we take $\mathbf{S}(\omega) = S(\omega)\mathbf{I}$ (i.e. a constant source wavelet for all shots and no directivity) and we assume the surface reflectivity to be a scalar $R^\cap$ (being equal to -1) we get:

$$\mathbf{P} = \mathbf{X}_0 S + \mathbf{X}_0 R^\cap \mathbf{P}. \quad (2.4.14)$$

In EPSI we use full waveform inversion to estimate the unknown primary impulse response $\hat{\mathbf{X}}_0$ and source wavelet $\hat{S}$ such that the estimated primaries $\hat{\mathbf{X}}_0\hat{S}$ together with the corresponding surface multiples $\hat{\mathbf{X}}_0 R^\cap \mathbf{P}$ can explain the total upgoing data $\mathbf{P}$. The unknown dataset $\hat{\mathbf{X}}_0$ is parameterized in the time domain with spikes. The objective function to minimize in EPSI is built by calculating the Frobenius norm of the difference between the estimated dataset $\hat{\mathbf{X}}_0\hat{S} + \hat{\mathbf{X}}_0 R^\cap \mathbf{P}$ and the measured dataset $\mathbf{P}$ for all frequencies ω as[2]

$$J = \sum_{\omega} ||\mathbf{P} - \hat{\mathbf{X}}_0(\hat{S} + R^\cap \mathbf{P})||^2. \quad (2.4.15)$$

The EPSI algorithm calculates $\hat{\mathbf{X}}_0$ and $\hat{S}$ such that J is driven to zero. This is done in an iterative way where the primary impulse response data volume $\mathbf{X}_0$ is built up slowly during the iteration process in the time domain (causality constraint) and sparsity is imposed on the updates of $\mathbf{X}_0$ (sparsity constraint). In this way the adaptive subtraction is avoided and interference between primaries and multiples is

---

[2] Equation 2.4.15 provides the objective function that is used for calculating the gradient in EPSI. However, for the algorithm to work properly, causality and sparsity constratints need to be applied to the gradient. These operations are applied in a element-wise fashion over the gradients and are not represented in equation 2.4.15. We also require S to be a compact filter in time.

better handled. Furthermore, missing near-offset data can be interpolated simultaneously with multiple elimination, improving the results especially for shallow water situations.

For a detailed description of the EPSI algorithm please refer to van Groenestijn and Verschuur [2009a,b, 2010, 2011]

Figure 2.6 compares the results of SRME and EPSI when applied to a shallow marine dataset from the North Sea. Due to the sparsity constraint and the involved near-offset reconstruction, the EPSI algorithm tends to produce primary estimates with a larger degree of multiple suppression [van Groenestijn, 2010].

A better alternative to the EPSI algorithm was introduced by Lin and Herrmann [2013]. This methodology (called robust-EPSI) presents advantages over the initial EPSI formulation [van Groenestijn and Verschuur, 2009a], namely a better estimation of deep primaries and a more robust inversion, allowing also a near-offset interpoaltion extension [Lin and Herrmann, 2016]. The main draw-back of this algorithm is its expensive computational efficiency, that make a possible 3D application unfeasible.

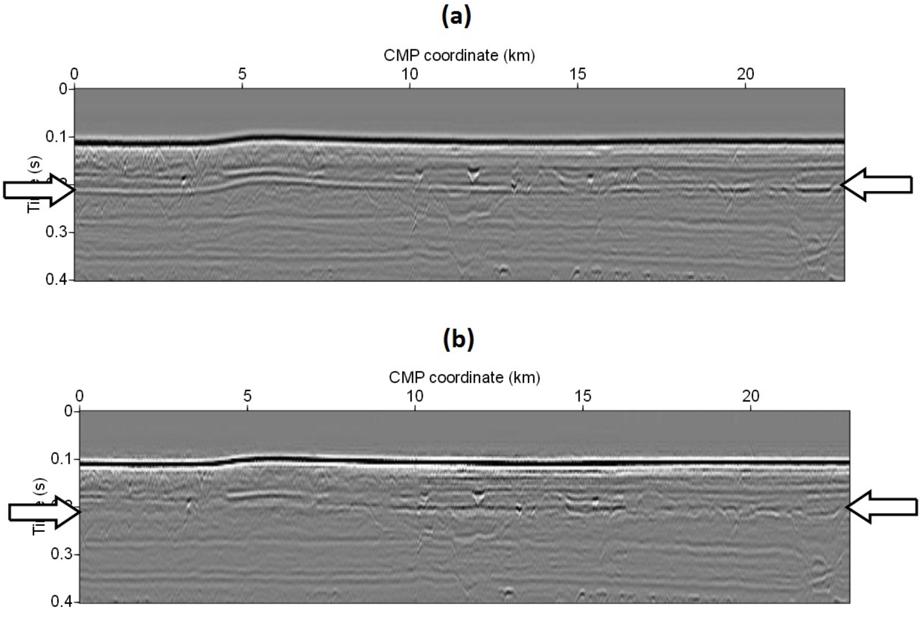

Figure 2.6: *Primary stacks from a moderately shallow water dataset acquired in the North Sea. (a) SRME primaries and (b) EPSI primaries. As we can see EPSI presents better multiple suppression when compared with SRME. Taken from van Groenestijn (2010).*

3

Closed-Loop SRME

In this chapter we will introduce the Closed-Loop SRME method which will serve as a theoretical frame-work for the next chapters. The basic theory, implementation, comparison with other methods and possible extensions are discussed.

3.1 Introduction

The Closed-Loop SRME algorithm combines the inversion (or ”closed-loop”) approach of EPSI with the simplicity of the SRME algorithm. The theoretical framework developed here can be used to generate many extensions to the presently described algorithm. For example, if additional sparsity requirements are added to the scheme, the minimum energy constraint (typical from SRME-related algorithms) might be dropped, providing better primary estimates. Little advantages are seen in 2D CL-SRME when compared to an algorithm like EPSI, however, important benefits will arise in the 3D case, were the EPSI algorithm is expected to have some practical difficulties.

In the following sections, we will introduce the CL-SRME method, discuss its properties and applications, as well as its similarities and differences with respect to both SRME and EPSI. By itself, the CL-SRME algorithm can be considerated as a inversion-based version of the SRME algorithm.

3.2 Theory

Let $\mathbf{P}$ describe the up-going wavefield at the surface, in the detail hiding notation [Berkhout, 1982; Kinneging et al., 1989] we can write the data $\mathbf{P}$ as a function of the primaries $\mathbf{P}_0$ and the surface operator $\mathbf{A}$ via

$$\mathbf{P} = \mathbf{P}_0(\mathbf{I} + \mathbf{AP}), \tag{3.2.1}$$

were $\mathbf{A}$ is is related to the inverse of the source signature $\mathbf{S}$ and the reflectivity of the surface $\mathbf{R}^\cap$ by

$$\mathbf{A} = \mathbf{S}^{-1}\mathbf{R}^\cap. \tag{3.2.2}$$

In expression 3.2.1 the total up-going wavefield is represented as a product of a down-going wavefield ($\mathbf{I}+\mathbf{AP}$) and an up-going wavefield $\mathbf{P}_0$. Alternatively, equation 3.2.1 can be seen as expressing the total wavefield $\mathbf{P}$ as a sum of the primary wavefield $\mathbf{P}_0$ and the surface-related multiple wavefield $\mathbf{P}_0\mathbf{AP}$. Here, by 'primaries' we actually mean all possible events with no bouncing points at the surface. This category includes primaries, internal multiples, converted waves and all other up-going events that are not related with scattering at the surface.

Closed-Loop SRME (CL-SRME) proposes an inversion-based method to obtain the primary response $\mathbf{P}_0$ from the measured data $\mathbf{P}$ in a fully data-driven manner. The strategy is to build an objective function $J = J(\hat{\mathbf{P}}_0, \hat{\mathbf{A}})$ that is minimized when the inversion parameters ($\hat{\mathbf{P}}_0$, $\hat{\mathbf{A}}$) accurately describe the input dataset $\mathbf{P}$ via equation 3.2.1. The solution must fulfill equation 3.2.1 for all frequencies ω. Using a Frobenious L2 norm we can express J as:[1]

$$J = \sum_\omega J_\omega = \sum_\omega ||\mathbf{P} - \hat{\mathbf{P}}_0(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})||^2. \tag{3.2.3}$$

The goal of the inversion within J is to find the set of parameters ($\hat{\mathbf{P}}_0$, $\hat{\mathbf{A}}$) that minimizes the residual. The process starts from the initial value $\hat{\mathbf{P}}_0 = 0$, $\hat{\mathbf{A}} = \mathbf{A}_{SRME}$, where $\mathbf{A}_{SRME}$ is the result of optimally matching the wavefields $\mathbf{P}$ and $\mathbf{PP}$, just like the first iteration of SRME[2]. The initial guesses for $\hat{\mathbf{P}}_0$ and $\hat{\mathbf{A}}$ are

---

[1] We will use the symbols $\hat{\mathbf{P}}_0^{(k)}$ and $\hat{\mathbf{A}}^{(k)}$ to refer to the parameter estimations in the k-th iteration of the algorithm. In order to avoid overloading the notation, we will intentionally drop the iteration superscript whenever possible. Then, the symbols $\hat{\mathbf{P}}_0$ and $\hat{\mathbf{A}}$ will refer to the parameter estimates in the current iteration.

[2] The selection of the initial estimates $\hat{\mathbf{P}}_0 = \mathbf{0}$ and $\hat{\mathbf{A}} \neq \mathbf{0}$ is necessary to obtain a correlation-type of update for $\hat{\mathbf{P}}_0$ in the first iteration of the algorithm. Using equation 3.2.4 the update takes the form $\Delta\mathbf{P}_0^{(1)} = 2\mathbf{P}(\mathbf{I}+\hat{\mathbf{A}}\mathbf{P})^H$. This type of update is important for near offset interpolation, as it produces a back-projection of the multiples into the primaries. Where this back-projection will allow the estimation of the near-offset data.

updated every iteration in such a way that J decreases.

Note that the J function given in equation 3.2.3 is not enough to fully constraint the solution of the inversion problem. Thus an additional constraint needs to be included to solve the inversion in a more unique way. We will discuss this later on.

The algorithm uses a descent method based on a first order Taylor expansion of J_ω to obtain the appropriated descent directions that correspond to the updates of the inversion parameters. For these parameters the update process is done in a flip-flop manner, alternating the estimation of $\mathbf{P}_0$ and $\mathbf{A}$ as the algorithm progresses. The appropriated descent direction $\Delta\mathbf{P}_0$ for the inversion of $\mathbf{P}_0$ is calculated using the extremallity condition of J_ω: $\nabla_{\hat{\mathbf{P}}_0} J_\omega(\mathbf{P}_0, \mathbf{A}) = 0$ to obtain the update $\Delta\mathbf{P}_0 = -\nabla_{\hat{\mathbf{P}}_0} J_\omega(\hat{\mathbf{P}}_0, \hat{\mathbf{A}})$, where the negative sign stands for the selection of a descent direction in J_ω. The gradient for $\mathbf{P}_0$, therefore, is given as:

$$\Delta\mathbf{P}_0 = 2[\mathbf{P} - \hat{\mathbf{P}}_0(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})](\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})^H. \tag{3.2.4}$$

Note that the term in the square brackets represents the current L2 inversion residual, which should be minimized. The $(\hat{\mathbf{A}}\mathbf{P})^H$ terms represent back-projection of the multiples into primary updates. This mechanism will allow multiple information to contribute to primary estimation, even at offsets originally not measured.

Given $\Delta\mathbf{P}_0$, the primary estimate $\hat{\mathbf{P}}_0$ can be updated in every iteration k by:

$$\hat{\mathbf{P}}_0^{(k+1)} = \hat{\mathbf{P}}_0^{(k)} + \alpha\Delta\mathbf{P}_0^{(k)}, \tag{3.2.5}$$

where the step length α is chosen using a line-search, such that the scaled update $\alpha\Delta\mathbf{P}_0$ minimizes the residual $J(\hat{\mathbf{P}}_0 + \alpha\Delta\mathbf{P}_0) \longrightarrow min$. Here, α is given by:

$$\alpha = \frac{\sum_\omega Re(Tr(\mathbf{V}^H\mathbf{K}))}{\sum_\omega Tr(\mathbf{K}^H\mathbf{K})}, \tag{3.2.6}$$

with residual data

$$\mathbf{V} := \mathbf{P} - \hat{\mathbf{P}}_0(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P}) \qquad \text{and} \qquad \mathbf{K} := \Delta\hat{\mathbf{P}}_0(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P}). \tag{3.2.7}$$

In the former equations we have dropped the explicit mention of the iteration index k to avoid unnecessary clutter. The procedure is shown for a steepest descent implementation. For a faster algorithm, a conjugate gradients implementation can be easily deduced from these equations, as will be shown later on.

For sake of simplicity, we will assume in the following that the surface operator $\mathbf{A} = \mathbf{S}^{-1}\mathbf{R}^\cap$ is angle-independent, so the $\hat{\mathbf{A}}$ matrix is diagonal (no directivity in neither $\mathbf{R}^\cap$ nor $\mathbf{S}$). For simplicity we will also assume that the source signatures of all the shots are the same, such that we can write $\hat{\mathbf{A}} = A(\omega)\mathbf{I}$. This assumption can be easily dropped in a practical implementation by making a shot-dependent calculation of $\mathbf{A}$.

Then the inversion equation for $A(\omega)$ becomes that of an optimum filter in the L2 norm, which is typically solved in the time domain by using

$$\hat{A} = \arg\min_{\hat{A}} \sum_{\omega} ||\mathbf{P} - \hat{\mathbf{P}}_0 \hat{A} \mathbf{P}||^2, \tag{3.2.8}$$

where $\hat{A}$ is forced to be a short convolutional operator in the time domain (Verschuur and Berkhout [1997]). Here, $\hat{A}$ can be easily solved by pre-computing $\hat{\mathbf{P}}_0\mathbf{P}$ and then solving a matching problem in the time domain. The fact that $\hat{A}$ is short in the time domain can be thought of as prior-knowledge.

Note that the non-directivity assumption ($\hat{\mathbf{A}}$ being diagonal) is not necessarily met in practice, and it corresponds to a very common simplification present in both EPSI and SRME algorithms, which have been shown to perform successfully in many datasets [Verschuur, 1991; van Groenestijn and Verschuur, 2009b; Baardman et al., 2010]. The objective behind this assumption is to simplify the processing and remove all the unnecessary degrees of freedom, which might map to larger cardinality of the residual null space if left unconstrained in the inversion[3]. In theory effects like ghosts can be included in the forward model [Verschuur, 2013; Vrolijk and Verschuur, 2015], but here we assume ghosts effects are removed in advance.

For the estimation of $A(\omega)$, no updating process is necessary. Instead $\hat{A}(\omega)$ is obtained in every iteration using the present value of $\hat{\mathbf{P}}_0$ and equation 3.2.8, the latter corresponding to the minimum energy constraint on primaries. In the following chapter we will relax this assumption by introducing a sparsity constraint in the primary estimate. Note that in order to solve the inversion problem for $\hat{\mathbf{P}}_0$ and $\hat{\mathbf{A}}$, two independent constraints are required, namely equation 3.2.1 and 3.2.8 or equation 3.2.1 and a sparsity constraint.[4]

---

[3] The residual null-space is the set of model space vectors $(\hat{\mathbf{P}}_0, \hat{\mathbf{A}})$ that make the residual zero. A larger cardinality of this set implies more non-uniqueness in the solutions. Note for example that the undesirable model space vector $(\hat{\mathbf{P}}_0, \hat{\mathbf{A}}) = (\hat{\mathbf{P}}, \mathbf{0})$ makes the residual in equation 3.2.3 equal to zero. Thus, this solution vector belongs to the global minimum set of the objective function. We will address this problem in the next chapter by introducing an extra constraint in the inversion.

[4] Unfortunately, equations 3.2.8 and 3.2.3 are not enough to obtain a unique solution. This implies that the present scheme might converge to solutions with some small degree of multiple leakage in the primaries. We will attenuate this problem in the following chapter by introducing extra constraints over J, together with smart starting points for $\hat{\mathbf{P}}_0$ and $\hat{\mathbf{A}}$. However, even then, the solution will remain non-unique. Multiple-leaking solutions are present up to some extent in all current multiple prediction algorithms.

3.3 Conjugate gradient implementation

Given the parameterization of CL-SRME in $\hat{\mathbf{P}}_0$ and $\hat{\mathbf{A}}$, the algorithm might become vulnerable to instabilities as the ones described by Nekut and Verschuur [1998]. These instabilities are generated when the descent scheme finds itself trapped between two instable solutions in the vicinity of an objective function minimum. Then the solution jumps back and forth between two undesirable model space solutions without efficiently reaching the minimum. In order to avoid this, we use a conjugate gradient approach such that the updates are always orthogonal (or at least as orthogonal as possible) and less dependent on the local structure of the objective function. A conjugate gradient approach is also more computationally efficient.

A Fletcher-Reeves [Aster et al., 2013] implementation of the conjugate gradient algorithm can be easily obtained by including

$$\Delta\hat{\mathbf{P}}_{0,CG}^{(k+1)} = \Delta\hat{\mathbf{P}}_{0}^{(k+1)} + \beta\Delta\hat{\mathbf{P}}_{0,CG}^{(k)}, \tag{3.3.9}$$

and

$$\beta = \frac{Tr(\Delta\hat{\mathbf{P}}_{0}^{(k+1)}\Delta\hat{\mathbf{P}}_{0}^{(k+1)})}{Tr(\Delta\hat{\mathbf{P}}_{0}^{(k)}\Delta\hat{\mathbf{P}}_{0}^{(k)})}. \tag{3.3.10}$$

The condition $\Delta\hat{\mathbf{P}}_{0,CG}^{(0)} = \Delta\hat{\mathbf{P}}_{0}^{(0)}$ will provide a starting point for the algorithm. From here it is easy to follow the standard implementation of the conjugate gradient algorithm [Tarantola, 1987; Aster et al., 2013].

Figure 3.1 depicts a typical CL-SRME example from a simple two-layered synthetic model (shown in figure 4.1). The CL-SRME algorithm separates the multiples after 20 iterations.

3.4 Comparison with EPSI and SRME

CL-SRME shares the same forward model and the same parameterization as SRME, so in many ways CL-SRME can be seen as a 'closed-loop' or 'inversion-oriented' version of the SRME algorithm. Using the same reasoning, the SRME algorithm is thought to operate in a "input-output" approach which doesn't involve a feedback process. This type of algorithm can also be called an "open-loop" method (to mark the contrast with respect to the "closed-loop" methods, i.e. EPSI and CL-SRME). Besides, their resemblance, fundamental differences between CL-SRME and SRME can be observed in the way the inversion estimates are obtained (inversion series). The CL-SRME algorithm involves both convolution and correlation type of updates,

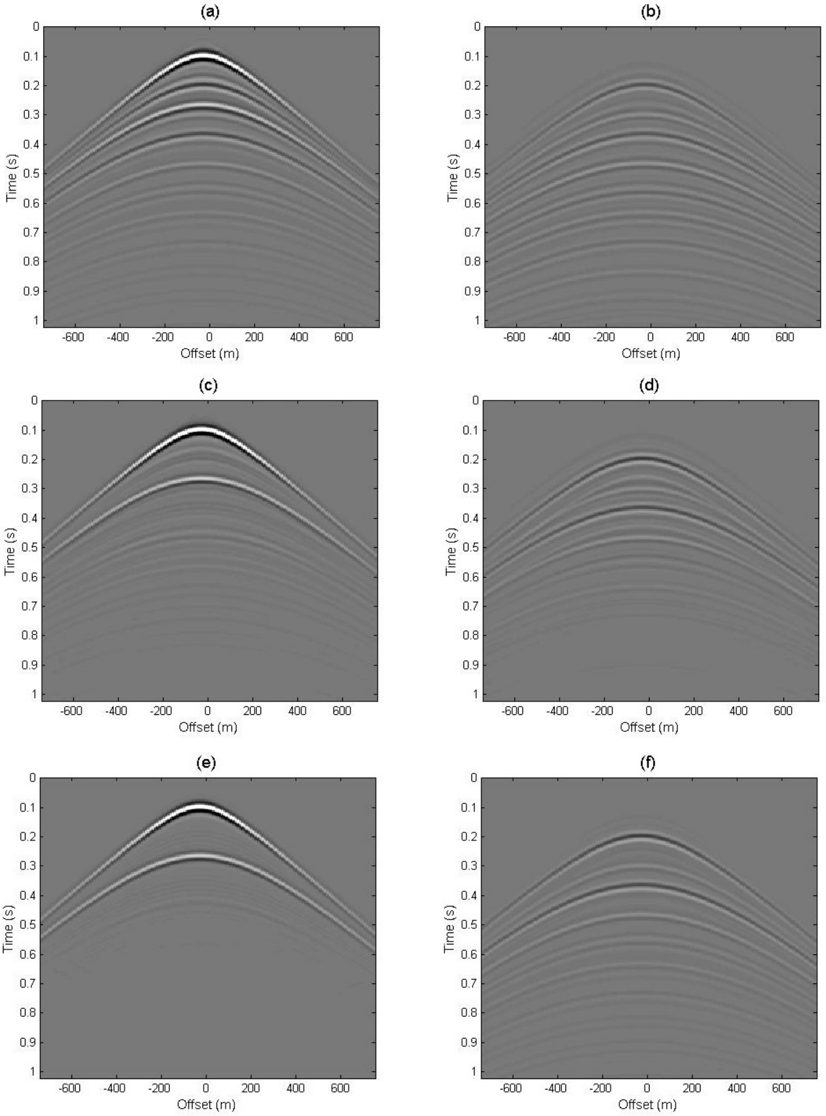

Figure 3.1: *Closed Loop-SRME algorithm used for multiple elimination. The model used to generate the data is the same one as in figure 4.1. To illustrate the procedure the primaries and multiples after the (a-b) first, (c-d) third and (e-f) and 20th iteration respectively.*

see equation 3.2.4, while SRME involves only convolution-type of updates. Correlation type of updates are important when considering interpolation in the scheme.

Introducing a gradient-based process also opens the room for several upgrades and extensions to the theory presented here. In the next chapter we will discuss an important extension that will allow the CL-SRME algorithm to incorporate reconstruction and multiple estimation in the same process. One of the key changes here, will be the addition of an additional constraint (sparsity) to fully resolve J without the need of equation 3.2.8. Note that if the minimum energy assumption on $\mathbf{P}_0$ is used in CL-SRME as additional constraint, its solution will converge to the SRME solution.

In a similar way as done in EPSI (van Groenestijn and Verschuur, 2009), in CL-SRME the primaries are iteratively estimated such that they, together with their associated multiples, explain the input data. However, in CL-SRME this is done in a quite different way. Here, the input data $\mathbf{P}$ is parameterized via the surface operator $\mathbf{A}$ and the primary wavefield $\mathbf{P}_0$, which is consistent with the parameterization in SRME [Verschuur et al., 1992]. EPSI on the other hand describes the input data $\mathbf{P}$ in terms of the source signature $\mathbf{S}$ and the primary impulse response $\mathbf{X}_0$. Mathematically, these two parameterizations are equivalent, but practically they involve quite different wavefields, thus different inversion strategies and different results are expected. Note, for instance, that in contrast to $\mathbf{P}_0$, $\mathbf{X}_0$ is expected to be spiky, thus justifying a sparsity constraint on $\mathbf{X}_0$. The relationship between the inversion parameters ($\mathbf{X}_0$, $\mathbf{S}$) in EPSI, and ($\mathbf{P}_0$, $\mathbf{A}$) in CL-SRME is given by: $\mathbf{P}_0 = \mathbf{X}_0\mathbf{S}$ and $\mathbf{A} = \mathbf{S}^{-1}\mathbf{R}^{\cap}$. More about EPSI and CL-SRME can be found in chapter 9.

3.5 Conclusions

In this chapter we have introduced a new method for multiple estimation using inversion that we have called Closed-Loop SRME. Even though by itself this new method doesn't provide better results than EPSI or SRME, we can use its mathematical description to generate useful extensions to the algorithm shown here. In many ways, the CL-SRME algorithm can be regarded as an inversion-based SRME algorithm. In the following chapters we will use the theory shown here to achieve data interpolation together with multiple separation. This development will have interesting results and consequences for more realistic 2D and 3D geometries.

The inversion methodology in CL-SRME also allows other extensions to the traditional SRME algorithm. Some direct extensions might be: multiple elimination on blended data, applications to passive data and internal multiple elimination with missing data.

4

The focal transform

In this chapter we will discuss the focal transform concept from both the theoretical and the practical point of view. We will start describing the forward propagation operators from the Raleigh integral concept. Then we will describe the inversion of those operators using the matched-filter approach. Later, some explicit formulas to calculate the operators will be given. Then we can describe the the focal transform as a two-sided inverse propagation, and describe its properties and possible uses. In the next chapter the focal transform will be integrated with the CL-SRME algorithm.

4.1 Propagation operators

The extrapolation of an acoustic wavefield inside a source-less, inhomogeneous and isotropic half-space medium is given by the Rayleigh integral, which in 2D reads [Wapenaar and Berkhout, 1989][1]:

$$P(x, z; \omega) = \int_{\mathbb{R}} W(x, z, x_0, z_0; \omega) P(x_0, z_0; \omega) dx_0, \tag{4.1.1}$$

where P is the wavefield to extrapolate, W is the propagation operator, ω is the frequency and $\{(x_0, z_0), (x, z)\}$ are the initial and target extrapolation points. Note that every point in the extrapolated wavefield is constructed by the summation of all lateral contributions in the initial wavefield. The former equation can be extended to 3D by adding an extra lateral dimension y:

---

[1] In order to provide a better link with other branches of Physics, we will start by describing our propagation operators in a continious fashion.

$$P(x, y, z; \omega) = \iint_{\mathbb{R}^2} W(x, y, z, x_0, y_0, z_0; \omega) P(x_0, y_0, z_0; \omega) dx_0 dy_0, \tag{4.1.2}$$

now (x_0, y_0, z_0) is the initial point and (x, y, z) is the target point for extrapolation. Note the summation over all secondary source points (lateral positions x_0 and y_0) according to Huygens' principle.

In practical applications, the wavefield is always measured in a discrete manner, by placing receivers at the surface. So, it is convenient to discretize equations 4.1.1 and 4.1.2 such that we can apply them in practical circumstances. After discretization equation 4.1.1 reads

$$P(x_i, z; \omega) = \sum_{k=1}^{N_x} W(x_i, z, x_{0,k}, z_0; \omega) P(x_{0,k}, z_0; \omega) \Delta x_0. \tag{4.1.3}$$

Similary, equation 4.1.2 reads

$$P(x_i, y_j, z; \omega) = \sum_{k=1}^{N_x} \sum_{l=1}^{N_y} W(x_i, y_j, z, x_{0,k}, y_{0,l} z_0; \omega) P(x_{0,k}, y_{0,l}, z_0; \omega) \Delta x_0 \Delta y_0, \tag{4.1.4}$$

where $\{x_i, y_j\}_{(i,j)}$ are now the discrete target positions, $\{x_{0,k}, y_{0,l}\}_{(k,l)}$ are the discrete initial positions, $(\Delta x, \Delta y)$ are the respective sampling intervals in each direction and (N_x, N_y) are the number of sampling points in each dimension.

Note that equation 4.1.3 and 4.1.4 apply now only for the discretized versions of the field $P(x, y, z; \omega)$. Care must be taken while spatially discretizing a wavefield, as aliasing might occur if the wavefields are too under-sampled.

4.2 Inverse operators

Together with the forward operators, expressions for their inverses are also necessary in practical applications. In principle, the formulas for the inverse operators can be obtained by directly calculating W^{-1}, however, there are two important reasons to avoid this:

[1] The exact formulation of the W operators for an arbitrary medium typically depends on the exact medium parameters, which are unknown in practical applications.

[2] The inversion of the forward extrapolation operators is typically unstable.

Fortunately, if we concentrate on inverse operators to remove propagation effects, only the low-frequency variations in the material parameters (i.e. the so-called macro-model) are needed [see Wapenaar and Berkhout, 1989]. In this case we can find an approximate inverse of the forward operator by just applying complex conjugation via:

$$W(x, z, x_0, z_0; \omega)^{-1} \approx W(x_0, z, x, z_0; \omega)^* \tag{4.2.5}$$

and

$$W(x, y, z, x_0, y_0, z_0; \omega)^{-1} \approx W(x_0, y_0, z, x, y, z_0; \omega)^*. \tag{4.2.6}$$

The former approximation for calculating the inverse of W is known as the *matched filter approach* [Wapenaar and Berkhout, 1989]. This approximation will limit the maximum attainable resolution.

4.3 Matrix form for multiple experiments

As discussed in Chapter 2, the discretized version of the W operators (equations 4.1.3 and 4.1.4) can be organized in a matrix in order to account for the complete seismic experiment with several sources and receivers in the acquisition grid. Then we can express the set of propagation operators to a given depth level as $\mathbf{W}$ and their approximate inverses as $\mathbf{W}^H$ (equations 4.2.5 and 4.2.6). Using reciprocity it's possible to change the propagation direction of the $\mathbf{W}$ operator from down-going to up-going, this is equivalent to take $\mathbf{W}^T$ when an upward propagation is needed. Due to the convolutional structure of wavefield propagation, the matrices $\mathbf{W}$ represent spatial convolutions. If the propagation medium is laterally invariant, then $\mathbf{W}$ is a Toeplitz matrix. Laterally variant media typically have $\mathbf{W}$ matrices with band-like structures, representing generalized convolutions. The matrices $\mathbf{W}$ are related with the more traditional monopolar Green's functions $\mathbf{G}(z_0, z_m)$ via the scalar expression [Wapenaar and Berkhout, 1989]:

$$W(x, y, z_m, x_0, y_0, z_0) = \frac{2}{\rho}\frac{\partial G(x, y, z_m, x_0, y_0, z_0)}{\partial z_0}, \tag{4.3.7}$$

where ρ is the density. The derivative in the vertical direction z makes explicit the fact that the W operator propagates dipole sources, as opposed to the G operator, which is related to monopole sources[2]. As seen from equation 4.3.7, the W operators

---

[2] Note that in Rayleigh theory, only dipolar Greens functions are necessary to extrapolate the wavefields. This is due to the following facts: (1) the integration is done over a flat surface and (2) we have separated up-going waves from down-going waves. For more information see Gisolf and Verschuur [2010].

Table 4.1: *Dipole responses for an isotropic homogeneous acoustic medium. Associated W operators can be obtain from these expressions through a $2/\rho$ scaling factor (equation 4.3.7). After discretization each of these values in the $x-\omega$ domain will correspond to entries in the* $\mathbf{W}$ *matrix.*

	3D dipole	2D dipole
$x-\omega$ exact	$\frac{e^{-ikr}}{r}$	$-ik\pi\cos\pi H_1^{(2)}(kr)$
$x-\omega$ far-field	$\frac{ik}{r}\cos\phi e^{-ikr}$	$\sqrt{\frac{2\pi ik}{r}}\cos\phi e^{-ikr}$
$k_x-\omega$ exact	$2\pi sign(z)e^{-ik_z\lvert z\rvert}$	$2\pi sign(z)e^{-ik_z\lvert z\rvert}$

are calculated from the dipolar Green's function with a $2/\rho$ scaling.

4.4 Explicit formulas for W

Finding explicit formulas for the propagation operators can be done analytically in the case of a homogeneous medium, but there are no analytical solutions for the general inhomogeneous case. Some expressions can be derived in the case of a smoothly varying medium [Gisolf and Verschuur, 2010]. In practice, however, the W's for arbitrary media typically have to be computed numerically. In the case of an homogeneous mediums, expressions in table 4.1 can be used [Berkhout, 1982].

In table 4.1, r describes the distance between source and receiver, ϕ represents the angle between the source-receiver vector and the z-axis, and $k=\omega/c$ is the magnitude of the wave-number vector $\mathbf{k}=(k_x,k_y,k_z)$ (with c the propagation velocity). Note that the k_z is defined differently for the 2D and 3D cases. In 2D $k_z=\sqrt{k^2-k_x^2}$, whereas in 3D $k_z=\sqrt{k^2-k_x^2-k_y^2}$. The function $H_1^{(2)}$ is a Hankel function of the second kind.

Though simple, these expressions can be used to construct approximate propagation operators (for an inhomogeneous medium) given some NMO velocities. This might look like a rough approximation to the real propagators, but as we will see in Chapter

5, for the current work, only this limited knowledge is required. The reasoning behind this lies in the fact that we will be only interested in removing the phase (or equivalently, the travel-time) information from selected events.

4.5 The focal transform

In the reminder of this chapter we will assume a given set of propagation operators $\mathbf{W}$ to be known, either as prior information or estimated via the expressions in table 4.1. The user would typically calculate the propagation operators $\mathbf{W}$ to match the most energetic reflectors in the data.[3]

The focal transform can be separated into two main types of transforms: the single-level transform and the multi-level transform.

■ 4.5.1 Single-level focal transform

The single-level focal transform can be understood as a redatuming operation to a given layer. It comprises the propagation of a given wavefield measured at the surface to given depth level or target layer. In principle, this operation can be loosely seen as taking the whole acquisition, both sources and receivers into the subsurface up to a given depth level, and repeating the seismic experiment there (see figure 4.1). The transform domain here is just the redatumed dataset, and it can be calculated numerically.

Assuming that $\mathbf{P}$ is the dataset to focus and $\mathbf{X}$ is the focused dataset (also known as the focal domain), then the forward focal transform is given by[4]

$$\mathbf{X} = \mathbf{W}^{*}\mathbf{P}\mathbf{W}^{H}, \tag{4.5.8}$$

and the inverse transform is given by

$$\mathbf{P} = \mathbf{W}^{T}\mathbf{X}\mathbf{W}. \tag{4.5.9}$$

Note that if $\mathbf{W}$ and $\mathbf{P}$ have the same spectral support (i.e. no frequencies or wavenumbers are being filtered out by the focusing) then $\mathbf{P} = \mathbf{W}^{T}\mathbf{W}^{*}\mathbf{P}\mathbf{W}^{H}\mathbf{W} = \mathbf{P}$.

---

[3] To match a given $\mathbf{W}$ operator with a given reflector response in our dataset we just need to match the travel-time curves of both events. This can be easily done by matching the velocities (described by $k = \omega/c$) and reflector depths (given by z) of our operator defined by table 4.1.

[4] Here we are using the relationship $\mathbf{W}^{-1} \approx \mathbf{W}^{H}$, which follows from the matched-filter approximation in equations 4.2.5 and 4.2.6.

This implies that the single-level focal transform is fully invertible.[5]

The single-level focal transform will be used in this thesis to focus the first water bottom primaries. This is done in order to produce better near-offset interpolations for the shallow section.

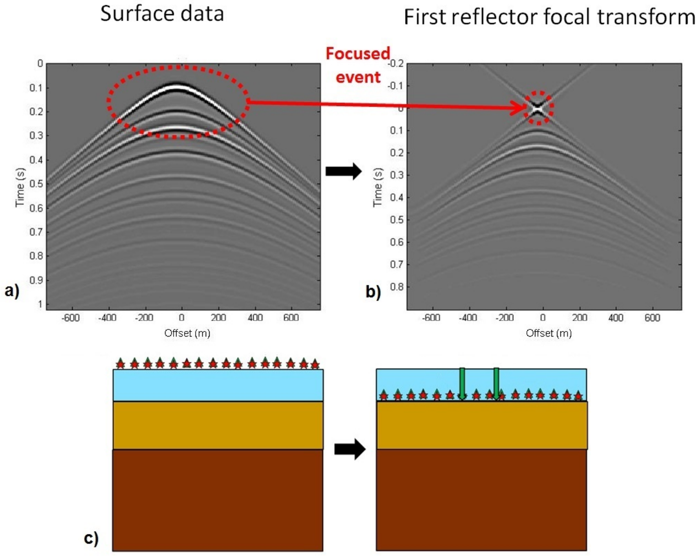

Figure 4.1: *The data acquired at the surface (a) is focal-transformed via the application of the propagation operators related to the first layer. In the transformed dataset (b), the primary event related with the first layer appears concentrated at t=0. The focal transform has mapped the information of the first reflection into a compressed event in the transformed domain. (c) Application of the focal transform related to one particular layer is equivalent to redatuming the dataset to that layer.*

■ 4.5.2 Multi-level focal transform

The multi-level focal transform can be understood as a redatuming operation to a given set of levels, where the energy from the input data is distributed over this set of depth levels. Here, we express the input wavefield in terms of a linear combination of transform domains and propagation operators. Because there are an infinite

---

[5] In general $\mathbf{P} = \mathbf{W}^T\mathbf{W}^*\mathbf{P}\mathbf{W}^H\mathbf{W} = \mathbf{\Phi}^T\mathbf{P}\mathbf{\Phi}$, where $\mathbf{\Phi} = \mathbf{W}^H\mathbf{W}$ is the "resolution function" [Wapenaar and Berkhout, 1989; Berkhout et al., 2001]. This function provides information about the bandwidth of the maximum obtainable spatial resolution in the inversion, which can be proven to be approximately $6/5k_{max}$, with k_{max} the maximum wavenumber in the data.

amount of possible decompositions of the input wavefield in terms of the output variables, typically the output transform domains are calculated via inversion, by exploiting some additional knowledge or assumption about the focal domains (very often, sparsity).

Assuming a set of z_m ($m \in [1, M]$) reflectors of interest, we can write the forward transform to a given layer z_m approximately as

$$\mathbf{X}_m \approx \mathbf{W}_m^* \mathbf{P} \mathbf{W}_m^H, \tag{4.5.10}$$

where the approximation stands for the fact that each $\mathbf{X}_m$ needs to be calculated from inversion (we will do this later by finding the set of $\mathbf{X}_m$ values that minimize the functional given by equation 5.2.3).

The inverse transform can be deduced easily from the condition that all the focal transforms must add up to the original data

$$\mathbf{P} = \sum_{m=1}^{M} \mathbf{W}_m^T \mathbf{X}_m \mathbf{W}_m. \tag{4.5.11}$$

Actually, given a set of $\mathbf{W}_m$ operators, this expression can also be considered the definition of the multi-level focal transform.

Figure 4.2 depicts the decomposition of a dataset in terms of two focal domains. Here a sparsity constraint on the focal domains is assumed in order to constraint the inversion (see also section 5.2). Note that only primaries are used.

■ 4.5.3 Properties and uses

Now we will list some of the most important properties and uses of the focal transform.

[1] *Event compression*: Reflection events, specifically in the shallow part of the data can be neatly compressed with the focal domain. A single element in the focal domain maps to a full reflection event (a difractor response) in the direct domain. The AVO behavior is then contained in the event tails. Figure 4.3 compares the focal transform with other types of transforms.

[2] *Robustness*: The focal transform is robust against focusing velocity errors. This is show in figure 4.4, in which the focusing velocity is perturbed and the resulting error in the data reconstruction is calculated.

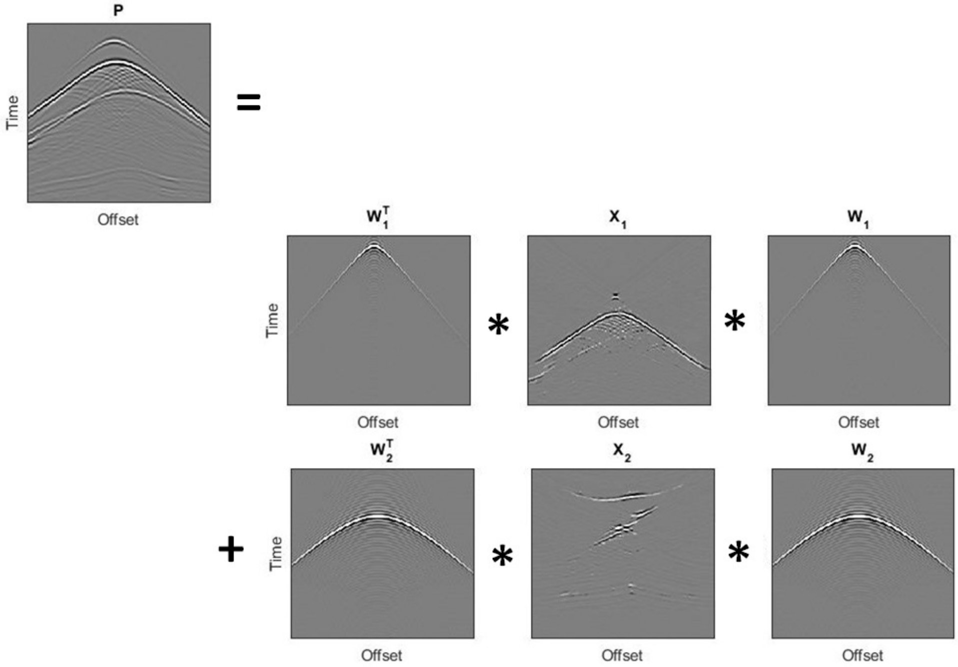

Figure 4.2: *Schematic representation of the decomposition of the primary dataset generated by figure 1.4 in terms of two focal domains. Sparsity is assumed when calculating the shown focal domains. Here* $\mathbf{P}$ *represent the data,* $\mathbf{W}_1, \mathbf{W}_2$ *the propagation operators and* $\mathbf{X}_1, \mathbf{X}_2$ *the focal domains. Each panel shows one column of the corresponding matrix in the x-t domain.*

[3] *Redatuming*: As discussed above, the focal transform can be used for redatuming up to a given target layer. A muting line needs to be imposed to eliminate possible anti-casual events or any other form of undesired energy (see figure 4.1).

[4] $\omega - k_x$ *filtering*: The focal transform, if used in a focusing/defocussing manner (i.e. apply the forward transform and then the inverse), acts as a dip-filter, with the propagation velocity providing the ”cutting-line” in the $\omega - k_x$ domain (see figures 4.5 and 4.6). This property might be important when designing propagation operators, as the $\omega - k_x$ filtering might be undesired.

[5] *Other uses*: Due to the previously mentioned properties, the focal transform can be used for many other purposes, including: interpolation ([Kutscha and Verschuur, 2012]), deblending ([Kontakis and Verschuur, 2006]), denoising (from the $\omega - k_x$ filtering), and also as a tool for primary/multiple separation.

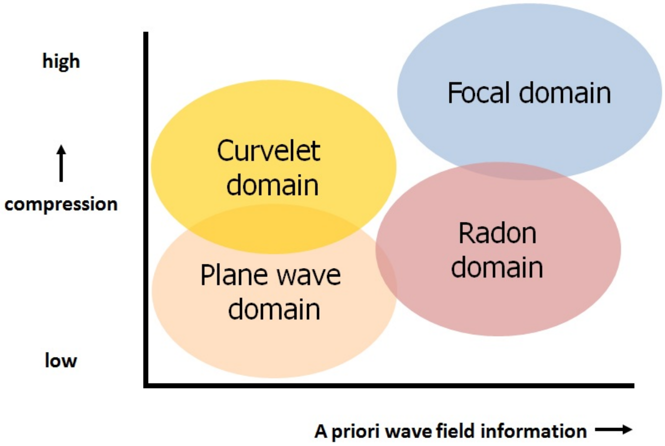

Figure 4.3: *Schematic comparison between some of the major transform domains used in the seismic literature for data reconstruction. A comparison is done between the different transform with respect to the amount of compression that they have and with respect to the a priori wavefield information they require. The compression can be understood as the amount of parameters needed to describe a given dataset in the transform domain. The smaller, the more compressed. The a priori information is the information that we need to include in order to carry the transform, and is not directly contained in the data (e.g. NMO times and velocities).*

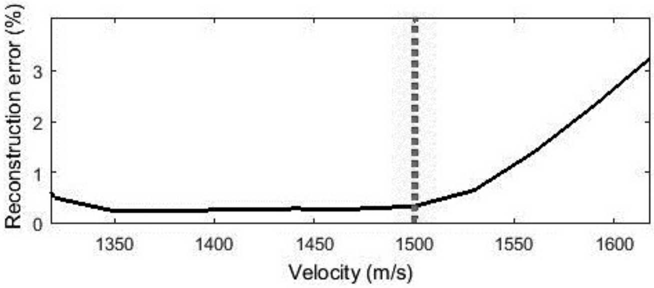

Figure 4.4: *Reconstruction error on a simple two-layer example (figure 4.1) as a function of the associated focal operator velocity. The true velocity* $c = 1500m/s$ *is shown in the dotted line. As we can see, errors in the operator propagation velocity do not produce big errors in the reconstruction estimates. The higher velocities tend to be more sensitive than the lower velocities, due to the mentioned* $\omega - k_x$ *filtering effect.*

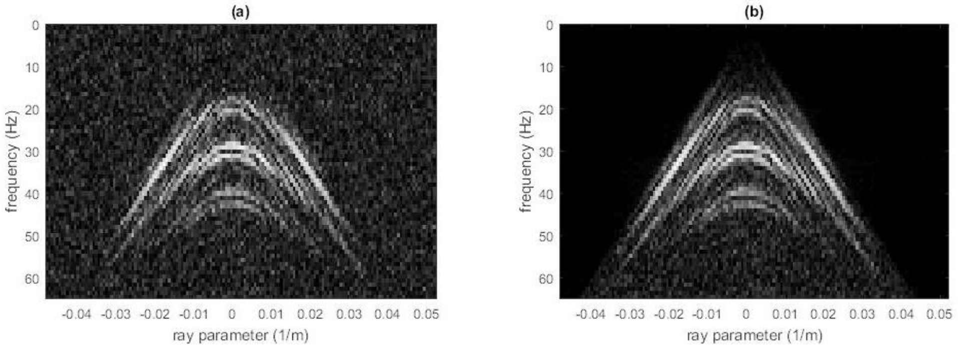

Figure 4.5: *$\omega - k_x$ filtering via the focal domain. (a) Input $\omega - k_x$ spectra with a 10% noise level and (b) $\omega - k_x$ of the input after applying the forward single-level focal tranform followed by its inverse tranform. As we can see a large amout of the initial noise has been removed. The propagation velocity of the $\mathbf{W}$ operators here is choosen so that it corresponds to the minimum velocity in the data.*

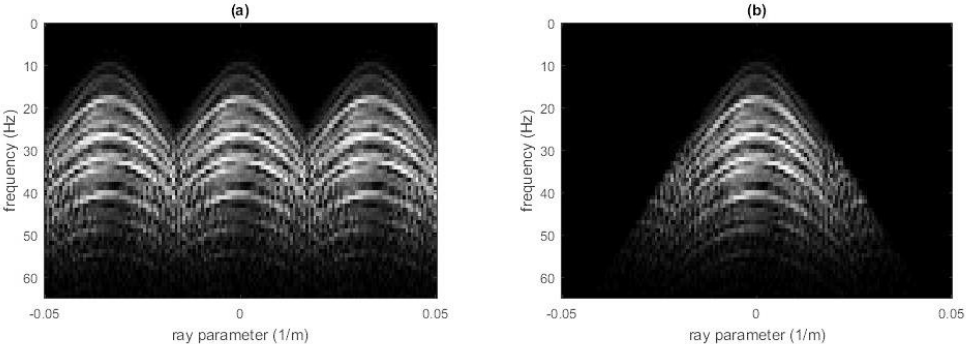

Figure 4.6: *$\omega - k_x$ filtering via the focal domain. (a) $\omega - k_x$ spectra of dataset with a 1:3 receiver under-sampling factor and (b) $\omega - k_x$ spectra of the input after applying the forward single-level focal transform followed by its inverse transform. As we can see a large amount of the initial noise has been removed. The propagation velocity of the $\mathbf{W}$ operators here is chosen so that it corresponds to the minimum velocity in the data. The sampling gaps (regular undersampling) in the $x - t$ domain map to blending noise in the $\omega - k_x$ domain, removing the blending noise produces interpolation (however, there will be aliasing artefacts at the higher frequencies).*

5

Focal Closed-Loop SRME

In this chapter we will introduce an extension of the Closed-Loop SRME algorithm presented in the previous chapter to include extended data interpolation capabilities. We will re-parameterize the CL-SRME algorithm using the focal transform in order to use some its interpolation capabilities. We will illustrate the theory with some synthetic and field data examples.

5.1 Introduction

In order to use CL-SRME to accurately predict multiples in a coarsely sampled dataset (e.g. for 3D application) a new parameterization must be adopted to allow CL-SRME to efficiently reconstruct data over large gaps. In our new approach we will use the focal transform [Berkhout and Verschuur, 2006; Kutscha et al., 2010; Kutscha and Verschuur, 2012], which aims at focusing primary reflections into localized events. The focal domain can compress the primary energy, a property that will be useful when separating primary signals from under-sampling noise in the focal domain. A sparse norm regularization is used to remove the under-sampling noise from the estimations and to eliminate the redundancy in the parameterization. Note that the multi-level focal transform is a redundant parameterization because every focal domain is as big as the input data. This implies that in a N-level focal transform, the number of inversion parameters in actually N times larger than the number of measurement points in the input dataset. This redundancy can be removed by applying an extra constraint.

Other parameterizations using the curvelet transform [Candes et al., 2006; Herrmann et al., 2007, 2008; Herrmann and Verschuur, 2008] or the wavelet transform [Liu et al., 1995] are also possible candidates for a sparse representation of the seismic data. Our choice of the focal transform is justified by the fact that, due to its focusing characteristics, the focal transform is able to compress the energy of highly curved events into localized events, making it ideal for shallow water layer applications, where the events to reconstruct are strongly curved in the near offsets. For deeper data and flatter events, other parameterizations like curvelets or linear Radon might become more beneficial.

In the following sections we will provide some theoretical explanations along with some practical 2D examples to explain the CL-SRME algorithm including the focal domain parameterization. Special attention will be given to near-offset reconstruction in shallow-water environments.

5.2 Focal Closed-Loop SRME

Closed-Loop SRME proposes an inversion-based method to obtain the primary impulse response $\mathbf{P}_0(\omega)$ from the measured data $\mathbf{P}(\omega)$ in a fully data-driven manner. The strategy is to build an objective function $J = J(\hat{\mathbf{P}}_0, \hat{\mathbf{A}})$ that is minimized when the parameter estimates $(\hat{\mathbf{P}}_0, \hat{\mathbf{A}})$ accurately describe the input dataset $\mathbf{P}$. Here the symbols $(\hat{\mathbf{P}}_0, \hat{\mathbf{A}})$ represent the estimates of the true parameters $(\mathbf{P}_0, \mathbf{A})$. A sparsity constraint L_S on the primaries can be added to the process to ensure the minimum amount of multiple leakage in the estimates. Note that this constraint must be applied in the time domain. Using the Frobenious L2 norm together with the sparsity promoting norm L_S (yet to be defined) we can express J as

$$J = \sum_{\omega} ||\mathbf{P} - \hat{\mathbf{P}}_0(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})||^2 + \lambda L_S(\hat{\mathbf{P}}_0), \tag{5.2.1}$$

where each data matrix contains one frequency component and L_S acts over the full data-cube of $\hat{\mathbf{P}}_0$ (see figure 2.1). Using J in a least-squares iterative inversion procedure, CL-SRME aims to estimate the primary wavefield $\mathbf{P}_0$ by a large scale inversion in which the inversion parameters $\hat{\mathbf{A}}$ and $\hat{\mathbf{P}}_0$ are estimated such that the input data $\mathbf{P}$ is explained. In this way the multiples help to estimate the primaries. Note that CL-SRME given by equation 5.2.1 will describe exactly the same problem as traditional SRME under the condition $L_S = L_2$, where the L_2 term represents the minimum-energy condition over the primaries. In that case, both algorithms tend to converge to the same solution. We will describe L_S later in this chapter.[1]

In order to use CL-SRME to remove multiples from datasets with severe undersampling (which is typical for 3D geometries), we look for a way to very effectively

---

[1] As we will see in equation 5.2.3, $L_S(\hat{\mathbf{P}}_0) = \sum_m ||\hat{\mathbf{x}}_m||_S$. This term applies a sparsity constraint on the focal domains of the primaries $\hat{\mathbf{x}}_m$. The terms $\hat{\mathbf{x}}_m$ contain the primary information represented in a sparse domain.

parameterize our primaries, such that both data reconstruction and multiple removal are achieved. Our goal is to use CL-SRME to both remove the multiples and reconstruct big portions of missing data. In order to create a proper reconstruction algorithm, the following conditions must be met:

[1] The primary wavefield must be represented in a transform domain with the smallest possible number of non-zero transform parameters, such that their information is compressed and efficiently represented. Signals arising from multiple reflections do not need to be represented in this way, as they simply follow any parameterization of the primaries $\mathbf{P}_0$ via $\mathbf{P}_0\mathbf{AP}$. Typically, the data gaps are mapped to artefacts in the transform domain. Due to their lack of coherency, these artefacts are often poorly compressed and are usually seen as aliasing noise in the transform domain.

[2] A parameter selection method in the transform domain must be implemented to separate the parameters representing the primary signals from the parameters accounting for the aliasing noise. Depending on the geometry of the transform domain these methods might vary, where typical examples might be: picking, thresholding, regularization, etc.

To fulfill the first condition we propose the use of the focal transform [Berkhout and Verschuur, 2006; Kutscha et al., 2010; Kutscha and Verschuur, 2012; Lopez and Verschuur, 2013] applied to the primary data. In this description, we will make use of some rough NMO-velocity information to create propagation operators $\mathbf{W}$ that will allow us to back-propagate the wavefields. If we select these operators to take the reflected energy back to where it came from (i.e. a strongly reflective layer), then the scattered energy will be focused in one small event in time and space. Focusing brings strong compression and coherency in the data, in such a way that it will easily allow reconstruction over large data gaps.

For the multi-level focal transform we consider the M strongest reflectors in the data and extract their approximate RMS velocities. Then, we construct a set of propagation operators from the surface to the selected reflectors. These operators have no need to be exact, as they are expected to provide only an approximation of the travel time information in our data. The set of operators for the whole seismic experiment are written as $\{\mathbf{W}_m\}_{m\in[1,M]}$, and are expected to extrapolate the data from the surface to a set of depth levels $\{z_m\}_{m\in[1,M]}$. These $\mathbf{W}_m$ will be later used to parameterize the primary estimates $\hat{\mathbf{P}}_0$ into a set of focal domains $\hat{\mathbf{X}}_m$ such that

$$\hat{\mathbf{P}}_0 = \sum_{m=1}^{M} \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m. \tag{5.2.2}$$

Note that each focal domain $\hat{\mathbf{X}}_m$ represents a full data volume (sources and receivers) for one frequency component. Thus the focal transform is redundant. The relationship between the $\hat{\mathbf{X}}_m$ term of equation 5.2.2 an the $\hat{\mathbf{X}}_0$ in EPSI is given by $\hat{\mathbf{X}}_0 S = \sum_{m=1}^{M} \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m$. So actually $\hat{\mathbf{X}}_0$ can be interpreted as a special case of equation 5.2.2 in which $M = 1$, the source wavelet is a Dirac delta function and the propagation operators are chosen to be matrix identities (target propagation level at the surface).

Despite their similar notation, the $\hat{\mathbf{X}}_m$ term of equation 5.2.2 is in principle not directly related with the $\hat{\mathbf{X}}_0$ in EPSI (which represents the primary impulse response).

To fulfill the second condition above we require a method to eliminate the aliasing noise from the transform domains $\hat{\mathbf{X}}_m$. In addition, with such a condition, the redundancy in the transform would be removed (e.g. non-sparse representations of events will be removed/attenuated to favor the sparser, and typically more unique, representations). We will use a sparsity-promoting regularization norm $||\cdot||_S$ applied to the focal domains $\hat{\mathbf{X}}_m$ to do this. With this extra constraint we drive the algorithm towards a sparse representation of the focal domains. This condition will remove the energy from the aliasing artefacts to concentrate it in the main $\hat{\mathbf{X}}_m$ primary events, thus obtaining a fully sampled representation of $\hat{\mathbf{P}}_0$ via equation 5.2.2.

Using equations 5.2.1 and 5.2.2, the current algorithm can be described by minimizing the following objective function:

$$\begin{aligned} J &= J^{(LS)} + J^{(reg)} \qquad (5.2.3) \\ &= \sum_{\omega} ||\mathbf{P} - \sum_{m=1}^{M} \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q}||^2 + \lambda \sum_{m} ||\hat{\mathbf{x}}_m||_S, \end{aligned}$$

with

$$\mathbf{Q} = \mathbf{I} + \hat{\mathbf{A}}\mathbf{P}, \qquad (5.2.4)$$

where $\hat{\mathbf{x}}_m$ is the inverse Fourier transform of the data-cube[2] of $\hat{\mathbf{X}}_m$ (see figure 2.1) and λ is a user-defined regularization constant (typically $\lambda \approx 10^{-2}$). The norm $||\cdot||_S$ represents any sparsity-promoting norm of preference (e.g. L1), which applies to every element in $\hat{\mathbf{x}}_m$. Note that the regularization norm is applied in the time domain, so sparsity is enforced in that domain.

The former equations represent the forward model for the unknown parameters $(\hat{\mathbf{X}}_m, \hat{\mathbf{A}})$, such that $J = J(\hat{\mathbf{X}}_m, \hat{\mathbf{A}})$. Using gradient-based inversion we can find a set of model parameters $(\hat{\mathbf{X}}_m, \hat{\mathbf{A}})$ that minimizes our objective function J.

The inversion updates for the focal domains $\hat{\mathbf{X}}_m$ are given by $\Delta\hat{\mathbf{X}}_m^{(LS)} = -\nabla_{\hat{\mathbf{X}}_m} J^{(LS)}$ and $\Delta\hat{\mathbf{x}}_m^{(reg)} = -\nabla_{\hat{\mathbf{x}}_m} J^{(reg)}$, which leads to (proof in appendix B.1.2):

---

[2] This means that the quantity $\hat{\mathbf{x}}_m$ is actually a 3D matrix.

$$\Delta \mathbf{X}_m = \Delta \mathbf{X}_m^{(LS)} + \Delta \mathbf{X}_m^{(reg)}, \tag{5.2.5}$$

in which $\Delta \mathbf{X}_m^{(LS)}$ and $\Delta \mathbf{x}_m^{(reg)}$ are given by

$$\Delta \mathbf{X}_m^{(LS)} = 2\mathbf{W}_m^*(\mathbf{P} - \hat{\mathbf{P}}_0\mathbf{Q})\mathbf{Q}^H\mathbf{W}_m^H \tag{5.2.6}$$

and

$$\Delta \mathbf{x}_m^{(reg)} = -\lambda\nabla_{\hat{\mathbf{x}}_m}||\hat{\mathbf{x}}_m||_S. \tag{5.2.7}$$

Here $\Delta \mathbf{X}_m$ is the total $m-$focal domain update, $\Delta \mathbf{X}_m^{(LS)}$ is the portion of $\Delta \mathbf{X}_m$ associated to the least-squares part of J and $\Delta \mathbf{X}_m^{(reg)}$ (the Fourier transform of the data-cube of $\Delta \hat{\mathbf{x}}_m^{(reg)}$) is the portion of $\Delta \mathbf{X}_m$ associated to the regularization part of J. The super-scripts $\cdot^*$ and $\cdot^H$ refer to a complex conjugation and a conjugate transpose operation respectively. Note that in equations 5.2.5, 5.2.6 and 5.2.7, $\Delta \mathbf{X}_m^{(LS)}$ and $\Delta \mathbf{X}_m^{(reg)}$ are related with $\Delta \mathbf{X}_m$ via an overall scaling factor λ which controls the strength of the sparsity regularization.

The updates $\Delta \mathbf{X}_m$ are used to renew the estimate of $\mathbf{X}_m$ in every iteration k via the recursion formula

$$\hat{\mathbf{X}}_m^{(k+1)} = \hat{\mathbf{X}}_m^{(k)} + \alpha_m \Delta \mathbf{X}_m^{(k)}, \tag{5.2.8}$$

where the scaling parameter α_m is chosen by a line search such that $J(\hat{\mathbf{X}}_m^{(k+1)}) = J(\hat{\mathbf{X}}_m^{(k)} + \alpha_m \Delta \mathbf{X}_m^{(k)})$ is minimized. The surface operator $\hat{\mathbf{A}}$ is strongly dependent on the estimate of $\hat{\mathbf{P}}_0$, so it can be calculated in every iteration directly from the $\hat{\mathbf{P}}_0$ estimate by

$$\hat{A} = \begin{cases} \hat{A} \quad s.t. \quad ||\mathbf{P} - \hat{\mathbf{P}}_0\hat{A}\mathbf{P}||^2 \rightarrow min & \text{in the early iterations} \\ \hat{A} \quad s.t. \quad ||\mathbf{P} - \hat{\mathbf{P}}_0 - \hat{\mathbf{P}}_0\hat{A}\mathbf{P}||^2 \rightarrow min & \text{in the rest of the iterations} \end{cases} \tag{5.2.9}$$

The matching of $\hat{A}$ is done in practice via a least-squares adaptive filter. The filter, which is restricted to be short in time, is found such that the amplitudes of $\mathbf{P}$ (or $\mathbf{P} - \mathbf{P}_0$) and $\hat{\mathbf{P}}_0\hat{A}\mathbf{P}$ are matched, driving their difference to a minimum [Verschuur and Berkhout, 1997]. Equation 5.2.9 provides an extra constraint (a-priori knowledge) and cannot be derived from equation 5.2.1. More about equation 5.2.9 can be found in section 5.4.

5.3 Data interpolation

The Focal CL-SRME algorithm is constructed to include data interpolation in its scheme. Two types of data interpolation are considered (see Chapter 1):

[1] *Under-sampling interpolation*: This type of interpolation is needed to remove the aliasing artefacts produced by the data gaps. These artefacts will mainly map to empty regions in the focal domain and they can be removed by imposing sparsity in the solutions. Data gaps and non-regular under-sampling can also be interpolated in a similar fashion. A reconstruction problem in the direct domain is mapped to a de-noising problem in the focal domain. The de-noising problem in the focal domain is solved by a sparsity constraint (see figure 5.1).

[2] *Near-offset interpolation*: This type of interpolation is needed to fill the near-offset gap and provide a good multiple separation. Its strength lies mainly in the back-projection of the multiples into the primaries (equation 5.2.6). The regularization and the compression in the focal domain are then useful to remove the cross-talk energy and to drive the solution closer to the true values. Due to the back-projection, we require larger offset coverage compared to the under-sampling case. Also, due to an initial amplitude difference in the gradient between the back-propagated traces and the measured traces, amplitude matching becomes necessary. The matching is done only in the first iteration to the term $[(\mathbf{P}-\hat{\mathbf{P}}_0\mathbf{Q})\mathbf{Q}^H]^{(1)} = \mathbf{P}(\mathbf{I}+\hat{\mathbf{A}}\mathbf{P})^H$ to scale up the amplitudes of the interpolated traces compared to non-interpolated traces. At this point, the traces to interpolate are already roughly filled by the back-projection mechanisim shown in equation 5.2.6 (see figure 5.3).

To interpolate the missing data, we replace the empty traces in our input dataset $\mathbf{P}$ with the interpolated traces in out modeled dataset. The replacement is done only for the traces to interpoalte, the measured traces are kept intact. The residual in the interpolated traces is set to zero by construction via the following condition:

$$\mathbf{P}'' = (\sum_{m=1}^{M} \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q})''; \qquad (5.3.10)$$

where $\mathbf{P}''$ is the data to reconstruct and the double prime in the right hand side means selecting only the data points corresponding to $\mathbf{P}''$ (missing data positions).

Equation 5.3.10 uses the prediction equation 3.2.1 to produce estimates of the missing data $\mathbf{P}''$ given the knowledge of the fully sampled primary wavefield $\hat{\mathbf{P}}_0$. Note that the relationship between $\mathbf{P}$ and $\mathbf{P}_0$ (equation 3.2.1), implies that the knowledge of $\mathbf{P}_0$ and $\mathbf{A}$ fully determines $\mathbf{P}$. This means that if we are able to reconstruct the missing data in $\hat{\mathbf{P}}_0$ (which we do via a sparsity constraint on the focal domain), then

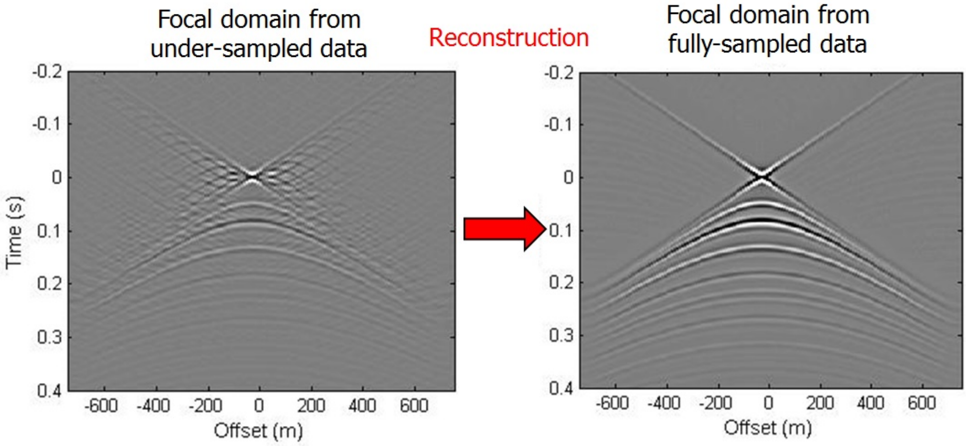

Figure 5.1: *Reconstruction in Focal CL-SRME is achieved by using the focal domain to transform an under-sampling problem into a de-noising problem. The de-noising problem is solved by imposing a sparsity-promoting regularization in the inversion. (Left) focal domain from under-sampled data, (right) focal domain from fully-sampled data. The dataset in the right has all the noise removed so its fully sampled.*

we automatically reconstruct $\mathbf{P}$ via equation 5.3.10. In fact, it can be shown that $\mathbf{P}$ can be written fully in terms of $\mathbf{P}_0$ and $\mathbf{A}$ via a scattering series as shown by equation 2.3.6. Note that equation 5.3.10 refers only to the missing data positions in the input dataset, such that the original data $\mathbf{P}'$ is kept intact.

Our initial guess for $\hat{\mathbf{P}}_0^{(0)}$, $\hat{\mathbf{A}}^{(0)}$, $\mathbf{P}''^{(0)}$, together with equations 5.2.5, 5.2.6, 5.2.7, 5.2.8, 5.2.9 and 5.3.10, constitute the essence of the proposed joint multiple prediction and interpolation method.

In order to obtain a first interferometric estimate of missing data (especially in the near-offsets), it is important to start the inversion with an initial $\hat{\mathbf{A}}$ value such that the first $\mathbf{P}_0$ estimate contains dominant back-projections of the multiples into the primaries. In the near-offset regions, the back-projection mechanism of equation 5.2.6 is the main mechanism driving the solutions. We will assume $\hat{\mathbf{P}}_0^{(0)} = 0$, $\hat{\mathbf{A}}^{(0)} = A_{initial}$ and $\mathbf{P}''^{(0)} = 0$ at the beginning of the algorithm. The $A_{initial}$ term is calculated using equation 5.2.9 and $\hat{\mathbf{P}}_0 = \mathbf{P}$.

Figure 5.2 depicts the current algorithm by showing the data to interpolate $\mathbf{P}$ together with the focal domain of its first primary $\hat{\mathbf{X}}_1$ (only one focal domain in this example). As we can see, as the algorithm progresses, the aliasing noise disappears, $\mathbf{P}$ gets reconstructed and the multiple energy is eliminated from $\hat{\mathbf{X}}_1$.

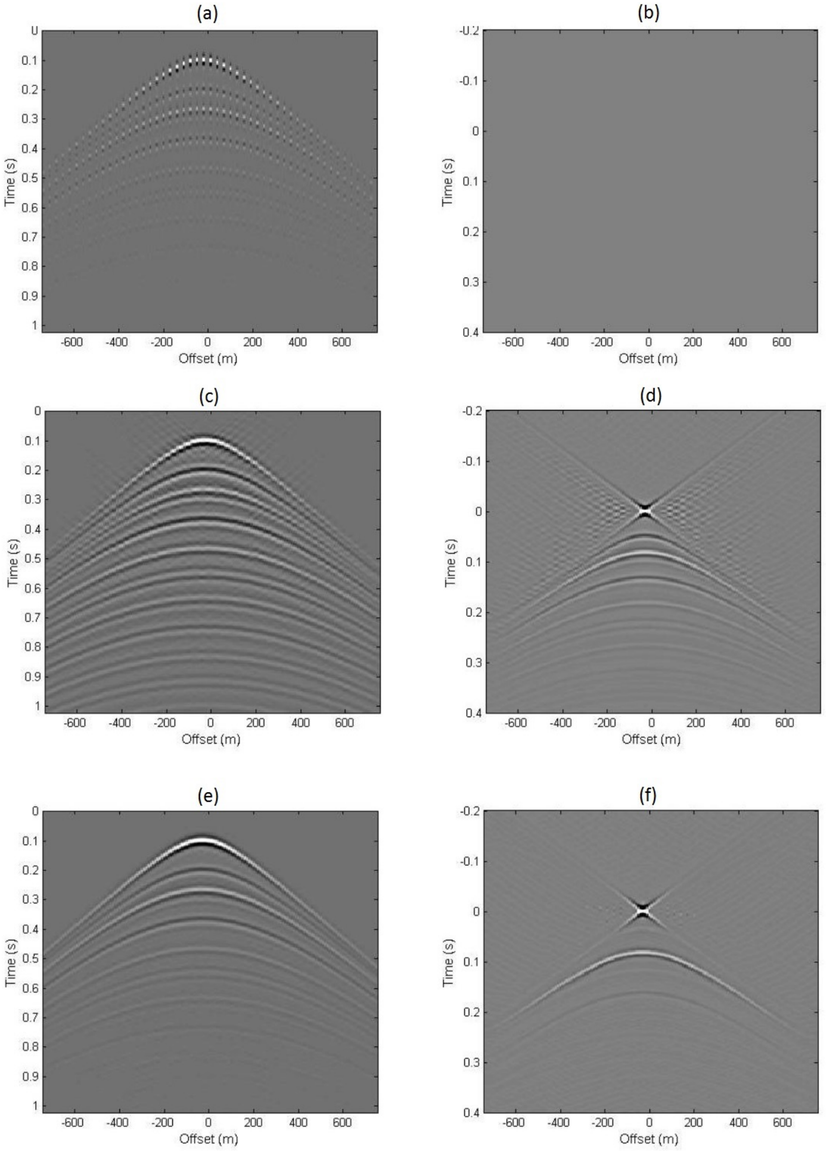

Figure 5.2: *Focal CL-SRME algorithm used for under-sampling reconstruction over a 1:3 under-sampling ratio. The model used to generate the data is the same one as in figure 4.1. The interpolated data* $\mathbf{P}$ *is shown next to the focal domain of its first primary* $\hat{\mathbf{X}}_1$ *(only one focal domain used here) for several iterations. At the beginning (a)* $\mathbf{P}^{(0)}$*, (b)* $\hat{\mathbf{X}}_1^{(0)}$*, after the first iteration (c)* $\mathbf{P}^{(1)}$*, (d)* $\hat{\mathbf{X}}_1^{(1)}$ *and after the 30th iteration (e)* $\mathbf{P}^{(30)}$*, (f)* $\hat{\mathbf{X}}_1^{(30)}$*. As the algorithm progresses, the aliasing noise disappears,* $\mathbf{P}$ *gets reconstructed and the multiple energy is eliminated from* $\hat{\mathbf{X}}_1$*.*

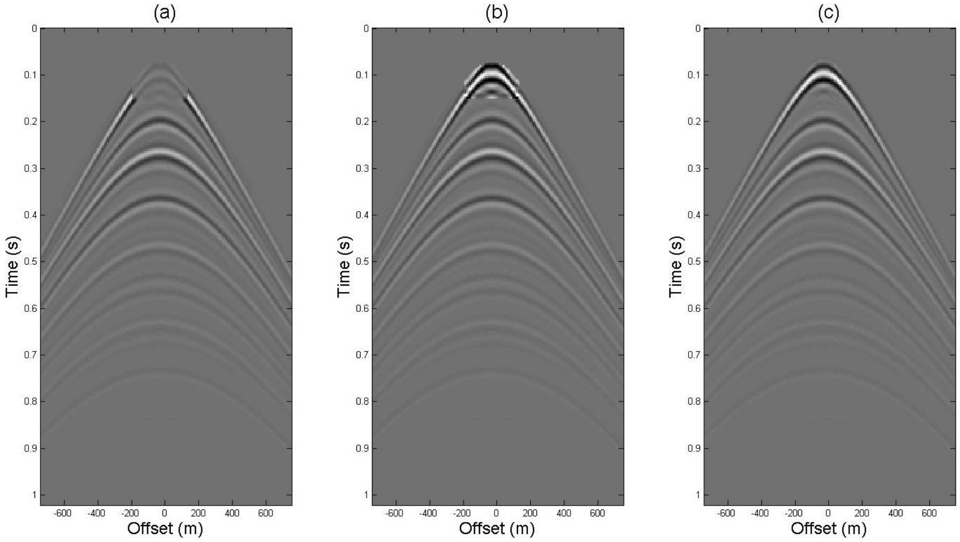

Figure 5.3: *The back-projected energy of the multiples can be used to interpolate missing near-offset data. However, in order to do so, it is necessary to balance the amplitude difference between the present and the missing data in the first iteration gradient. Here, the defocused 1st iteration gradient* $[(\mathbf{P}-\hat{\mathbf{P}}_0\mathbf{Q})\mathbf{Q}^H]^{(1)}$ *is shown (a) before and (b) after amplitude matching. The interpolation will be improved with iterations until arriving to (c) the final estimate of the data. The velocity model used here is shown figure 4.1. In this example, a 320m near-offset gap over a 0.15s time range was interpolated.*

5.4 Strategy for estimating $\hat{\mathbf{A}}$

Equation 5.2.9 uses the minimum primary energy constraint [Verschuur and Berkhout, 1997] in the early iterations to estimate $\mathbf{A}$ (assuming $\hat{\mathbf{A}} = \hat{A}\mathbf{I}$). This is necessary to avoid the trivial solution given by $\hat{A} = 0$ and $\mathbf{P}_0 = \mathbf{P}$. It is also necessary to maintain the minimum energy constraint until the sparsity starts playing a relevant role in the updates. Typically, the constraint can be dropped after $\sqrt{J}$ has decayed to 20% of its value at the initial guess.

Dropping the minimum energy criterium after some iterations is useful for the following reasons:

[1] Its theoretically more accuarate to minimize the residual $\mathbf{P}-\hat{\mathbf{P}}_0-\hat{\mathbf{P}}_0\hat{A}\mathbf{P}$ rather than the estimated primaries $\mathbf{P}-\hat{\mathbf{P}}_0\hat{A}\mathbf{P}$.

[2] We can better resolve the residual if we drop the minimum energy criterium.

In figure 5.4 we can see the comparison between the inversion residuals obtained by starting with the minimum energy constraint and then relaxing it, and the residual obtained by keeping the minimum energy constraint throughout the entire inversion. The normalized input data is also shown. A lower residual is obtained when we switch off the minimum energy constraint after a given number of iterations. The switch-off strategy seems to perform better for this and most of the observed examples.

5.5 Algorithm outline

We can outline the focal CL-SRME algorithm by resuming its main steps. The iteration index has been omitted for simplicity. The M variable represents the desired number of focal domains. The $\mathcal{F}$ operator represents a Fourier Transform mapping time to frequency coordinates. Here $\mathbf{P}$ is the total data, $\mathbf{P}'$ is the input data (with zeros in the missing-data positions) and $\mathbf{P}''$ is the interpolated data (with zeros in the measured data positions).

Focal Closed-Loop SRME algorithm

```
Calculate the $\mathbf{W}_m$ operators using NMO times and velocities.
$\hat{\mathbf{X}}_m \Leftarrow \mathbf{0}$, $\mathbf{P} \Leftarrow \mathbf{P}'$
Calculate $\hat{\mathbf{A}}$ via an optimum filter such that $||\mathbf{P} - \mathbf{P}\hat{A}\mathbf{P}||^2 \rightarrow min$
while ($iter \leq maxIterations$) do
    If $iter = 1$ then apply near-offset amplitude matching
    while ($m \leq M$) do
        $\Delta\mathbf{X}_m^{(LS)} \Leftarrow 2\mathbf{W}_m^*(\mathbf{P} - \hat{\mathbf{P}}_0 - \hat{\mathbf{P}}_0\hat{\mathbf{A}}\mathbf{P})(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})^H\mathbf{W}_m^H$
        $\Delta\mathbf{X}_m^{(reg)} \Leftarrow \mathcal{F}\{-\lambda\nabla_{\hat{\mathbf{x}}_m}||\hat{\mathbf{x}}_m||_S\}$ (Appendix C)
        $\Delta\mathbf{X}_m \Leftarrow \Delta\mathbf{X}_m^{(LS)} + \Delta\mathbf{X}_m^{(reg)}$
        $\hat{\mathbf{X}}_m \Leftarrow \hat{\mathbf{X}}_m + \alpha_m\Delta\mathbf{X}_m$ ($\alpha_m$ s.t. $J \longrightarrow min$, Appendix B.2)
    end while
    $\hat{\mathbf{P}}_0 \Leftarrow \sum_{m=1}^{M}\mathbf{W}_m^T\hat{\mathbf{X}}_m\mathbf{W}_m$
    Find $\hat{\mathbf{A}}$ according to equation 5.2.9
    $\mathbf{P}'' \Leftarrow (\hat{\mathbf{P}}_0 + \hat{\mathbf{P}}_0\hat{\mathbf{A}}\mathbf{P})''$
    $\mathbf{P} \Leftarrow \mathbf{P}' + \mathbf{P}''$
end while
```

5.6 Discussion

To allow efficient reconstruction, the chosen parameterization must represent the wavefields with the smallest amount of non-zero parameters possible. Reconstruc-

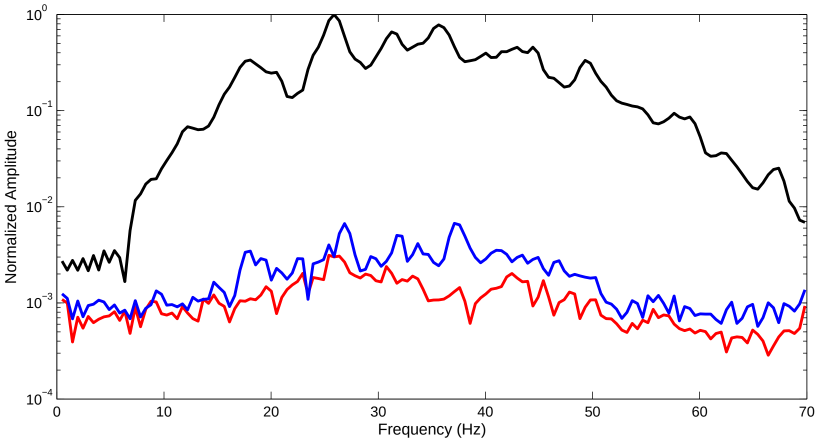

Figure 5.4: *Spectral amplitudes of the central traces in the input data (black), the residual obtained by switching off the minimum energy constraint after a given number of iterations (red), and the residual obtained by keeping the minimum energy constraint for all iterations (blue). To avoid interference effects, spectral amplitudes are calculated over a set of central traces rather than a single trace. We computed these results using the dataset presented in figure 1.5a.*

tion is typically achieved after a thresholding/regularization process in which the under-sampling artefacts are taken out of the estimates, leaving only the meaningful parameters to describe the data. As a result, the data then becomes fully sampled after the inverse transformation. Note that this type of reconstruction can be used in both deep and shallow data reconstruction. The focal transform is especially suited for shallow reflectors, as typically these event are strongly curved and can be efficiently represented in the focal domain. Other parameterizations like curvelets [Lin and Herrmann, 2009], Fourier [Zwartjes and Gisolf, 2006] or Linear Radon [Trad et al., 2003] transforms are also possible, but provide poorer data compression when applied to highly-curved events (e.g. around the near-offsets with shallow water data).

In terms of computation cost, the Focal CL-SRME algorithm can be described as relatively fast, with typically ~ 30 iterations needed for convergence. The speed is in principle directly affected by the number of focal domains used, but this is relaxed by the fact that typically we don't require many focal domains to achieve a good result (1 to 3 focal domains are enough in most cases). When a single focal domain is used, the current algorithm requires at least 6 data products per iteration, SRME on the other hand, requires only one data product per iteration. A larger number of data products is a consecuence of a more complex description, involving both correlations and convolutions in the updates, together with the focus/defocus steps. With the same settings, the algorithm seems to be around 10 times faster than EPSI. This is due to the fact that EPSI requires a top-down approach in which the strongest primaries are explained first. This process typically involves many more iterations (~ 100).

Though important in 2D, we expect the former method to have a stronger impact in 3D, as typically it is there where the data interpolation faces most of its challenges. However, practical problems must be faced when handling 3D datasets, the most important of them being the large data volumes present, and the fact that in Focal CL-SRME we require to store in memory several datasets with dense sampling (i.e. missing traces replaced by traces filled with zeros). A practical strategy to circumvent this issue would be to use an hybrid approach, such that the near-offsets in the primaries are estimated by the accurate (but more expensive) Focal CL-SRME, and the far offsets are estimated by a less accurate (but lighter) differential NMO interpolation method. This idea will be developed in Chapters 7 and 8.

Note that the Focal CL-SRME algorithm is built to work in under-sampling scenarios in which algorithms like EPSI would not be applicable (due to the large amounts of missing data). Rather than competing with EPSI, we propose the Focal CL-SRME algorithm as a method complementary to EPSI, to be used whenever EPSI fails to apply properly.

Focal CL-SRME reduces to the previosly introduced CL-SRME if we choose only one focal domain with $\mathbf{W} = \mathbf{W}^H = \mathbf{I}$.

5.7 Conclusions

In this chapter we have presented the theoretical basis of the Focal CL-SRME method. This method incorporates data interpolation in the multiple prediction process. Its most important achievement being the capacity of back-projecting the multiples into the primaries to reconstruct missing information in the input dataset. Examples of this algorithm will be presented in the next chapter.

6

Results of the 2D algorithm

We have designed two experiments to test both the under-sampling and the near-offset interpolation capabilities of the present algorithm in the 2D case. Furthermore, we show results for 2D field data.

6.1 Near-offset interpolation

■ 6.1.1 30m sea-floor salt model example

In order to account for the near-offset reconstruction capabilities of the present algorithm, a synthetic 2D model with a shallow sea-floor at 30m (figure 6.1) is used for the generation of the synthetic data via an acoustic finite difference algorithm. We remove entirely the information of the shallow reflector up to 0.18s, simulating the situation that, in practice, for such a shallow water hardly any information from the water bottom reflection is measured. We will concentrate our interpolation efforts in the shallower zone of the data (early times), as the deeper parts of the data can be accurately interpolated with other methods. Note that with a 30m sea-floor depth, NMO-based interpolation methods would suffer from strong wavelet stretching, rendering the interpolation very inaccurate.

The results of the Focal CL-SRME algorithm applied to this dataset are displayed in figure 6.3. As we can see, the Focal CL-SRME algorithm reconstructed the shallow reflector information, providing also accurate multiple separation in the deeper parts of the data. Only one focal domain was used here for inversion (centered in the first reflector and based on NMO parameters). The ability of mapping the multiples into the primaries inherent in CL-SRME, together with the sparsity promoting regularization in a compressed domain, are the keys for large-gap near-offset reconstruction in this example.

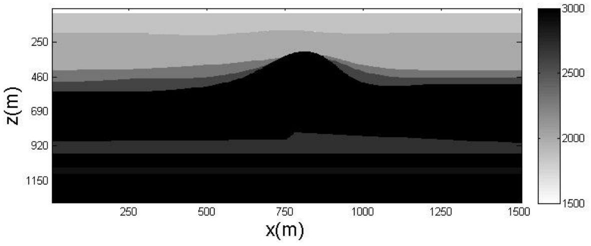

Figure 6.1: *Velocity model used for the synthetic data generation in the 30m sea-floor salt-model example.*

■ 6.1.2 50m sea-floor salt model example

Next we will address the synthetic model shown in figure 6.2, with a shallow reflector at 50m. We remove entirely the information of the shallow reflector, eliminating the information in all offsets. In practice this is equivalent to not measuring the first reflector primary at all. Again, traditional methods based on NMO or Radon reconstruction are impossible to apply in these circumstances. The Focal CL-SRME algorithm is used for primary estimation and near-offset reconstruction of this dataset. The results are shown in figure 6.4.

Figure 6.4a depicts the input dataset and figures 6.4b and 6.4c depict the reconstructed dataset and the estimated primaries. As we can see, the shallow event has been reconstructed almost completely, and its associated multiples are largely removed from the estimations. Some multiple leakage is still observed in the estimates, but these effects are small if we consider the amount of interpolation that had to be done. Note, in particular, that the multiples in the deeper regions of the data are well removed.

■ 6.1.3 100m sea-floor field data example

In this section we will show an example of the Focal CL-SRME algorithm applied to a North-Sea dataset (courtesy of PGS). The data is taken from a 2D line from which a subset is selected with 201 receivers and 201 sources, with a 12.5m receiver-spacing. By applying reciprocity a split-spread dataset is obtained. Note that the same dataset has also been used by Baardman et al. [2010]. In this example we use Focal CL-SRME to generate a near-offset interpolation over the shallow part of the

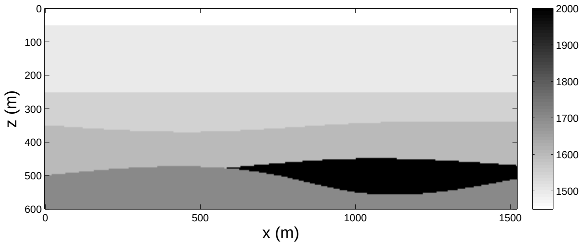

Figure 6.2: *Velocity model used for the synthetic data generation in the 50m sea-floor salt-model example.*

data. A box of 375m in the near-offset and 0.24s in time will be interpolated from the data. Other methods can be used to interpolate the data at the deeper times. A dense grid of 300 sources and receivers is used.

Figure 6.5a shows the input dataset and figures 6.5b and 6.5c display the reconstructed dataset and the estimated primaries. As we can see, the shallow event has been reconstructed almost completely, and its associated multiples are largely removed from the estimations. Again, the multiples in the deeper regions of the data are largely removed. Only one focal domain was used here for inversion (centered around the first reflector and based uniquely on NMO parameters). A zoom in the shallow section can be seen in figure 6.6.

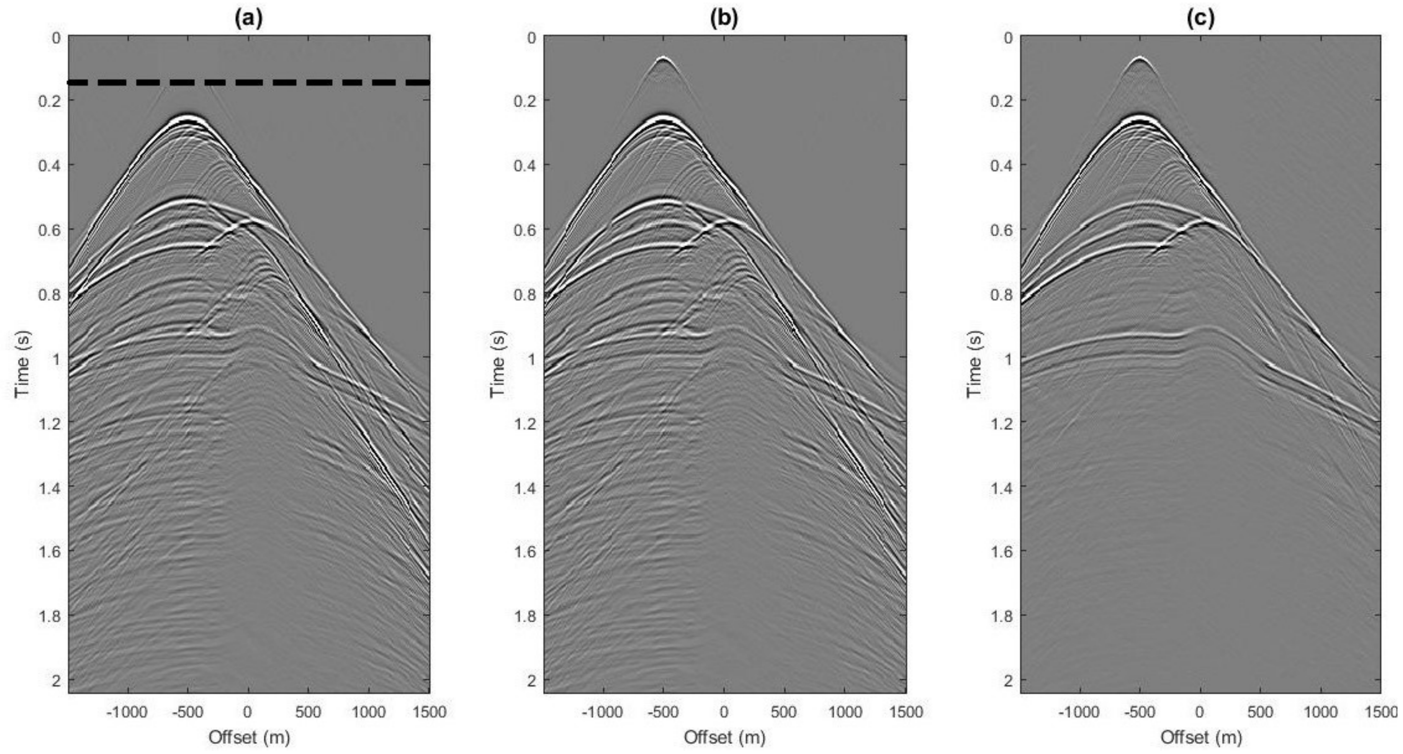

Figure 6.3: *Focal CL-SRME in a shallow water environment (30m water-depth) using a salt-model and near-offset reconstruction. (a) Input dataset with the first primary muted up to 0.18s (muting done above the dashed-line), (b) reconstructed dataset and (c) estimated primaries. Full near-offset reconstruction is achieved in the shallow region, with the multiples in the deep regions largely removed. 30 iterations are needed for this example.*

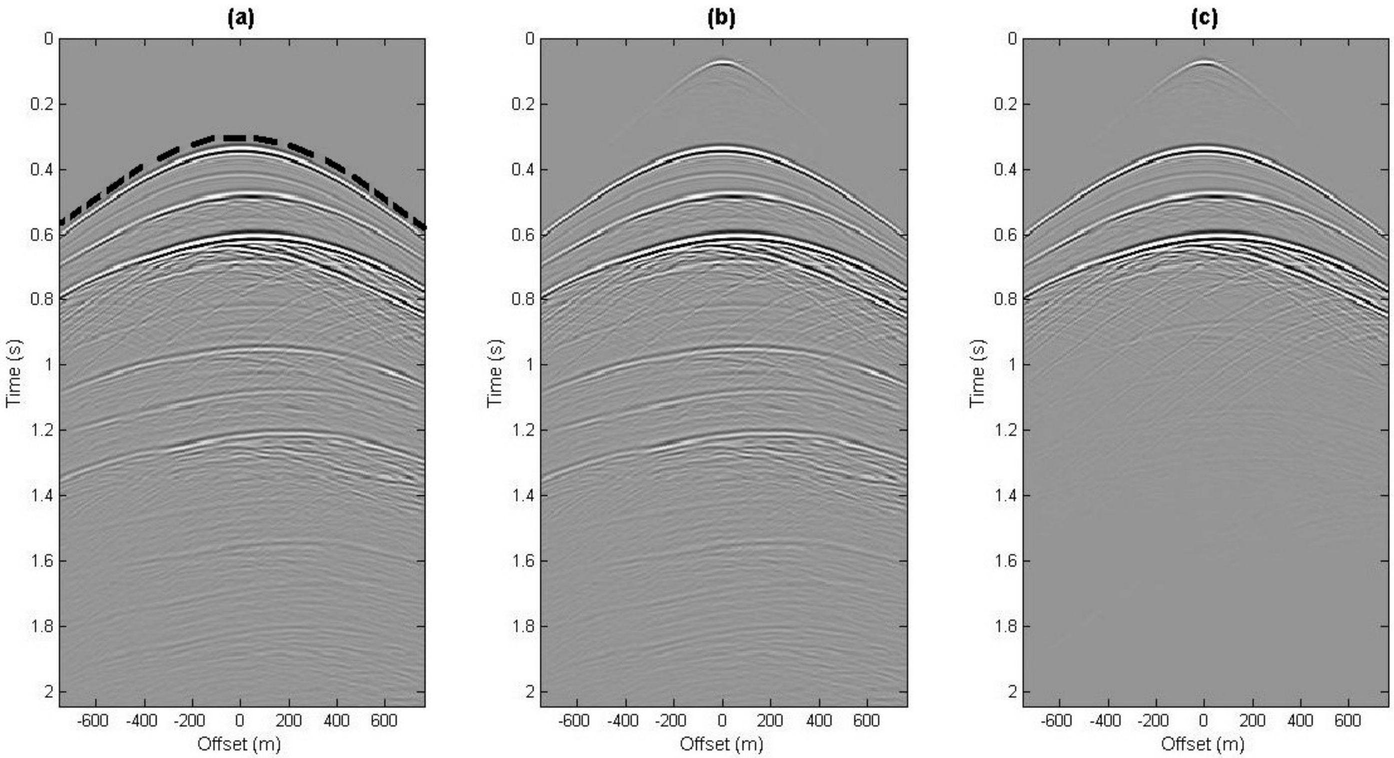

Figure 6.4: *Focal CL-SRME in a shallow water environment (50m water-depth) with complete near-offset reconstruction (the entire offset range is reconstructed). (a) Input dataset with the first primary completely muted (muting done above the dashed-line), (b) reconstructed dataset and (c) estimated primaries. Full near-offset reconstruction is achieved in the shallow region, with the multiples in the deep regions largely removed. 30 iterations are needed for this example.*

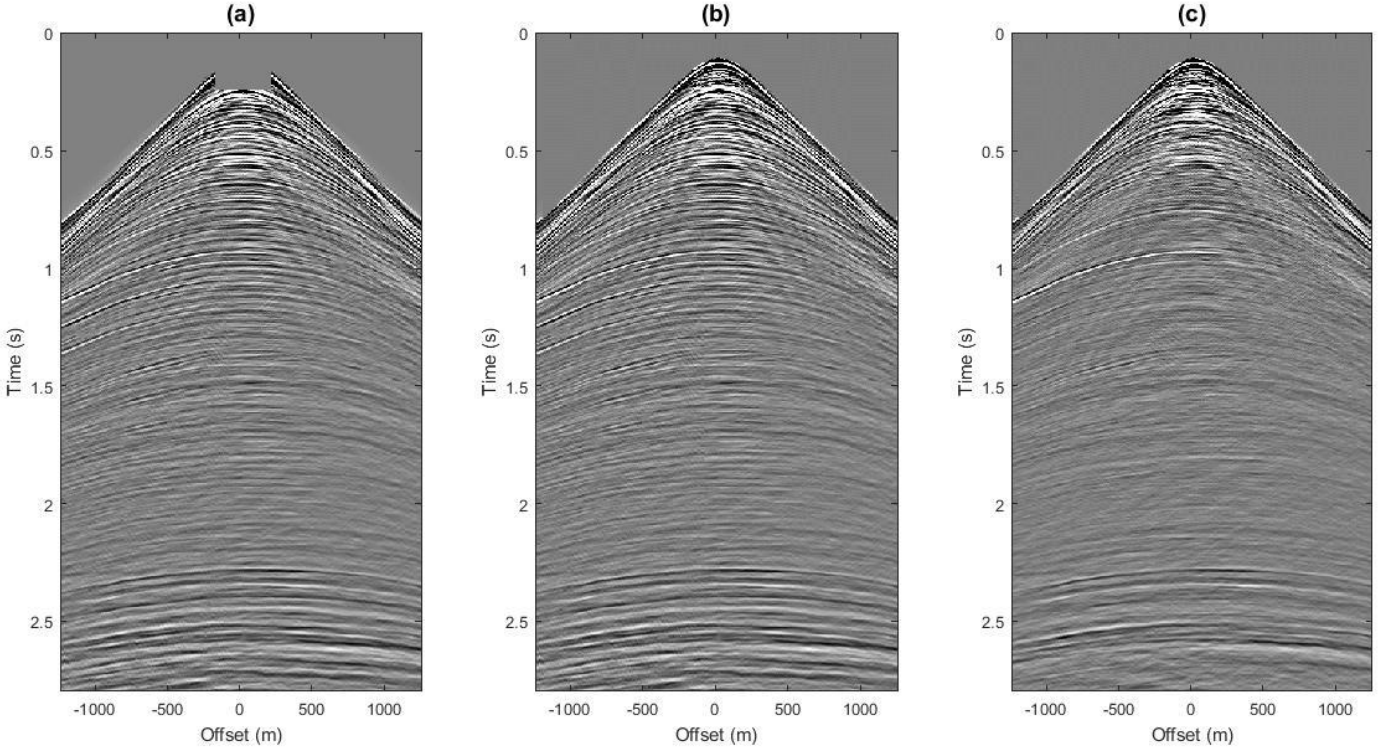

Figure 6.5: *Focal CL-SRME for field data (around 100m water-depth) with near-offset reconstruction. (a) Input dataset with the shallow near-offsets muted out, (b) reconstructed dataset and (c) estimated primaries. 30 iterations are needed for this example.*

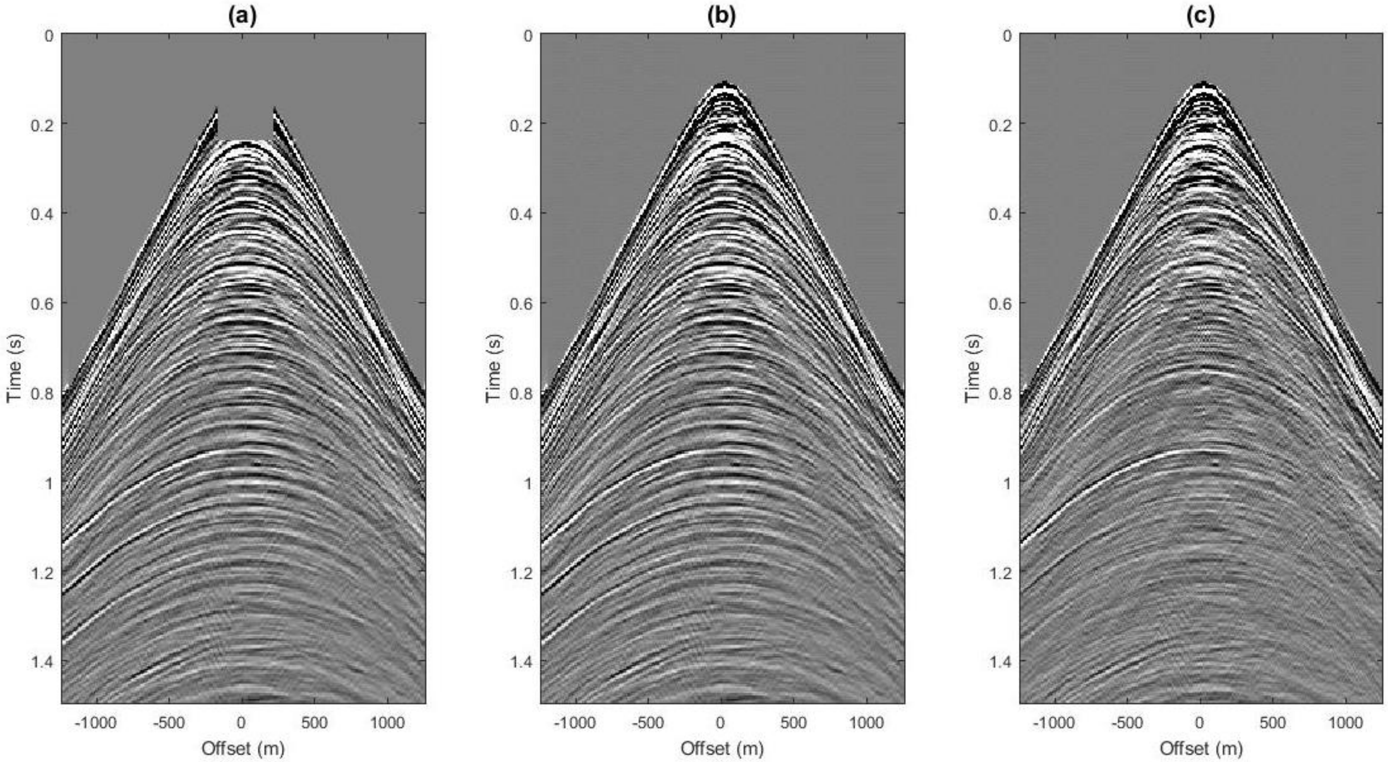

Figure 6.6: *Shallow section zoom (up to 1.5s) for the gathers shown in figure 6.5. (a) Input dataset with the shallow near-offsets muted out, (b) reconstructed dataset and (c) estimated primaries.*

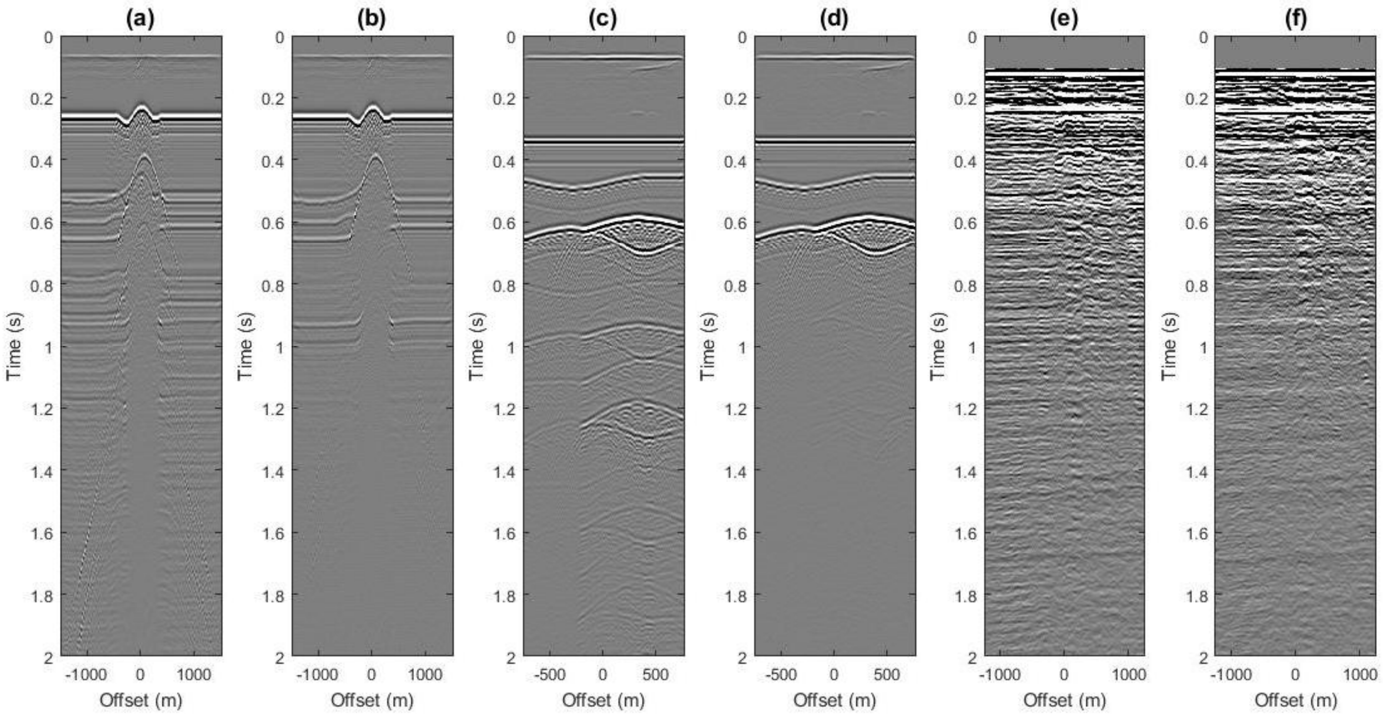

Figure 6.7: *Zero-offset gathers for: the 30m sea-floor example (a) data, (b) estimated primaries; the 50m sea-floor example (c) data, (d) estimated primaries; the 100m sea-floor example (e) data and (f) estimated primaries. In all the examples the early times are interpolated with Focal CL-SRME.*

■ 6.1.4 Variation of the sea-floor depth

In order to better account for the differences in primary estimation of SRME (with initial interpolation) and Focal CL-SRME, a new experiment is done for many models with several sea-floor depths and near-offset gaps. This, time, however we will keep part of the input offset range and we will interpolate only up to 0.32s, such that we focus our interpolation efforts in the shallow part of the data. We will generate the synthetic data using the model shown in figure 6.2. In each experiment the difference between the true primaries and the estimated primaries is computed and the relative energy error E is calculated via

$$E = \frac{\sum(P_0^{(true)} - P_0^{(estimated)})^2}{\sum(P_0^{(true)})^2} \times 100\%. \qquad (6.1.1)$$

The results are shown in figure 6.8. As expected, errors in both algorithms tend to become larger when the water bottom becomes shallower. The Focal CL-SRME tends to produce always better primary estimations, with less residuals for all water depth values. The difference between the primary estimates of the CL-SRME algorithm and the traditional SRME is larger for shallower models (where more interpolation power is needed) and becomes less important for deeper models (where the initial interpolation is accurate enough). Overall, the Focal CL-SRME algorithm appears to better handle shallow data. Small differences in parameter settings (of both algorithms) generate some of the variations observed in the shallower depths in figure 6.8.

6.2 Under-sampling interpolation

Here we will use again the North-sea dataset from section 6.1.3. The data is taken from a 2D line from which a subset is selected with 201 receivers and 201 sources, with a 12.5m receiver-spacing. By applying reciprocity a split-spread dataset is obtained. We have chosen sources to have a 1:5 irregular under-sampling ratio in the sources. Then, a coarse NMO-based shot interpolation is applied to the data as a pre-processing step to reduce the aliasing artifacts in the focal domain. The initial interpolation does not have to be accurate, as the interpolation errors will be mapped to aliasing artifacts in the focal domain, which will be removed later on by the sparsity constraint. In this example we will use three focal operators which are chosen to focus the energy of the three most prominent reflectors in the data. The operators are calculated using the stacking velocities of the reflectors of interest in the data, and assuming an homogeneous propagation medium for each operator. The NMO times and velocities used here are t_{NMO}=0.16, 0.52, 0.96 seconds and v_{NMO}=1500, 1600, 1800m/s. Besides the stacking velocities and the apex times of the events of interest, no other model information is used to generate the operators.

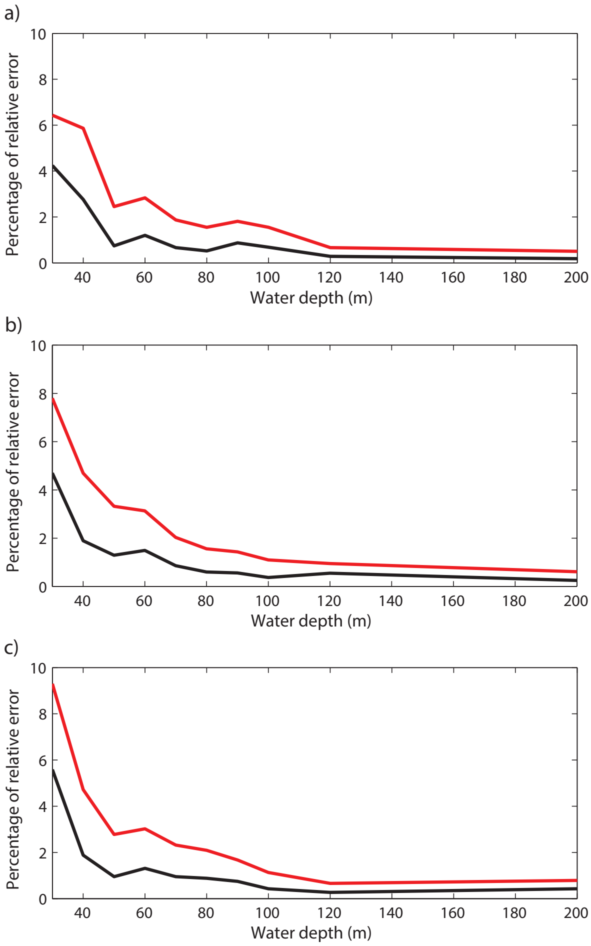

Figure 6.8: *Relative energy error of SRME with interpolation (red) and Focal CL-SRME (black) for different sea-floor depths. With (a) a 64m near-offset gap, (b) a 80m near-offset gap and (c) a 96m near-offset gap.*

Figure 6.9 shows the results of applying the Focal CL-SRME algorithm to this dataset. Because the interpolation is done along the source coordinate, results are presented in the common receiver gather domain. The input data is shown in figure 6.9a. The shot-interpolated data, primaries and multiples (at the first iteration) are shown in figures 6.9b, 6.9c and 6.9d. The shot-interpolated data, primaries and multiples (at the 30th iteration) are shown in figures 6.9e, 6.9f and 6.9g.

As we can see from figure 6.9, full data reconstruction has been achieved. By comparing the input (figure 6.9a) and the output (figure 6.9e) datasets we can see the effect of the interpolation in filling up the missing data. By comparing the first and last iterations in the data, primaries and multiples, we can see how the amplitude errors get healed, together with some phase errors. By comparing figures 6.9c and 6.9f we can also see the multiple energy eliminated from the primary recordings. Note that we have achieved a reasonably accurate primary-multiple separation, even in the case of a very coarsely-sampled input dataset. By comparing the input dataset (figure 6.9a) with the output primaries (figure 6.9f) we can note the appearance of two weak primaries in the deep part of the recordings (between 1.4 and 1.8 seconds), which are hardly visible in the input data. However, due to the huge interpolation effort in this example, some minor aliasing/phase artifacts are still visible in the output data, especially in the top part.

Figure 6.10 shows the associated primary focal domains, used for obtaining the results in figure 6.9. In this example three operators are chosen to focus the main reflectors, thus compressing the largest amount of energy possible. The associated energy of these events maps at zero time, while the other events map to either negative or positive times. The following remarks can be deduced from figure 6.10: (1) The initial redundancy in the focal domains has largely reduced. This is important for the uniqueness of the parameterization, as initially one single focal domain contains the information of the entire dataset. The former can be seen by events disappearing from one focal domain, as they are more effectively represented in another focal domain. (2) The aliasing artifacts produced by the under-sampling have been largely removed. This will produce data reconstruction in the original domain. The interpolation errors (due to un-physical data) in the initial primaries (figure 6.9c), might be difficult to separate from the data in the original domain. However, they become easy to separate in the focal domain (e.g. figure 6.10b) as they appear as aliasing artifacts.

The k_x-ω domain plots of the input and reconstructed data are presented in figure 6.11. As we can see the undersampling in the x-t domain is mapped to blending noise in the k_x-ω domain. After inversion the full reconstruction of the k_x-ω response is obtained with no visible presence of any remaining aliasing artifacts.

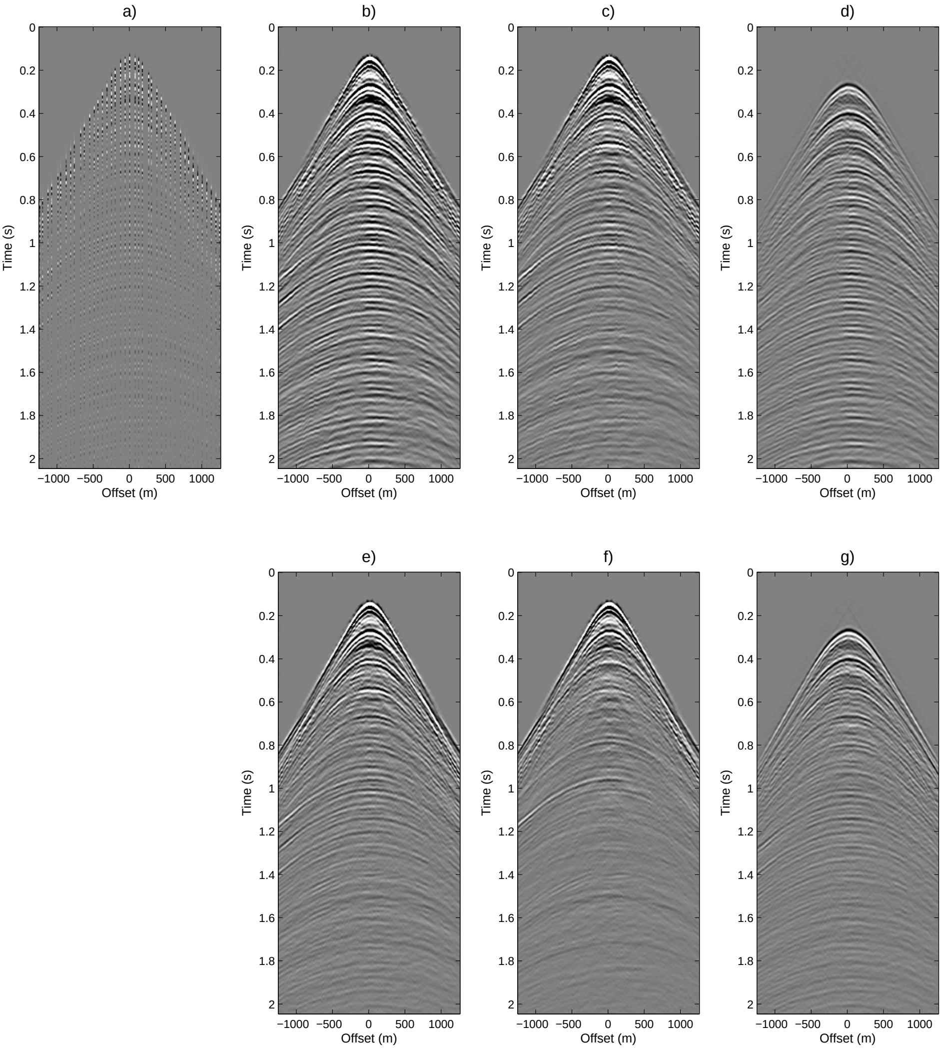

Figure 6.9: *Focal CL-SRME interpolation in a North-Sea dataset with a 1:5 shot undersampling ratio. A common-receiver gather is shown. Here we show the (a) original undersampled data, together with the first iteration (b) reconstructed data, (c) primaries and (d) multiples, and the 30th iteration (e) reconstructed data, (f) primaries and (g) multiples.*

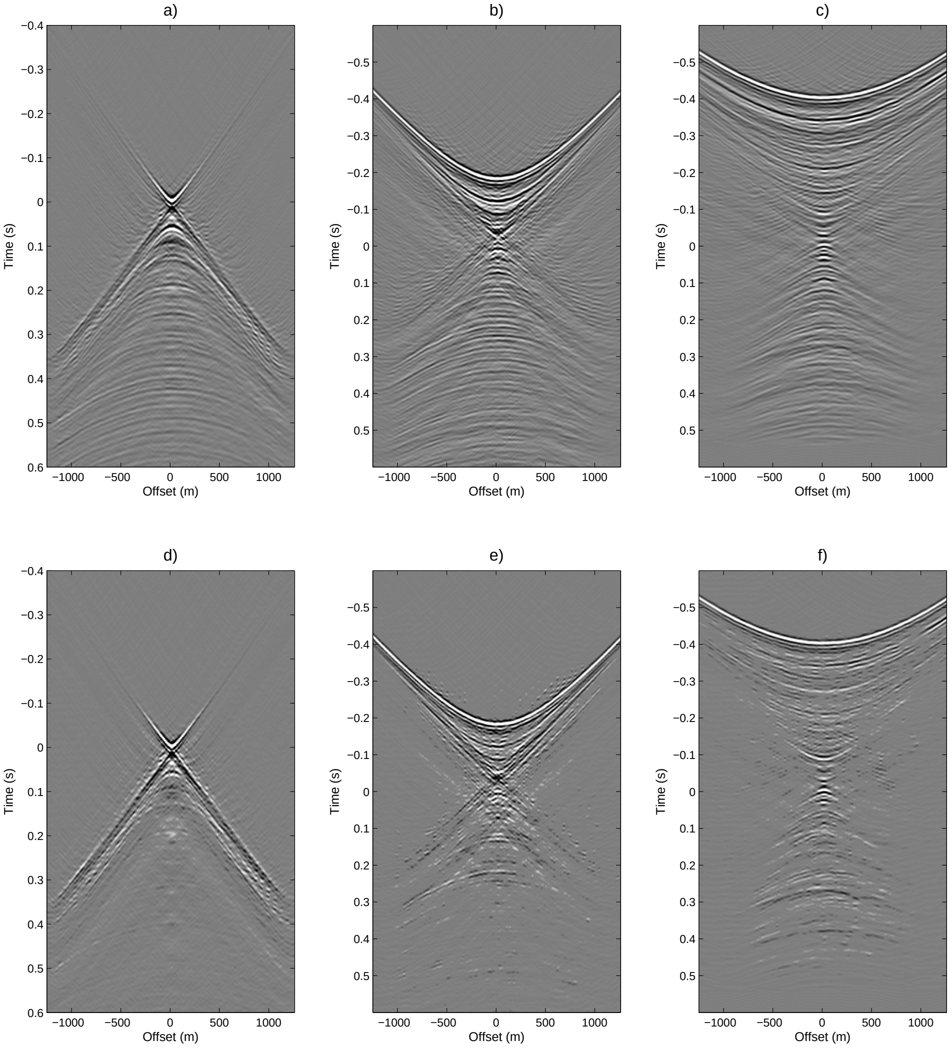

Figure 6.10: *Focal domains of the primaries associated with the results in figure 6.9. Figures (a),(b) and (c) represent the first iteration primary focal domains. Figures (d), (e) and (f) represent the 30th iteration primary focal domains. A common-receiver gather is shown. Some important differences can be appreciated between the upper panel (a-c) and the lower panel (d-f): (1) the elimination of the aliasing artefacts (e.g. diffraction-like events around* $t = 0$*), (2) the elimination of the multiples (events completely missing in the lower panel) and (3) the sparse representation of the primaries (events present in both panels but sparsely represented in the lower one).*

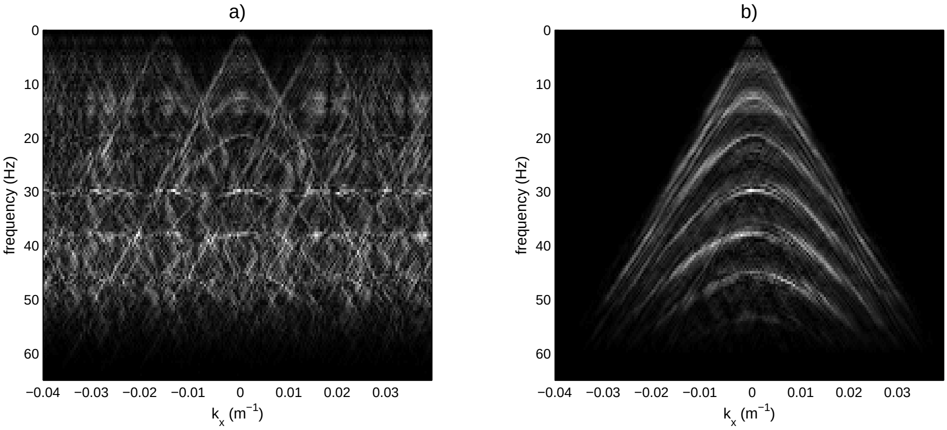

Figure 6.11: *k_x-ω domains of (a) the input under-sampled data and (b) the output reconstructed data. Note that a 1:5 undersampling ratio in the sources was applied to the input. Both figures have the same color scale.*

7

3D Focal Closed-Loop SRME

In this chapter we will present the 3D Focal CL-SRME algorithm. The fundamental theory will still be the same as the 2D algorithm developed in Chapter 7, but now practical considerations (regarding the data volume) will produce a different processing strategy. We will start with an introduction and a literature review on the field of 3D multiple separation. Then we will introduce the GSMP ('Generalized Surface Multiple Prediction') method. This method will allow the efficient calculation of large matrix products. Two types of data products will be discussed, that play an important role: the 'correlation product' and the 'convolution product'. Once all tools are in place, we can describe our scheme for the 3D Focal CL-SRME method. Without any loss of generality, we will make constant emphasis on the shallow reflector geologies, as this ones are the most relevant for the current work.

7.1 Introduction

In previous chapters we have introduced the 2D Focal CL-SRME algorithm. The main limitation of the former approach is the fact that it requires dense source and receiver coverage, with missing traces replaced by zero traces. Due to the huge data volumes in typical 3D acquisitions, this requirement becomes cumbersome to fulfil in practice. In order to overcome this issue we introduce here the 3D Focal CL-SRME approach. In this new approach we combine the ideas used for 3D SRME with the Focal CL-SRME method. We use the 3D SRME methodology to calculate the necessary data products within the Focal CL-SRME scheme, technology that is used to interpolate the sensitive near-offset information. Thus, near-offset multiple prediction is done mainly by Focal CL-SRME interpolation while far-offset prediction is based mainly on on-the-fly differential NMO (DNMO) interpolation. The result is an algorithm with the flexibility of 3D SRME and the built-in interpolation capabilities of Focal CL-SRME.

7.2 Literature review of 3D multiple elimination

With the introduction of surface related multiple attenuation (SRME) [Berkhout, 1982; Verschuur et al., 1992; Fokkema and van den Berg, 1993; Berkhout and Verschuur, 1997; Weglein et al., 1997] it became possible to estimate surface multiples on a fully data-driven manner, for both 2D and 3D datasets. The 3D version of the SRME algorithm would then follow the exact same equations as the 2D version, but with extended data matrices [Kinneging et al., 1989]. The ideal 3D SRME algorithm, however, posed a problem, as it required a dense coverage of sources and receivers along the sampled areas. This was impossible due to economical constraints, making data interpolation (prior to the application of SRME) a necessity.

The problem was resolved with the introduction of the 3D general surface multiple prediction (GSMP) method [Bisley et al., 2005; Kurin et al., 2006; Dragoset et al., 2008]. This method avoids any initial interpolation of the data before the prediction process (thus keeping the data volume within acceptable bounds), choosing instead an on-the-fly interpolation approach in which the multiples are predicted using contributions coming from differential NMO-interpolated traces [van Dedem and Verschuur, 1998]. The interpolation is done in a trace-by-trace fashion, such that only the strongest contributing traces are convolved to produce the multiple prediction result. Because there are no requirements on the ordering of the data structure, this type of approach also extended the SRME algorithm to different types of acquisition geometries.

Though very successful in a variety of scenarios, the 3D SRME method with interpolation [van Dedem and Verschuur, 1998] also presents some practical short-comings. Due to its reliance on DNMO interpolation, the method is sensitive to missing near-offset data, especially in the case of shallow reflections. In that case, the event's DNMO interpolation becomes inaccurate, and after convolution, these inaccuracies are mapped to wrong multiple (and primary) estimations. Additional errors in the method are also brought by errors in the interpolation itself, which assumes a laterally invariant medium, and presents known deficiencies when interpolating reflection data from dipping reflectors or complex structures.

To overcome the difficulties in near-offset interpolation when combined with multiple estimation, the Focal CL-SRME algorithm was proposed (chapter 5). In this scheme Closed-Loop SRME (chapter 3) is used to generate the primary predictions, such that, in combination with the previously introduced focal domain parameterization (chapter 4) we produce wave-equation-consistent data interpolation. The resulting algorithm has the capacity of performing simultaneous multiple prediction and data interpolation. The main advantage of this algorithm is the fact that interpolation in the near-offsets is done by back-projection of the multiples into the primaries, being therefore more amplitude-consistent than other less physical methods. By being wave-equation consistent, the algorithm also has the capacity of interpolating missing data over larger data gaps with better precision than the traditional interpolation. The disadvantage, however, is the fact that it required an input data matrix with fully-sampled source and receiver positions (with missing traces replaced by zero traces), thus making it cumbersome to apply to a real 3D dataset.

Here we introduce an extension of the Focal CL-SRME algorithm to make it feasible in practical 3D scenarios. We use the GSMP technology to generalize the possible acquisition geometries, such that we can generate all the necessary convolution/correlation products even when the sampling is coarse and/or irregular. In order to provide better interpolation results in the places where it is really needed, we will use the interpolation of the Focal CL-SRME algorithm for the near-offsets (accurate but computationally expensive), and a DNMO-based on-the-fly interpolation for the far-offsets (less accurate but computationally inexpensive). This extension will allow the estimation of primaries under generalized 3D acquisition geometries in a computationally feasible way.

7.3 3D multiple estimation

The earth is a 3D environment, therefore the structural changes in the earth will also have a 3D geometry. This implies that the primaries and the multiples will in general have 3D propagation paths. In many cases, structural complexities in the subsurface lead to multiples with propagation paths with bouncing points out of the source-receiver plane (see figure 7.1). The problem with this type of multiples is that, in order to be predicted, they require a source and receiver positions around the bouncing point. These source and receiver combinations would be missing if the processing is done in a 2D fashion (i.e. using only data along the source-receiver line) or if there was no data acquired around the bouncing point. Therefore, the need of a 3D multiple prediction scheme (to account for the geometry) and a fine areal sampling (to measure all possible bouncing points).

In order to describe multiples in a 3D geometry let's start with the description in the 2D discrete situation. Then the multiple wavefield M due to a source at x_s and a receiver at x_r can be written as

$$M(x_r, x_s, \omega) = \sum_{x_k} R_0 X_0(x_r, x_k, \omega) P(x_r, x_s, \omega) \tag{7.3.1}$$

where $R_0 \approx -1$ is the reflectivity of the surface, X_0 is the primary impulse response and P is the up-coming wavefield. We can extend equation 7.3.1 to the 3D case by adding extra lateral coordinates y_s and y_r such that

$$M(x_r, y_r, x_s, y_s, \omega) = \sum_{x_k} \sum_{y_k} R_0 X_0(x_r, x_k, \omega) P(x_r, x_s, \omega), \qquad (7.3.2)$$

which describes a convolution in the spatial coordinates. This process requires fine sampling in both sources and receivers along the summation surface.

Because in practice a fine sampling in both sources and receivers is never achieved in 3D acquisition, alternative solutions need to be found to generate the necessary traces for multiple prediction. In the following section we will discuss one option, the GSMP method, which we will later incorporate in the 3D Focal CL-SRME scheme.

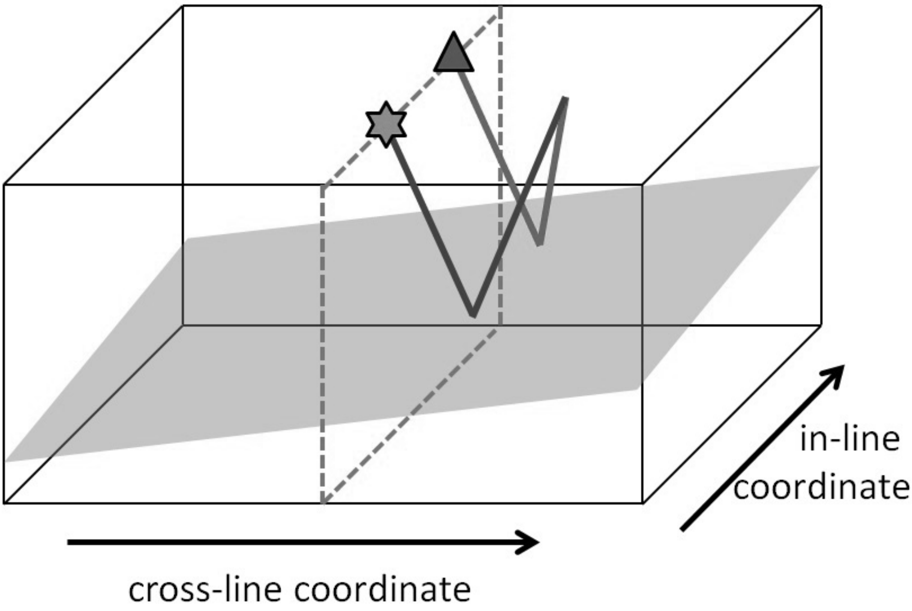

Figure 7.1: *A first order surface-related multiple for a model with a dipping reflector in the cross-line. The bouncing point in the surface is located outside source-receiver line. This means that this multiple will not be accurately predicted using a 2D prediction method in the in-line direction, here a 3D multiple prediction is needed (after Verschuur [2006]).*

7.4 The GSMP method

In order to properly discuss the 3D Focal CL-SRME algorithm, the general surface multiple prediction (GSMP) method [van Dedem and Verschuur, 1998; Bisley et al., 2005; Kurin et al., 2006; Dragoset et al., 2008, 2010] must be outlined first.

The GSMP method extends the SRME algorithm to generalized acquisition geometries via on-the-fly DNMO-interpolation for the missing traces in the dataset. Instead of trying to change the data to fit the algorithm, GSMP changes the algorithm to fit the data. In general, the GSMP algorithm presents a way of calculating the convolution of two large data matrices in a computationally efficient way, given that the input matrices are sparse wavefields and using the fact that their convolution represents the sampled version of a Kirchhoff summation.

Given an input dataset, the GSMP algorithm can be described as follows [Dragoset et al., 2008]:

[1] Input the recorded traces with their corresponding NMO velocities. Compute midpoint, offset and azimuth of each trace.

[2] Select a target trace and define the aperture of the computational grid for that trace.

[3] For every grid node in the aperture, use a nearest-neighbor search to select from the input traces the two best matching traces for that convolution.

[4] Compensate the two selected traces for offset errors using DNMO.

[5] Convolve the two traces and add to the current multiple prediction result.

[6] Go to step 3 until all grid nodes in the aperture are processed.

[7] Go to step 2 until the desired product is calculated for all input traces.

In step 3, the nearest-neighbor search is done by finding the input trace among all input traces that has minimum error E. Several choices are available for E. A typical implementation would include a Euclidean norm of the form:

$$\begin{aligned} E^2 &= [W_{hx}(hx_D - hx_I)]^2 + [W_{hy}(hy_D - hy_I)]^2 \\ &\quad +[W_x(x_D - x_I)]^2 + [W_y(y_D - y_I)]^2, \end{aligned} \tag{7.4.3}$$

where (hx,hy) and (x,y) represent the two offset and midpoint coordinates respectively. The subscripts D and I refer to the desired and input traces. The W factors are weights that mediate the relative importance of the different terms in the overall E value. The stronger the resemblance between the input trace and the desired trace, the smaller the E value.

A comparison between ideal SRME and GSMP methodologies is given in figure 7.2.

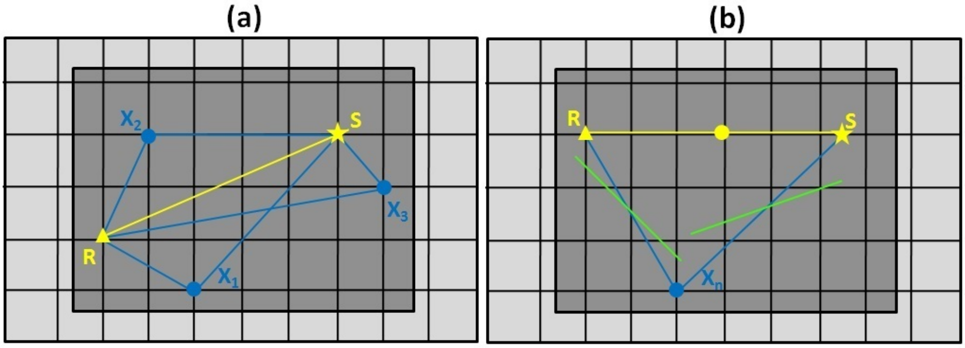

Figure 7.2: *Plan views of ideal 3D SRME and GSMP. a) In ideal 3D SRME, data are assumed to be completely regular so that the end points of* $\mathbf{R}$ *and* $\mathbf{S}$ *of each target trace (yellow line) fall on the nodes of a fixed grid oriented in a fixed direction. Each grid node in the aperture is considered a possible downward reflection point (DRP); There are three DRP's shown here, represented by* $\mathbf{X}$*'s. The contribution of the* n^{th} *DRP to an MCG is computed by convolving the corresponding traces* $\mathbf{RX}_n$ *and* $\mathbf{SX}_n$ *(blue lines), both of which must exist in the input dataset. b) In GSMP, the aperture grid is redefined for each target trace, based on its midpoint (yellow circle) and azimuth. The input to GSMP is a non-regularized, non-interpolated field dataset. Because, in general, the traces* $\mathbf{RX}_n$ *and* $\mathbf{SX}_n$ *required for the* n^{th} *convolution will not exist in the input dataset, a nearest-neighbor search selects the most suitable alternatives (green lines) from among the traces that are present (after Dragoset et al. [2010]).*

7.5 Generalization of the GSMP product for correlation and convolution

Due to the presence of multidimensional convolutions and correlations in the Focal CL-SRME theory (chapter 5), the construction of such products is necessary in the GSMP methodology. This implies the generalization of the concept to accommodate a correlation-type of product.

In 3D SRME this is not necessary as the SRME estimate involves convolution-type products only. This can be seen directly from the SRME prediction equation

$$\mathbf{P}_0 = \mathbf{P} - \mathbf{P}_0\mathbf{A}\mathbf{P} = \mathbf{P} - \mathbf{M}, \tag{7.5.4}$$

where $\mathbf{P}_0$ are the primaries, $\mathbf{P}$ is the input data, $\mathbf{M}$ are the multiples and $\mathbf{A}$ is the surface operator [Berkhout and Verschuur, 1997]. The notation here is the same as the one used in Berkhout and Verschuur [1997], were the bold quantities refer to monochromatic data matrices with shot coordinates in their columns and receiver coordinates in their rows.

From equation 7.5.4 the multiples at a point (i,j) in the output matrix are given by $[\mathbf{M}]_{ij} = \sum[\mathbf{P}_0\mathbf{A}]_{ik}[\mathbf{P}]_{kj}$, and the set of contributing traces for each (i,j) trace is given by[1]

$$ConvGather_{(i,j)}(k) = [\mathbf{P}_0\mathbf{A}]_{ik}[\mathbf{P}]_{kj}. \tag{7.5.5}$$

On the other hand, in CL-SRME (and later in its focal version), we estimate $\mathbf{P}_0$ via an iterative process such that these primaries $\hat{\mathbf{P}}_0$ and their associated multiples $\hat{\mathbf{P}}_0\hat{\mathbf{A}}\mathbf{P}$ explain the input data. This is done by minimizing $J = ||\mathbf{P} - \hat{\mathbf{P}}_0 - \hat{\mathbf{P}}_0\hat{\mathbf{A}}\mathbf{P}||^2$. This minimization involves the following update

$$\Delta\mathbf{P}_0 = 2(\mathbf{P} - \hat{\mathbf{P}}_0 - \hat{\mathbf{P}}_0\hat{\mathbf{A}}\mathbf{P})(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})^H = 2\mathbf{V}(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})^H, \tag{7.5.6}$$

where $\mathbf{V}$ represents the residual data, i.e. the difference between the input data $\mathbf{P}$ and the modeled data $\hat{\mathbf{P}}_0 + \hat{\mathbf{P}}_0\mathbf{A}\mathbf{P}$. From equation 7.5.6 the set of contributing traces for each (i,j) trace is given by

$$CorrGather_{(i,j)}(k) = 2[\mathbf{V}]_{ik}[(\mathbf{I} + \mathbf{A}\mathbf{P})^H]_{kj}, \tag{7.5.7}$$

which requires a correlation-type product.

In order to extend the GSMP concept to include also correlation, it is necessary to change the convolution of the traces in step 5 of the previous section for a trace-by-trace correlation. If a $\mathbf{P}^H$ product is involved, it is also necessary to transpose (exchange source and receiver coordinates) the corresponding dataset. The lateral summation (index k in equations 7.5.6 and 7.5.5) can be done using one grid for the correlation and another for the convolution as explained in section 7.7.

The result of a convolution $\mathbf{PP}$ and a correlation $\mathbf{PP}^H$ can be seen in figure 7.3, the velocity model for synthetic data generated in the figure 1.4.

The associated contribution gathers from the datasets shown in figure 7.3 can be seen in figure 7.4. Only the casual components are kept.

7.6 Focal CL-SRME algorithm in 3D

Though successful in many scenarios, the original CL-SRME algorithm proposed by Lopez and Verschuur [2015b] has practical limitations when working in realistic 3D datasets. The main problem resides in the fact that, in order to apply the theory described there, the data must be organized inside a large dense data matrix [Berkhout, 1982; Kinneging et al., 1989], with the traces to interpolate replaced by

---

[1] The "set of contributing traces" is also known as the multiple contribution gather (MCG).

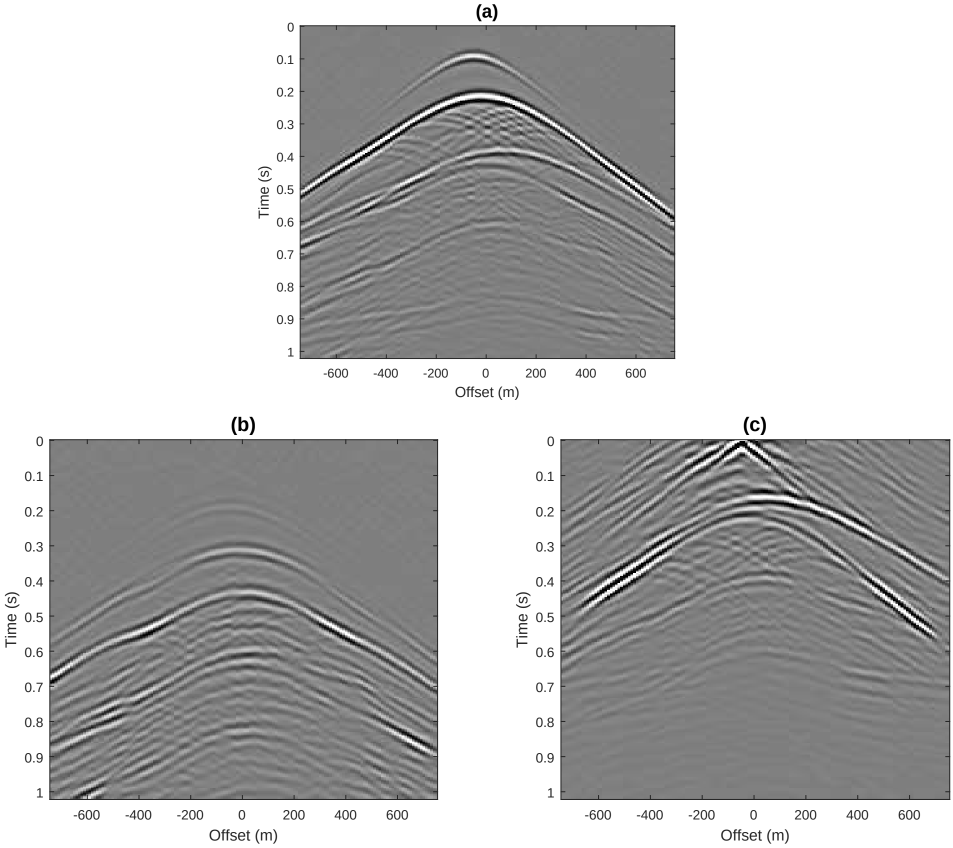

Figure 7.3: *Shot gathers from different types of products. (a) Input data* $\mathbf{P}$*, (b) convolution product* $\mathbf{PP}$ *and (c) correlation product* $\mathbf{PP}^H$*. Only the casual part of the datasets is shown here.*

zeros (or an initial interpolation). Taking into account that the data volume for real 3D acquisitions is generally very large ($\approx$ 10Tb), and that the input data matrix should be finely sampled, it becomes very unpractical to apply the former method in 3D, as it would "blow up" the memory just by creating the necessary matrix to start the process. We want to generalize the Focal CL-SRME algorithm to account for sparse trace distribution and generalized acquisition geometries.

In order to circumvent the former practical problems, we propose the 3D Focal CL-SRME method, which can be described as the Focal CL-SRME algorithm with the following adaptations:

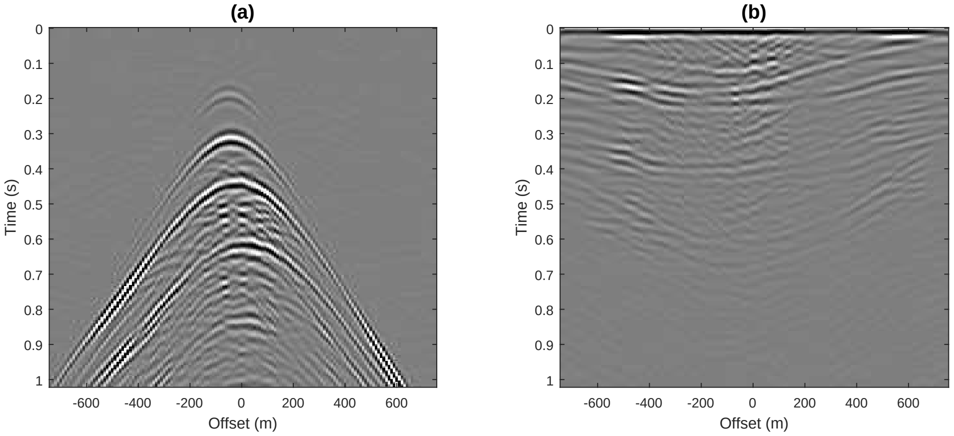

Figure 7.4: *Contribution gathers of the central trace from the products shown in figure 7.3. (a) Convolution contribution gather and (b) correlation contribution gather. The datasets in figures 7.4a and 7.4b should reproduce the central trace of figures 7.3b and 7.3c, respectively, when summed laterally. Only the casual components are shown.*

[**1**] The matrix products in the theory (data-data convolutions/correlations) are replaced by GSMP products.

[**2**] We generalize the notion of GSMP methodology to include both convolution and correlation type of products. This implies the construction of secondary wavefields by adding the contributions of two wavefields convolved or correlated with each other.

[**3**] For efficiency, we use simplified two-way focusing operators (see Berkhout and Verschuur [2006]). If the operator is assumed to be azimuth-invariant, it becomes is necessary to store only one finely-sampled shot gather. From which all the necessary traces for the focusing/defocusing operations can be extracted.

[**4**] To accelerate the convergence, we mute out the auto-correlation energy generated in the back-projection process. This can be easily achived with a first-break muting mask. See section 7.8 and figure 7.8.

[**5**] We apply a muting mask in noisy regions of the focal domain update (figure 7.8). This will eliminate extra aliasing noise to provide better convergence.

The most crucial shallow near-offset information is interpolated with the algorithm. If only near-offset information is required, only one focal domain is necessary (focused at the water-bottom). In this case, no additional data injection is required. Note that the back-projection mechanism will only provide reliable results if enough far-offset information is present (next section).

7.7 Contribution grids

From the theoretical point of view, in order to construct appropriate products, a summation must be done over all lateral coordinates to properly account for all possible scattering points. In practice, however, this summation is done over a finite grid in the lateral coordinates. The dimensions of this grid are important from a computational point of view. The smaller the contribution grid, the more efficient the calculation. However, if the grid is chosen to be too small, the resulting product will be erroneous [see also Lopez and Verschuur, 2015a].

In order to study the effect of a limited summation grid size in the products, we consider two data-product experiments using the dataset shown in figure 6.1. In the first experiment, the summation will add contributions from the near-offset locations only. In the second experiment, the summation will add contributions from the far-offset only. As we will see, traces within the near-offsets are the most contributing for convolution products, while traces within the far-offsets are the most contributing for correlation products.

Figure 7.5 shows the convolution and correlation gathers of the central trace for the synthetic shallow example shown in figure 6.1. Figure 7.5a,b shows the contribution gathers with the complete summation aperture, figure 7.5c,d shows the contribution gathers with the near-offsets removed and figure 7.5e,f shows the contribution gathers with the far-offsets removed. The result of the incomplete summations can be seen in figures 7.7 and 7.6. Convolution-products are strongly determined by the near-offsets, while correlation-products are determined mostly by the far-offsets.

The results in figures 7.5, 7.7 and 7.6 suggest a differential treatment of the contribution grid geometry for 3D. The proposed geometry for this research is shown in figure 7.9. For efficiency, we advice the use of a near-offset summation for convolution-products and a far-offset summation for correlation-products.

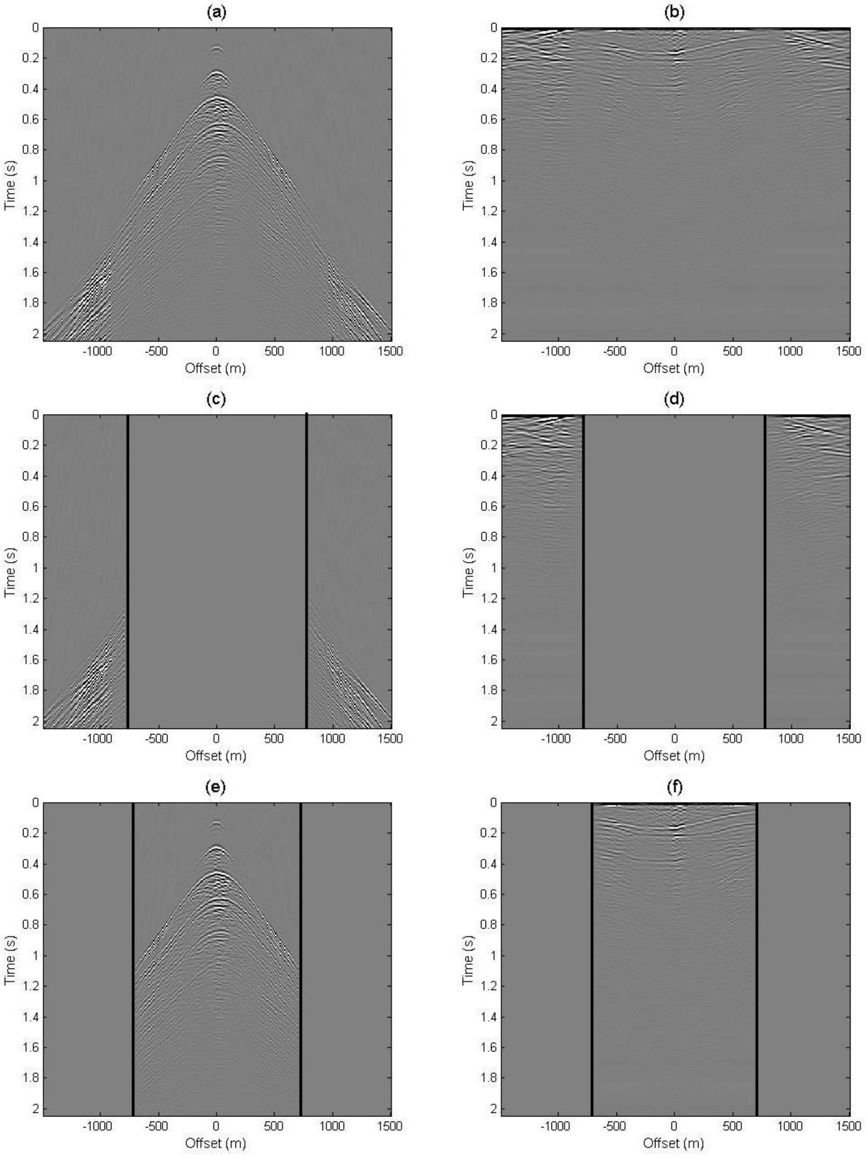

Figure 7.5: *Different summation apertures in the correlation and convolution contribution gathers. Full summation aperture for (a) convolution and (b) correlation. Far-offset summation only for (c) convolution and (d) correlation. Near-offset summation only for (e) convolution and (f) correlation. In experiments (c)-(f) 50% of the original summation aperture is removed. The black line shows the summation limits.*

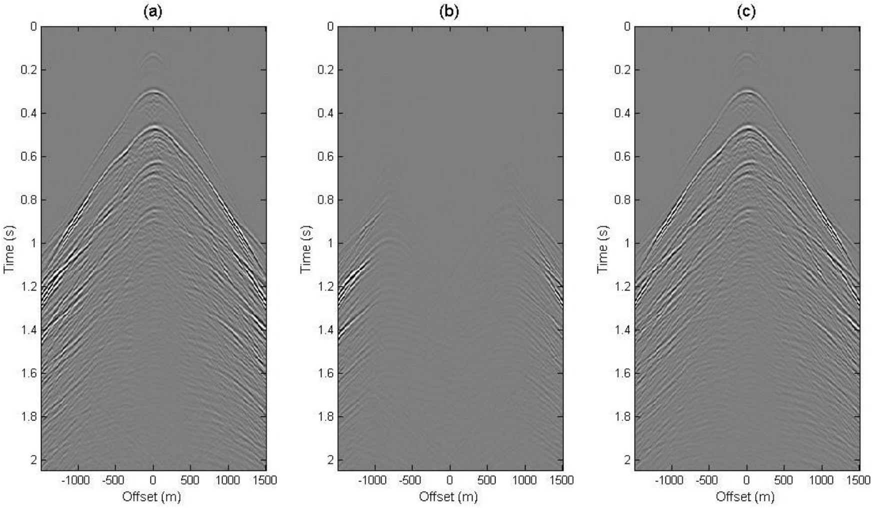

Figure 7.6: *Convolution products with different summation apertures. (a) Full summation, (b) Far-offset summation, (c) Near-offset summation.*

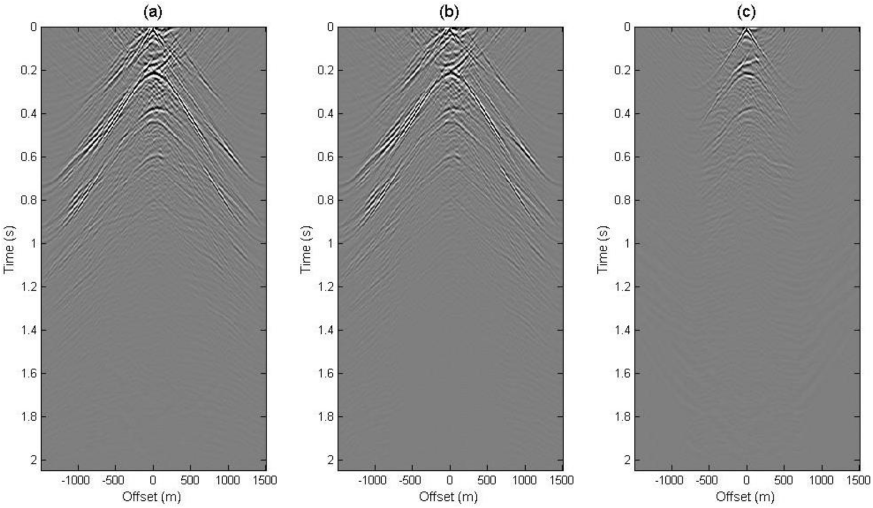

Figure 7.7: *Correlation products with different summation apertures. (a) Full summation, (b) Far-offset summation, (c) Near-offset summation.*

7.8 Algorithm outline

We can outline the 3D Focal CL-SRME algorithm by resuming its main steps. The iteration index has been omitted for simplicity. The M variable represents the desired number of focal domains. The $\mathcal{F}$ operator represents a Fourier Transform mapping time to frequency coordinates. Here $\mathbf{P}$ is the total data, $\mathbf{P}'$ is the input data (with zeros in the missing-data positions) and $\mathbf{P}''$ is the interpolated data (with zeros in the measured data positions). The operator $\mathbf{F}_m$ is a two-way focusing operator related to a given depth level [Berkhout and Verschuur, 2006].

3D Focal Closed-Loop SRME algorithm

```
Calculate the $\mathbf{F}_m$ operators using NMO times and velocities.
$\hat{\mathbf{X}}_m \Leftarrow \mathbf{0}$, $\mathbf{P} \Leftarrow \mathbf{P}'$
Calculate $\hat{\mathbf{A}}$ via an optimum filter such that $||\mathbf{P} - \mathbf{P}\hat{A}\mathbf{P}||^2 \rightarrow min$
while ($iter \leq maxIterations$) do
    $\mathbf{VQ}^H \Leftarrow (\mathbf{P} - \hat{\mathbf{P}}_0 - \hat{\mathbf{P}}_0\hat{\mathbf{A}}\mathbf{P})(\mathbf{I} + \hat{\mathbf{A}}\mathbf{P})^H$
    Apply a first-break muting to $\mathbf{VQ}^H$ (figure 7.8)
    If $iter = 1$ then apply near-offset amplitude matching
    while ($m \leq M$) do
        $\Delta\mathbf{X}_m^{(LS)} \Leftarrow 2\mathbf{F}_m^H\mathbf{VQ}^H$
        Apply focal domain muting to $\Delta\mathbf{X}_m^{(LS)}$ (figure 7.8)
        $\Delta\mathbf{X}_m^{(reg)} \Leftarrow \mathcal{F}\{-\lambda\nabla_{\hat{\mathbf{x}}_m}||\hat{\mathbf{x}}_m||_S\}$ (Appendix C)
        $\Delta\mathbf{X}_m \Leftarrow \Delta\mathbf{X}_m^{(LS)} + \Delta\mathbf{X}_m^{(reg)}$
        $\hat{\mathbf{X}}_m \Leftarrow \hat{\mathbf{X}}_m + \alpha_m\Delta\mathbf{X}_m$ ($\alpha_m$ s.t. $J \longrightarrow min$, equation 7.9.9)
    end while
    $\hat{\mathbf{P}}_0 \Leftarrow \sum_{m=1}^{M} \mathbf{F}_m\hat{\mathbf{X}}_m$
    Find $\hat{\mathbf{A}}$ according to equation 5.2.9
    $\mathbf{P}'' \Leftarrow (\hat{\mathbf{P}}_0 + \hat{\mathbf{P}}_0\hat{\mathbf{A}}\mathbf{P})''$
    $\mathbf{P} \Leftarrow \mathbf{P}' + \mathbf{P}''$
end while
```

7.9 Practical Considerations

The following considerations must be noted with respect to the 3D implementation of the Focal CL-SRME algorithm

- **Near-offset traces:** Near-offset interpolation is necessary (up to zero-offset) in order to properly apply the algorithm. The current algorithm interpolates the data in the shallow near-offset region via back-projection. However, other method must be employed for interpolating the deeper parts of the data. In most cases this is easily achieved.

- **Interpolation:** Due to its simplicity, we have chosen a DNMO interpolation method to generate the necessary traces for the required products. However, as noted in van Dedem and Verschuur [1998] this method might introduce artefacts in the predictions, especially if the offset correction is large. Alternatives to DNMO interpolation can be implemented in the algorithm without any significant change. Using a primitive on-the-fly interpolation method will limit the bandwidth of the predicted primaries to low frequencies only.

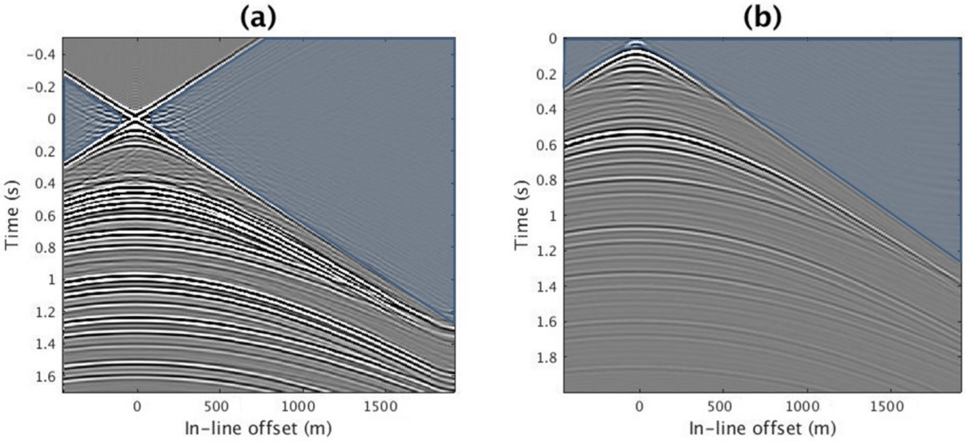

Figure 7.8: *Muting masks applied in 3D Focal CL-SRME (shaded area). (a) Muting over the focal domain* $\Delta\mathbf{X}_m^{(LS)}$*, (b) muting over the* $\mathbf{VQ}^H$ *variable. The muting over* $\Delta\mathbf{X}_m^{(LS)}$ *eliminates some of the aliasing noise generated by an incomplete acquisition. The muting over* $\mathbf{VQ}^H$ *eliminates the unwanted autocorrelation energy.*

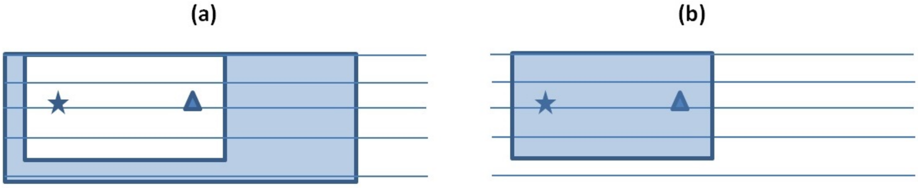

Figure 7.9: *Proposed contribution gathers for both (a) correlation and (b) convolution. The summation geometry is chosen to capture the largest contributions to the output products.*

- **Processing time:** The Focal CL-SRME algorithm requires more data products per iteration than conventional SRME. This implies a larger computational cost.

- **Focal operators:** We use a 3D homogeneous focal operator (see table 4.1) which we create along a fine offset grid from cero to the largest offset in the dataset. This operator is parameterized uniquely by the offset coordinate and the target NMO parameters (appex travel time and velocity). Whenever a focusing operation is needed, the GSMP product must be adopted to scan the stored traces for the necessary offset. Other implementations are possible (e.g. traces created on-the-fly inside the GSMP products), but they result less practical in terms of memory and/or processing time.

- **Step-lenght calculation:** Calculating the optimal step-lenghts (see appendix B.2) will be more efficient if all explicit data products are avoided. This can be achieved by using the identity $Tr(\mathbf{A}\mathbf{A}^H) = \sum_{i,j} |A_{ij}|^2$ (for any matrix $\mathbf{A}$ with lateral indices $\{i, j\}$) to transform equation B.2.18 into the component form

  $$\begin{aligned} J(\alpha_k) &= \sum_{ijt} |\mathcal{F}^{-1}\{\mathbf{P} - \sum_m \mathbf{F}_m \hat{\mathbf{X}}_m \mathbf{Q} - \alpha_k \delta_{km} \mathbf{F}_m \Delta \hat{\mathbf{X}}_m \mathbf{Q}\}|^2_{ijt} \quad (7.9.8) \\ &\quad + \lambda \sum_m ||\hat{\mathbf{x}}_m + \alpha_k \delta_{km} \Delta \hat{\mathbf{x}}_m||_S, \end{aligned}$$

  where the least-squares part of $J(\alpha_k)$ is now calculated over space-time coordinates.

7.10 Conclusions

We have introduced a primary estimation method aimed to applications in 3D data, especially in those cases in which accurate data interpolation becomes important (e.g. complex subsurface or shallow layers). The presented method brings computational advantages over the previously described Focal CL-SRME method, as it allows a generalized acquisition and keeps the processing data volumes within reasonable bounds. Examples of this method will be shown in the following chapter.

8

Results of the 3D algorithm

In this chapter we will present some initial results of the 3D Focal CL-SRME algorithm applied to both a 3D synthetic and a 3D field dataset.

8.1 3D synthetic example

Figure 8.1 shows the results applying the proposed method on a simple 3D dataset. The data is obtained from a velocity model with flat reflectors at 100m and 350m. The offset range is kept limited to prevent the data volume from growing too much. Figures 8.1a and 8.1b show the input data and the primary estimation for one selected 3D shot record. These figures are composed by a set of 2D panels placed next to each other. Each of these panels represents a 2D subset (i.e. like data from one streamer) from the full 3D shot, so they can be seen as transversal sections from the set of 3D reflection events. A grid of 60×60 shot positions and 60×60 receiver positions is used for the simulation, in which 2 out of 3 receiver positions are fully reconstructed. As we can see from figure 8.1, after 3D Focal CL-SRME, missing data is reconstructed and primaries are estimated. A slight multiple leakage in the primaries is observed due to the limited aperture in the example. However, this leakage will quickly disappear with growing acquisition apertures.

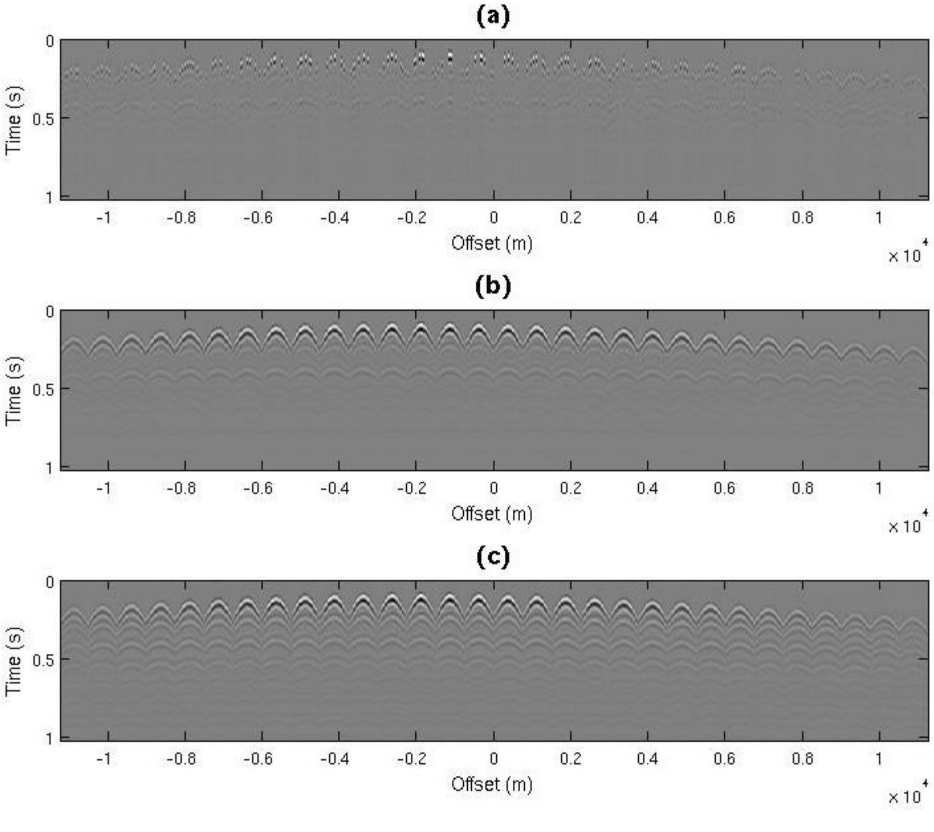

Figure 8.1: *Results from the synthetic 3D dataset generated using a velocity model with two horizons. (a) Input data with a 1:3 under-sampling ratio, (b) estimated primaries using 3D Focal CL-SRME and (c) estimated total data. The depth and velocity of the first layer (z=100m, c=1800m/s) were the only pieces of a-priori model information used to generate the focal operator and to obtain the results shown here.*

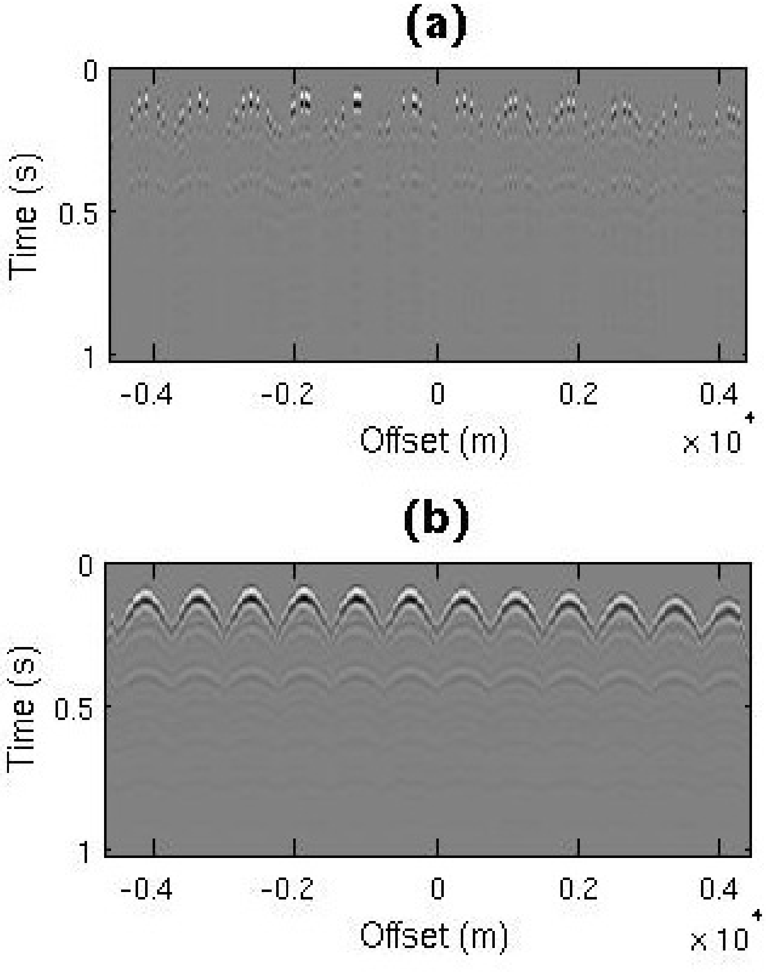

Figure 8.2: *A zoom from panels in figure 8.1 corresponding to (a) the input data with a 1:3 under-sampling ratio and (b) the estimated 3D Focal CL-SRME primaries.*

8.2 3D field example

A field data experiment is now presented. The dataset comes from a PGS marine survey. A subset of 6 streamers with 153 channels and 30 sources is selected for processing. The streamer in-line spacing is 12.5m and the cross-line spacing is 50m. The sources are located over a single line with a source-source spacing of 50m (see figure 8.4). The data corresponds to a relatively complex subsurface with an almost flat 150m water depth section, as depicted in figure 8.3. We will apply 3D Focal CL-SRME to this dataset in order to estimate primaries and multiples.

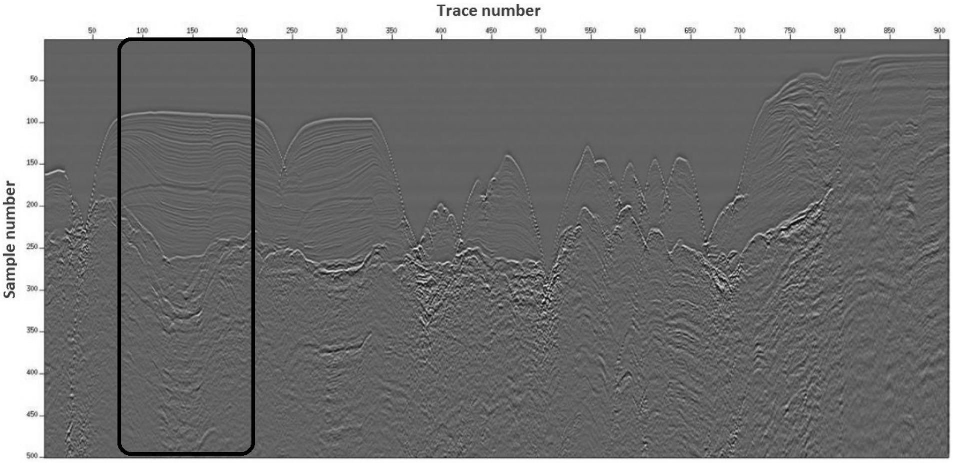

Figure 8.3: *Common channel gather showing the selected area (rectangular box) for the inversion. The water depth in the selected area is about 150m.*

As pre-processing steps, data is taken to a regular grid and near-offsets are interpolated with a parabolic-Radon technique [Kabir and Verschuur, 1995].

Here a single focal domain is calculated following the specifications in section 7.9. The focal operator is chosen to be that of a 3D homogeneous medium. It's inverse operator is estimated using the matching filter approach (section 4.2).

The results of the inversion are displayed in figures 8.5, 8.6, 8.7. The input dataset is shown in figure 8.5, the estimated primaries are shown in figure 8.6 and the estimated multiples are shown in figure 8.7. As we can see primaries and multiples are estimated in all cables.

Due to limited computational resources, the inversion was done over a relatively small dataset (only one sail-line, limited number of shots). Better results will be expected in a denser acquisition geometries. The inaccuracies introduced by the

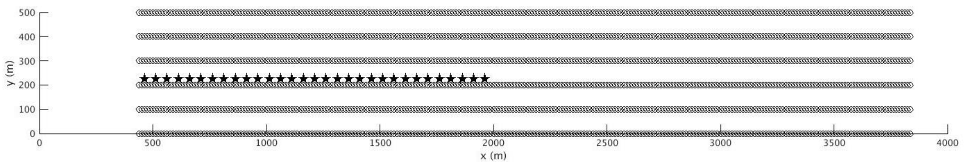

Figure 8.4: *Geometry of the acquisition used in the 3D field dataset example. The stars denote the sources and the circles the receivers. The figure shows 30 shots positions one sail-line with 6 streamers, each streamer with 153 channels.*

DNMO interpolation are also limiting factor for the primary/multiple separation, as interpolation errors produce estimations with frequencies lower than those expected. More advanced on-the-fly interpolation methods are expected to produce considerably better results and can be incorporated in the 3D Focal CL-SRME context without any significant change in the processing scheme.

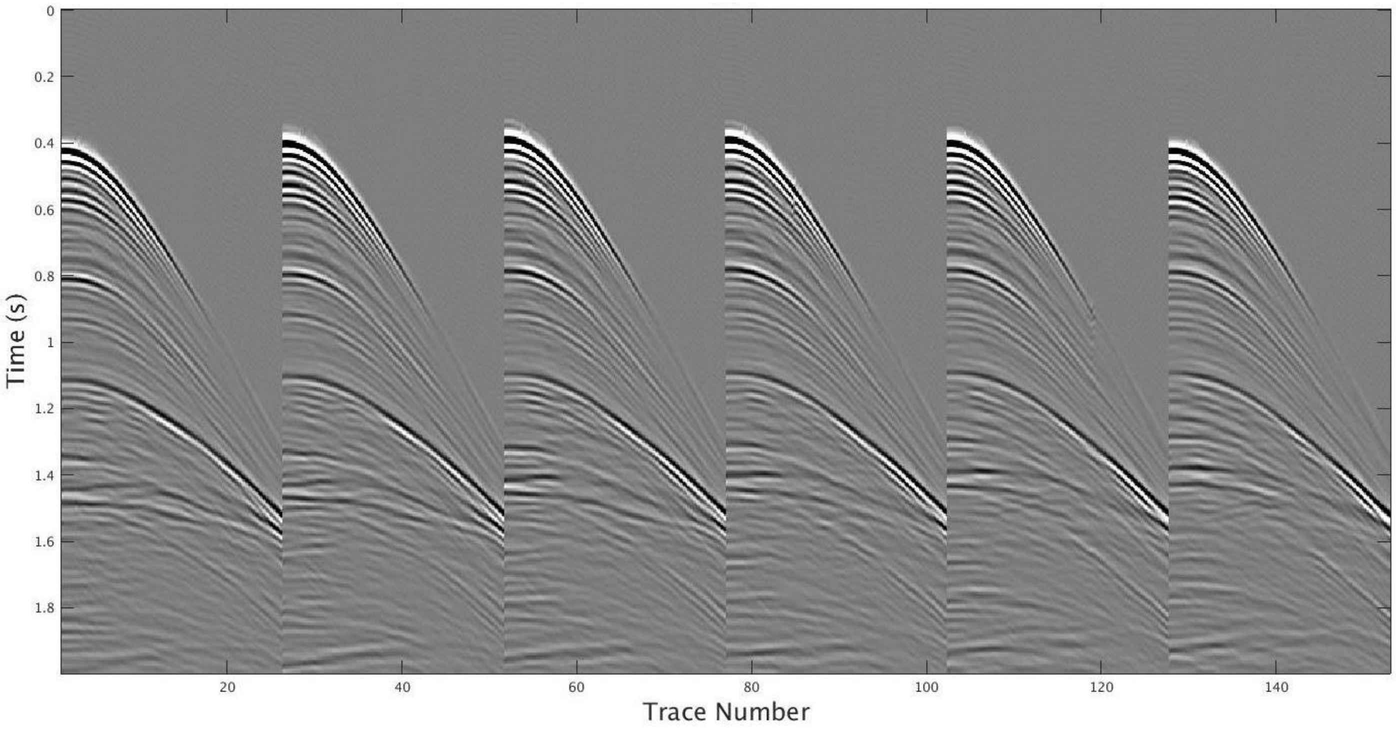

Figure 8.5: *3D common shot gather of the input data. Six streamer lines shown.*

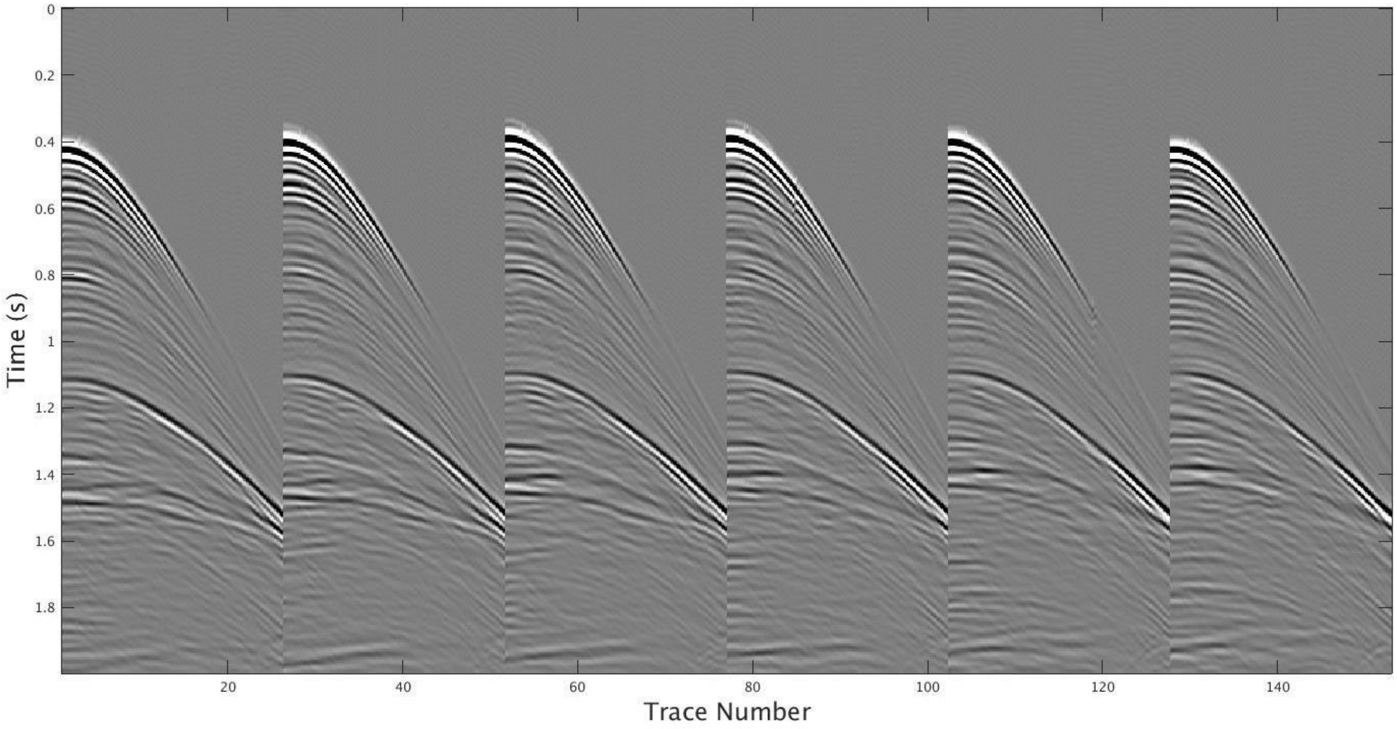

Figure 8.6: *3D common shot gather of the predicted primaries. Six streamer lines shown.*

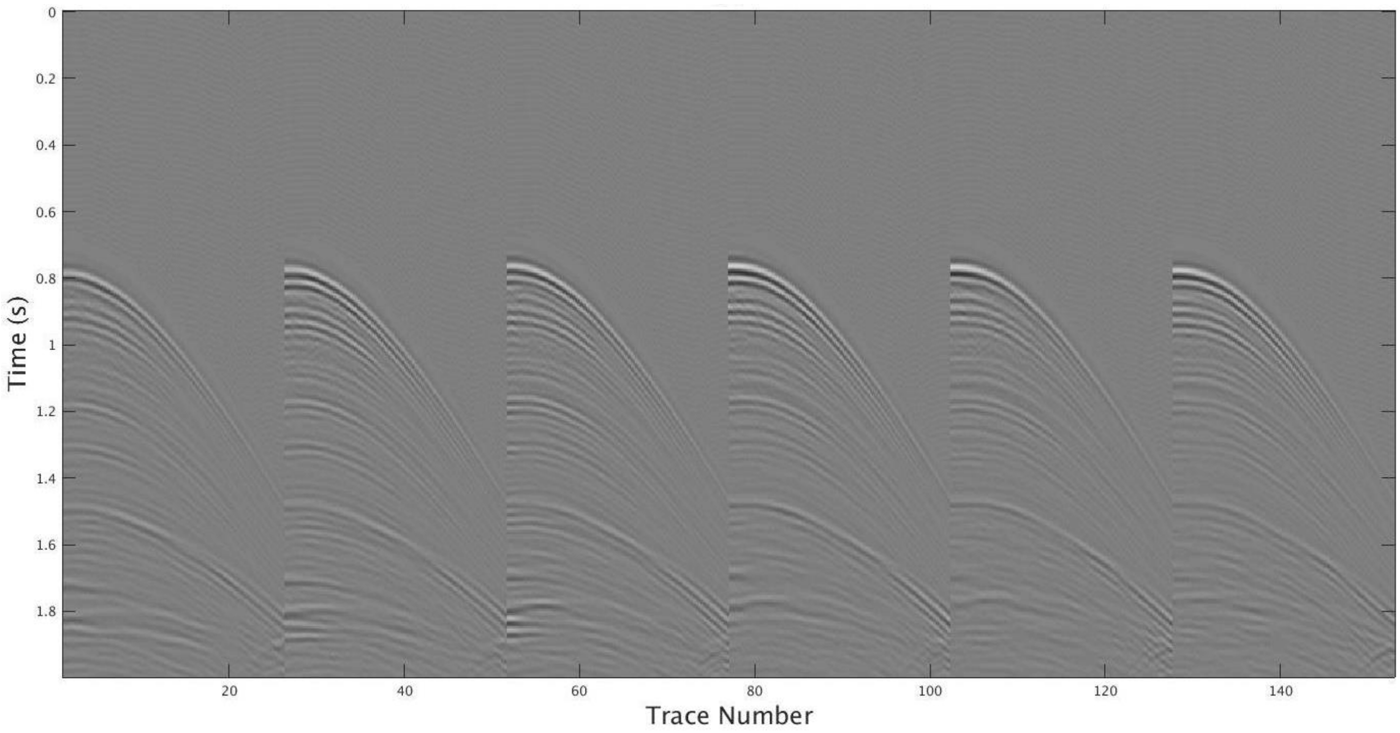

Figure 8.7: *3D common shot gather of the predicted multiples. Six streamer lines shown.*

9

Conclusions, recommendations, discussion and future research

9.1 Conclusions

We have introduced a new algorithm for primary estimation that uses inversion to estimate the primaries $\mathbf{P}_0$ and the surface operator A. This algorithm uses SRME's decomposition of $\mathbf{P}$ into $(\mathbf{P}_0, A)$. But instead of solving $\mathbf{P}_0$ via the iterative reconstruction of the inverse scattering series of $\mathbf{P}$, CL-SRME uses inversion to estimate $\mathbf{P}_0$ from the data itself, choosing the prediction equation from SRME as the forward model for the inversion of the $(\mathbf{P}_0, A)$ parameters.

We have also extended the CL-SRME methodology such that it allows primary estimation together with data interpolation. In order to achieve an accurate reconstruction we take the primary wavefield $\mathbf{P}_0$ to a transform domain in which the data gaps and/or interpolation errors are mapped to aliasing artifacts. The ultimate goal of the transform domain is to represent the primary reflections with the smallest possible number of parameters.

We use the focal domain, which is based on wavefield extrapolation operators, as it provides a high level of data compression. Being the product of a physics-based transform rather than a mathematical one, the focal domain also introduces additional physical constraints on the estimations, namely: wavefield continuity and wave-equation consistency. A sparse-promoting regularization is chosen as a method to guide the estimates of the focal domain inversion towards the elimination of the aliasing artifacts in these domains. Once all the artifacts have been cleared, the primary estimate becomes fully sampled.

Being a simultaneous reconstruction/multiple estimation algorithm, we expect the proposed method to perform better than other methods based on interpolation prior to the primary estimation. By allowing the missing data and the primary information to be simultaneously estimated, we allow their estimations to be simultaneously adjusted to optimally fit each other. Here, data reconstruction helps obtaining better multiple estimates and the physical primary-multiple relationship helps constraining the data interpolation. In other words, information from the multiples is used to reconstruct the missing data.

Given an under-sampled dataset, the Focal CL-SRME (FCL-SRME) algorithm can perform simultaneous data reconstruction and primary-multiple separation. The required focal operators can be obtained via simple estimation of the RMS velocities of the target reflectors.

In order to practically incorporate the large data volumes present in typical 3D datasets we have modified the processing scheme of the Focal CL-SRME algorithm in order to provide a practical solution for 3D datasets. The 3D Focal CL-SRME algorithm is thus a primary estimation method aimed to applications in 3D data, especially in those cases in which accurate data interpolation becomes important (e.g. complex subsurface and/or shallow water). The presented method combines the accurate interpolation advantages of the Focal CL-SRME algorithm with the efficiency of the GSMP method. A more precise interpolation scheme gives the current algorithm more advantages over traditional methods in the places where accuracy is really crucial (i.e. at the near-offsets). In this way, the current scheme provides more flexibility in the larger offsets while keeping the accuracy in the short offsets, which are the most relevant in primary/multiple separation. The capacities of the proposed algorithm have been demonstrated through several examples on both synthetic and field data.

Besides the current reconstruction-oriented methodology, the mathematical theory for multiple prediction presented in chapter 3 can serve as a starting point for many applications seeking to use the information contained in the multiples without the need of migration. An easy and efficient implementation should make the present algorithm attractive for future applications. Also, the theory presented in the appendices might serve as useful background for future researchers, even beyond the fields of multiple prediction or data reconstruction.

9.2 Recommendations

FCL-SRME has been shown to have very good interpolation capabilities, as it can handle large near-offset gaps, very shallow data and large under-sampling reconstruction (see chapter 6 for examples). Therefore, it is possible to consider FCL-SRME as an interpolation tool by itself. In this sense, FCL-SRME can be thought to be a robust version of the algorithm presented by Kutscha and Verschuur [2012]. The advantage is that in FCL-SRME an initial multiple separation step is not necessary. Actually, in FCL-SRME the primary/multiple relationship is used in to aid the reconstruction.

9.3 Discussion: Which algorithm to use?

In the literature, there are currently two possible parametrizations for inversion-oriented multiple estimation, namely: inverting for $\{\mathbf{P}_0, \mathbf{A}\}$ or inverting for $\{\mathbf{X}_0, \mathbf{S}\}$. The first one is used in the previously described robust EPSI (REPSI) algorithm[1] [Lin and Herrmann, 2013, 2016] and the second one is presented in this thesis [see also Lopez and Verschuur, 2015b]. It is important to compare this two methodologies in order to decide which algorithm is more suitable for a given problem[2].

Now, we will discuss the differences and similarities between REPSI and FCL-SRME. The main properties of both algorithms are shown below:

REPSI

- Parametrization in terms of $\{\mathbf{X}_0, \mathbf{S}\}$.
- Solves a basis persuit denoising algorithm.
- Uses an SPGL1 solver (based on local gradients) [van den Berg and Friedlander, 2008, 2011].
- Starts from a sparse unpopulated estimate for the primaries and move to more populated solutions with iterations.
- Proposes an L1 convexification of the L0 norm in the original EPSI (better mathematical properties).
- Minimum energy replaced by a sparsity constraint.

FCL-SRME

- Parametrization in terms of $\{\mathbf{P}_0, \mathbf{A}\}$.
- Solves a L1-L2 regularized least squares problem[3].
- Uses a conjugate gradients methodology (based on local descent) [Aster et al., 2013].

---

[1] We address REPSI rather than the original EPSI as REPSI is known to be a more robust algorithm, providing always the most possible sparse solution and presenting less artefacts than its predecesor.

[2] This section provides a experience-based quantitative comparison between REPSI and FCL-SRME. Most of the remarks are derived from our practical experience.

[3] Other norms are also possible, see appendix C.

Table 9.1: *Main benefits and limitations of REPSI and FCL-SRME*

Algorithm	Benefits	Limitations
REPSI	• Estimates the wavelet • Near-offset reconstruction for the entire time range	• Computationally expensive • No under-sampling reconstruction • Not applicable in 3D
FCL-SRME	• Computationally efficient (tens of iterations required) • Can handle data gaps due to both near-offsets and under-sampling • Easy implementation • Applicable for 2D and 3D datasets	• Uses minimum energy in the firsts iterations • Near-offset reconstruction only for the early reflectors (early times)

- Starts with a primary estimate close to the input dataset and remove the multiples with iterations.
- Requires a focal transform to impose sparsity in the solutions.
- Minimum energy present in the early iterations.

The main benefits and limitations of REPSI and FCL-SRME are given by table 9.1.

Which algorithm to choose will then depend on the application interest. If a method for estimating Green's functions (e.g. for interferometric applications) or source wavelet (e.g. for imaging) is desired, then the REPSI algorithm is probably the best choice. REPSI might be also very useful for 2D datasets. On the other hand, if the application is tailored to 3D primary/multiple separation, then the FCL-SRME algorithm is the best choice. Note that the differences between these algorithms lie mainly in the shallow-data sections. For the deeper data sections the performance of both FCL-SRME and REPSI is comparable to that of SRME.

9.4 Future research

For some applications the use of the FCL-SRME algorithm by itself may result in a very useful framework to build extensions/improvements over the SRME algorithm. The advantages of an inversion-type approach for SRME over the series-based methodology are the following:

[1] The convergence to the final solution might be controlled by manipulating the gradient (pre-conditioning). This can improve the convergency to the desired solution, e.g. applying lateral artefact-muting windows in the focal domain gradients is known to boost-up the data reconstruction capabilities of FCL-SRME (see figure 5.1).

[2] Joint parameter inversion becomes possible. This entails the use of the information in the multiples to provide a better estimate of other parameters of interest. For example, Vrolijk and Verschuur [2016] proved that a deghosting algorithm can benefit from the information in the multiples using Closed-Loop SRME with integrated deghosting.

[3] Multiple prediction with a complex source wavefield is also an option. This might open the possibility of using multiples for improving deblending algorithms. A similar extension was already developed for EPSI [van Groenestijn and Verschuur, 2011].

[4] Prior-knowledge can be introduced by controlling the initial estimates of both $\mathbf{A}$ and $\mathbf{P}_0$. This will guide the solution towards a different inversion path, which will lead to different (and hopefully better) results. This is expected from a local descent algorithm.

The disadvantage of FCL-SRME over SRME lies mainly in the fact that an inversion scheme is typically more computationally expensive than its series-based counterpart.

Compared with 3D SRME, the 3D FCL-SRME algorithm is still a computationally expensive solution, due to a larger amount of necessary matrix products. This can make the 3D algorithm impractical if limited computational resources are available. Future research can be tailored in methods to relax these constraints.

A

Equivalent expression for the multiple scattering wavefield

A.1 Expressions in the ($\mathrm{X_0}$,S) parameterization

The following list of equivalent expressions for the surface multiple scattering problem can be obtain directly from an similar list given in Chapter 2 (equations 2.3.5-2.3.10) using the replacement $\mathbf{P}_0 = \mathbf{X}_0\mathbf{S}$ and $\mathbf{A} = \mathbf{S}^{-1}\mathbf{R}^{\cap}$. Here we are define $\mathbf{X}$ as the total impulse response of the earth (i.e. primaries+multiples), such that $\mathbf{P} = \mathbf{XS}$ is fulfilled[1].

- **Series**:

$$\begin{aligned} \mathbf{X} &= \mathbf{X}_0 + (\mathbf{X}_0\mathbf{R}^{\cap})\mathbf{X}_0 + (\mathbf{X}_0\mathbf{R}^{\cap})^2\mathbf{X}_0 + (\mathbf{X}_0\mathbf{R}^{\cap})^3\mathbf{X}_0 + \cdots && \text{(A.1.1)} \\ \mathbf{X}_0 &= \mathbf{X} - \mathbf{X}(\mathbf{R}^{\cap}\mathbf{X}) + \mathbf{X}(\mathbf{R}^{\cap}\mathbf{X})^2 - \mathbf{X}(\mathbf{R}^{\cap}\mathbf{X})^3 + \cdots . && \text{(A.1.2)} \end{aligned}$$

- **Recursive**:

$$\begin{aligned} \mathbf{X} &= \mathbf{X}_0 + \mathbf{X}_0\mathbf{R}^{\cap}\mathbf{X} && \text{(A.1.3)} \\ \mathbf{X}_0 &= \mathbf{X} - \mathbf{X}_0\mathbf{R}^{\cap}\mathbf{X}. && \text{(A.1.4)} \end{aligned}$$

- **Inverse space**:

$$\begin{aligned} \mathbf{X}^{-1} &= \mathbf{X}_0^{-1} - \mathbf{R}^{\cap} && \text{(A.1.5)} \\ \mathbf{X}_0^{-1} &= \mathbf{X}^{-1} + \mathbf{R}^{\cap}. && \text{(A.1.6)} \end{aligned}$$

---

[1] The replacements $\mathbf{X}_0\mathbf{R}^{\cap} = \mathbf{P}_0\mathbf{A}$ and $\mathbf{X}\mathbf{R}^{\cap} = \mathbf{PA}$ might be useful here.

A.2 Expressions in the (P_0,A) parameterization

For completeness we also include the list of expressions in the ($\mathbf{P}_0$,$\mathbf{A}$) parameterization. The derivation of these expressions can be found in Chapter 2.

- **Series**:

$$\begin{aligned}\mathbf{P} &= \mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})\mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})^2\mathbf{P}_0 + (\mathbf{P}_0\mathbf{A})^3\mathbf{P}_0 + \cdots && \text{(A.2.7)}\\ \mathbf{P}_0 &= \mathbf{P} - \mathbf{P}(\mathbf{A}\mathbf{P}) + \mathbf{P}(\mathbf{A}\mathbf{P})^2 - \mathbf{P}(\mathbf{A}\mathbf{P})^3 + \cdots . && \text{(A.2.8)}\end{aligned}$$

- **Recursive**:

$$\begin{aligned}\mathbf{P} &= \mathbf{P}_0 + \mathbf{P}_0\mathbf{A}\mathbf{P} && \text{(A.2.9)}\\ \mathbf{P}_0 &= \mathbf{P} - \mathbf{P}_0\mathbf{A}\mathbf{P}. && \text{(A.2.10)}\end{aligned}$$

- **Inverse space**:

$$\begin{aligned}\mathbf{P}^{-1} &= \mathbf{P}_0^{-1} - \mathbf{A} && \text{(A.2.11)}\\ \mathbf{P}_0^{-1} &= \mathbf{P}^{-1} + \mathbf{A}. && \text{(A.2.12)}\end{aligned}$$

For more details about these expressions, the reader is refered to Verschuur et al. [1992]; Verschuur and Berkhout [1997]; Fokkema and van den Berg [1993].

B

Calculations on complex-valued matrix functionals with the Frobenius norm

In this appendix we will review the procedure to calculate the gradients and the optimal-lenghts of complex-valued matrix functionals with the Frobeniuos norm. These type of functionals are very common in the DELPHI literature, so its important to be able to derive the expressions necessary for an inversion algorithm.

B.1 Gradients

B.1.1 General Theory

Let

$$J(\mathbf{Z}) = Tr(\mathbf{V}(\mathbf{Z})\mathbf{V}(\mathbf{Z})^H) = ||\mathbf{V}(\mathbf{Z})||_{Fr}^2, \tag{B.1.1}$$

be a matrix functional such that $J \in M_n(\mathbb{C}) \rightarrow \mathbb{C}$ where $M_n(\mathbb{C})$ is the space of $n \times n$ matrices over the complex field $\mathbb{C}$. We choose $\mathbf{V} : M_n(\mathbb{C}) \rightarrow M_n(\mathbb{C})$ to be a differentiable function of $\mathbf{Z}$ and $||\cdot||_{Fr}^2$ to be the Frobenius norm. Typically $\mathbf{V}$ would correspond to the inversion residual in a descent algorithm. We are interested in calculating the gradient $\nabla_{\mathbf{Z}} J(\mathbf{Z})$ such that we can use it in an inversion algorithm. However, due to the appearance of both $\mathbf{Z}$ and $\mathbf{Z}^H$ in the definition of $J(\mathbf{Z})$, the functional defined in equation B.1.1 is non-analytical, thus non-differentiable. This implies that we need to redefine the concept of gradient for this type of functionals. We can resolve this problem by defining our gradient to be:

$$\nabla_{\mathbf{Z}} J(\mathbf{Z}) := 2\frac{dJ(\mathbf{Z})}{d\mathbf{Z}^*} = \frac{\partial J}{\partial \mathcal{R}\mathbf{Z}} + i\frac{\partial J}{\partial \mathcal{I}\mathbf{Z}}, \tag{B.1.2}$$

where $\mathcal{R}\mathbf{Z}$ and $\mathcal{I}\mathbf{Z}$ represent the real and imaginary parts of $\mathbf{Z}$ respectively. Equation B.1.2 is equal to zero when $J(\mathbf{Z})$ is an analytic function of $\mathbf{Z}$ and non-zero other-wise.

The gradient in equation B.1.2 applied to equation B.1.1 can be calculated using the following algorithm:

[1] Take the exterior derivative of J: $J \longrightarrow dJ$.

[2] Use the linearity and the Leibniz rule to write dJ in $d\mathbf{Z}^*$ terms. Note that the gradient definition uses only $d\mathbf{Z}^*$, so any $d\mathbf{Z}$ variation is zero:

$$dJ = dTr(\mathbf{V}^H\mathbf{V}) = Tr(d\mathbf{V}^H\mathbf{V} + \mathbf{V}^H d\mathbf{V}) \tag{B.1.3}$$

[3] Take dJ into a $dJ = Tr(f(\mathbf{Z})d\mathbf{Z}^*)$ form (with f a function of $\mathbf{Z}$).

[4] Use the property:

$$dJ = Tr(f(\mathbf{Z})d\mathbf{Z}^*) \longrightarrow \frac{dJ(\mathbf{Z})}{d\mathbf{Z}^*} = f(\mathbf{Z}) \tag{B.1.4}$$

to calculate the required derivative.

[5] Scale the result by 2 to obtain $\nabla_{\mathbf{Z}} J(\mathbf{Z})$.

Note descent algorithms require the negative of the gradient $-\nabla_{\mathbf{Z}} J(\mathbf{Z})$ as descent direction. Also note that if $\mathbf{Z} = \mathbf{Z}^T$, $\mathbf{Z}^*$ can be replaced by $\mathbf{Z}^H$ in all the steps above.

B.1.2 Proof of equation 5.2.6

In this section we will calculate the least-squares gradient of a functional J (in equation 5.2.6) with respect to the variable $\hat{\mathbf{X}}_i(\omega')$. Start by introducing the variable $j^{(LS)} = ||\mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q}||_{Fr}$, then $J^{(LS)}$ can be written as

$$J^{(LS)} = \sum_{\omega} ||\mathbf{P} - \sum_{m} \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q}||_{Fr} = \sum_{\omega} j^{(LS)}(\hat{\mathbf{X}}_m, \hat{\mathbf{X}}_m^H; \omega). \tag{B.1.5}$$

Taking the gradient $\nabla_{\hat{\mathbf{X}}_i(\omega')}$ on this expression we get

$$\begin{aligned}
\nabla_{\hat{\mathbf{X}}_i(\omega')} J^{(LS)} &= \frac{2dJ^{(LS)}}{d\hat{\mathbf{X}}_i(\omega')^H} \\
&= \sum_{\omega} \frac{2dj^{(LS)}(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_i^H; \omega)}{d\hat{\mathbf{X}}_i(\omega')^H} \\
&= \frac{2dj^{(LS)}(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_i^H; \omega')}{d\hat{\mathbf{X}}_i(\omega')^H} \\
&= \nabla_{\hat{\mathbf{X}}_i(\omega')} j^{(LS)}(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_i^H; \omega'). \qquad \text{(B.1.6)}
\end{aligned}$$

The elimination of $\sum_\omega$ in the above expression is due to the fact that all the ω frequencies are mutually independent, then the only non-zero contribution to the derivative comes when $\omega = \omega'$. The calculation of the derivative of $j^{(LS)}$ will now follow the steps outlined above (for simplicity we will drop now the explicit dependence of ω'). Start rewritting $j^{(LS)}$ as

$$j^{(LS)}(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_i^H) = Tr[(\mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q})^H (\mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q})] = Tr[\mathbf{V}^H \mathbf{V}], \qquad \text{(B.1.7)}$$

where $\mathbf{V} := \mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q}$ is the data residual. Now taking the exterior derivative in $j^{(LS)}$ we get

$$dj^{(LS)}(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_i^H) = dTr[\mathbf{V}^H \mathbf{V}] = Tr[d\mathbf{V}^H \mathbf{V} + \mathbf{V}^H d\mathbf{V}], \qquad \text{(B.1.8)}$$

where the last step corresponds to the Leibnitz product rule. Here, the variables $\hat{\mathbf{X}}_i^H$ and $\hat{\mathbf{X}}_i$ can be taken as independent. Then, because we are considering variations over the $\hat{\mathbf{X}}_i^H$ variable only ($d\hat{\mathbf{X}}_m^H = 0$ if $m \neq i$), we can see that $d\mathbf{V}(\hat{\mathbf{X}}_i) = \mathbf{V}(d\hat{\mathbf{X}}_i) = \mathbf{0}$. Therefore we can write

$$\begin{aligned}
dj^{(LS)}(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_i^H) &= Tr[d\mathbf{V}^H \mathbf{V}] \\
&= Tr[d(\mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q})^H \mathbf{V}] \\
&= Tr[-(\mathbf{W}_i^T d\hat{\mathbf{X}}_i \mathbf{W}_i \mathbf{Q})^H \mathbf{V}] \\
&= Tr[-\mathbf{Q}^H \mathbf{W}_i^H d\hat{\mathbf{X}}_i^H \mathbf{W}_i^* \mathbf{V}] \\
&= Tr[-\mathbf{W}_i^* \mathbf{V} \mathbf{Q}^H \mathbf{W}_i^H d\hat{\mathbf{X}}_i^H] \\
&= Tr[-\mathbf{W}_i^* (\mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q}) \mathbf{Q}^H \mathbf{W}_i^H d\hat{\mathbf{X}}_i^H], \qquad \text{(B.1.9)}
\end{aligned}$$

so

$$\frac{dj^{(LS)}(\hat{\mathbf{X}}_i,\hat{\mathbf{X}}_i^H)}{d\hat{\mathbf{X}}_i^H}=\frac{dJ^{(LS)}(\hat{\mathbf{X}}_i,\hat{\mathbf{X}}_i^H)}{d\hat{\mathbf{X}}_i^H}=-\mathbf{W}_i^*(\mathbf{P}-\sum_m \mathbf{W}_m^T\hat{\mathbf{X}}_m\mathbf{W}_m\mathbf{Q})\mathbf{Q}^H\mathbf{W}_i^H, \tag{B.1.10}$$

which leads to:

$$\nabla_{\hat{\mathbf{X}}_i}J^{(LS)}=-2\mathbf{W}_i^*(\mathbf{P}-\sum_m \mathbf{W}_m^T\hat{\mathbf{X}}_m\mathbf{W}_m\mathbf{Q})\mathbf{Q}^H\mathbf{W}_i^H, \tag{B.1.11}$$

which is the desired gradient. For more information about derivatives over complex matrices please refer to Schreier [2010] and Hjørungnes [2011].

B.2 Optimal-descent lengths

■ B.2.1 General Theory

Calculating the step-length $\alpha \in \mathbb{R}$ such that the updating $\mathbf{Z}^{(i+1)} = \mathbf{Z}^{(i)} + \alpha\Delta\mathbf{Z}^{(i)}$ minimizes J is easily done by imposing the extremallity condition over $J(\alpha)$, i.e. solving the scalar valued equation $\partial_\alpha J(\mathbf{Z} + \alpha\Delta\mathbf{Z}) = 0$. Given the function $J(\alpha)$ other line-search methods are also avilable in the literature [Pedregal, 2013]. The form of the solution will depend upon the particular structure of $J(\mathbf{Z})$.

■ B.2.2 Step-lenght calculation with no sparsity

We will solve for α_i value in equation 5.2.8 assuming that there is no sparsity constraint. The value found here will be used as starting value for a the line-searching algorithm in the case of non-zero sparsity (next section).

Start with the objective function

$$J=\sum_\omega Tr[(\mathbf{P}-\sum_m \mathbf{W}_m^T\hat{\mathbf{X}}_m\mathbf{W}_m\mathbf{Q})^H(\mathbf{P}-\sum_m \mathbf{W}_m^T\hat{\mathbf{X}}_m\mathbf{W}_m\mathbf{Q})], \tag{B.2.12}$$

then $J(\hat{\mathbf{X}}_i + \alpha_i\Delta\hat{\mathbf{X}}_i)$ can be written as

$$\begin{aligned} J(\alpha_i) \;&=\; \sum_\omega Tr[(\mathbf{P}-\sum_m \mathbf{W}_m^T\hat{\mathbf{X}}_m\mathbf{W}_m\mathbf{Q}-\alpha_i\mathbf{W}_i^T\Delta\hat{\mathbf{X}}_i\mathbf{W}_i\mathbf{Q})^H \\ &\quad (\mathbf{P}-\sum_m \mathbf{W}_m^T\hat{\mathbf{X}}_m\mathbf{W}_m\mathbf{Q}-\alpha_i\mathbf{W}_i^T\Delta\hat{\mathbf{X}}_i\mathbf{W}_i\mathbf{Q})], \end{aligned} \tag{B.2.13}$$

choosing $\mathbf{V} := \mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q}$ and $\mathbf{K}_i := \mathbf{W}_i^T \Delta\hat{\mathbf{X}}_i \mathbf{W}_i \mathbf{Q}$ this simplifies to

$$\begin{aligned} J(\alpha_i) &= \sum_\omega Tr[(\mathbf{V} - \alpha_i \mathbf{K}_i)^H (\mathbf{V} - \alpha_i \mathbf{K}_i)] \\ &= \sum_\omega Tr[\mathbf{V}^H \mathbf{V} - \alpha_i \mathbf{V}^H \mathbf{K}_i - \alpha_i \mathbf{K}_i^H \mathbf{V} + \alpha_i^2 \mathbf{K}_i^H \mathbf{K}_i], \end{aligned} \tag{B.2.14}$$

imposing the extremallity condition we get

$$\partial_{\alpha_i} J(\alpha_i) = \sum_\omega Tr[-\mathbf{V}^H \mathbf{K}_i - \mathbf{K}_i^H \mathbf{V} + 2\alpha_i \mathbf{K}_i^H \mathbf{K}_i] = 0, \tag{B.2.15}$$

$$\alpha_i = \frac{\sum_\omega Tr[\mathbf{V}^H \mathbf{K}_i + \mathbf{K}_i^H \mathbf{V}]}{2\sum_\omega Tr[\mathbf{K}_i^H \mathbf{K}_i]} = \frac{\sum_\omega Re\{Tr[\mathbf{V}^H \mathbf{K}_i]\}}{\sum_\omega Tr[\mathbf{K}_i^H \mathbf{K}_i]}. \tag{B.2.16}$$

■ B.2.3 Step-lenght calculation with sparsity

In the case of a non-linear regularization like the one described in Chapter 5 our objective function looks like

$$J = \sum_\omega Tr[(\mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q})^H (\mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q})] + \lambda \sum_m ||\hat{\mathbf{x}}_m||_S \tag{B.2.17}$$

with $\hat{\mathbf{x}}_m$ is the inverse Fourier transform of the data-cube of $\hat{\mathbf{X}}_m$ and λ is a user-defined regularization constant (typically $\lambda \approx 10^{-2}$). The norm $||\cdot||_S$ represents any sparsity-promoting norm of preference (e.g. L1), which applies to every element in $\hat{\mathbf{x}}_m$.

If we include a variation $J(\hat{\mathbf{X}}_i + \alpha_i \Delta\hat{\mathbf{X}}_i)$ our objective function becomes

$$J(\alpha_i) = \sum_\omega Tr[\mathbf{V}^H \mathbf{V} - \alpha_i \mathbf{V}^H \mathbf{K}_i - \alpha_i \mathbf{K}_i^H \mathbf{V} + \alpha_i^2 \mathbf{K}_i^H \mathbf{K}_i] + \lambda \sum_m ||\hat{\mathbf{x}}_m + \alpha_i \delta_{im} \Delta\hat{\mathbf{x}}_m||_S, \tag{B.2.18}$$

with

$$\mathbf{V} := \mathbf{P} - \sum_m \mathbf{W}_m^T \hat{\mathbf{X}}_m \mathbf{W}_m \mathbf{Q} \qquad \text{and} \qquad \mathbf{K}_i := \mathbf{W}_i^T \Delta\hat{\mathbf{X}}_i \mathbf{W} \mathbf{Q}. \tag{B.2.19}$$

To find the proper step-lenght we impose the condition $J(\hat{\mathbf{X}}_i + \alpha_i \Delta\hat{\mathbf{X}}_i) \longrightarrow min$ and then calculate the α_i that satisfies such condition. Note that, the function $J(\alpha_i)$ above is a simple one-parameter function, which is in principle easy to minimize using any line-search method. The value obtained in the no-sparsity constraint case can be used as starting value for many of these algorithms.

C

The sparsity constraint

A sparsity constraint is a condition applied to an inverse problem in order to drive the solution towards a minimum in the objective function in which the energy appears concentrated in a small number of parameters rather than spread around in many them. Typically this type of constraint is used to regularize a problem, thus reducing the number of solutions available for convergence. Sparsity constraints usually involve low-order norms (such as L1 or L0) applied to the inversion parameters.

In practice the sparsity constraints might be incorporated into the inversion problem for several reasons. Sometimes they can account for additional knowledge about the inversion parameters (for instance, in geophysics, the impulse response of the earth $\mathbf{X}$ is assumed to be sparse). In other cases they can be used to select the solutions with "concentrated" parameters over the "disperse" solutions (e.g. to eliminate unwanted reverberations in the estimates). In the case of data reconstruction, a sparsity constraint can be used to eliminate the aliasing artefacts in the estimations.

We know from inversion theory [Aster et al., 2013] that to incorporate an extra constraint into our algorithm all we need to is include such a constraint in to objective functional via $J^{new} = J^{old} + J$, were J^{old} is the original objective function, J^{new} is the new objective function and J is the objective function of the extra constraint. To solve our inversion problem, we need to be able to calculate the gradients of J^{new} with respect to the inversion variables and the scaling constant.

The implementation of a sparsity constraint can be achieved by using the following generalized function:

$$J(\mathbf{x}) = ||\mathbf{x}||_{S^n} = \left\{ \begin{array}{ccc} \sum_{ijk}(|x_{ijk}|^2 + \epsilon^2)^{-n+1} & if & n \neq 1 \\ \sum_{ijk} ln(|x_{ijk}|^2 + \epsilon^2) & if & n = 1 \end{array} \right\} \tag{C.0.1}$$

Table C.1: *Norms generated by variating the n and ϵ parameters in equation C.0.1*

norm	n	ϵ
L2	0	0
L1-L2	1/2	$\mathbb{R}$
L1	1/2	0
Cauchy	1	$\mathbb{R}$
German	2	$\mathbb{R}$

Where $\mathbf{x}$ is the model-space parameter array to sparsify and ϵ is a stabilization parameter related with the noise level in $\mathbf{x}$. By calibrating n and ϵ in equation C.0.1 we can obtain the set of norms shown in table C.2.

The calculation of the gradients might be cumbersome using the method described in appendix B given that the point-wise operations in equation C.0.1 prevent $J(\mathbf{x})$ to be expressed explicitly in a matrix-form. If this is the case, a component-by-component calculation of the gradient is required. Using equation B.1.2 and C.0.1 and assuming $n \neq 1$, the $\nabla_{\mathbf{x}} J$ gradient is then given by

$$\begin{aligned}
[\nabla_{\mathbf{x}} J]_{ijk} &= 2\frac{d}{dx^*_{ijk}} \sum_{lmo} (x_{lmo}x^*_{lmo} + \epsilon^2)^{-n+1} && \text{(C.0.2)}\\
&= 2\sum_{lmo} \frac{x_{lmo}}{(|x_{lmo}|^2 + \epsilon^2)^n} \frac{dx^*_{lmo}}{dx^*_{ijk}} && \text{(C.0.3)}\\
&= 2\sum_{lmo} \frac{x_{lmo}}{(|x_{lmo}|^2 + \epsilon^2)^n} \delta_{li}\delta_{mj}\delta_{ok} && \text{(C.0.4)}\\
&= 2\frac{x_{ijk}}{(|x_{ijk}|^2 + \epsilon^2)^n}. && \text{(C.0.5)}
\end{aligned}$$

Similarly, the $\partial_\alpha J(\mathbf{x}+\alpha\Delta\mathbf{x})$ term (which can useful in the calculation of the inversion step-lenght) is given by

$$\partial_\alpha J(\mathbf{x} + \alpha\Delta\mathbf{x}) = \sum_{ijk} \frac{(x_{ijk} + \alpha\Delta x_{ijk})\Delta x_{ijk}}{((x_{ijk} + \alpha\Delta x_{ijk}) + \epsilon^2)^n}. \tag{C.0.6}$$

Note that equations C.0.5 and C.0.6 remin valid even when $n = 1$ (Cauchy-type of constraint). A special case must be taken into account for the L1 norm ($n = 1/2$ and $\epsilon = 0$) in which a discontinuity is found at $x = 0$. Thought presented here for completeness, this discontinuity will limit the practical application of the gradient

Table C.2: *Gradients for common regularizations*

norm	n	ϵ	gradient
L2	0	0	$2x_{ijk}$
L1-L2	1/2	$\mathbb{R}$	$2\frac{x_{ijk}}{\sqrt{\lvert x_{ijk}\rvert^2+\epsilon^2}}$
L1	1/2	0	$2sign(x_{ijk})$
Cauchy	1	$\mathbb{R}$	$2\frac{x_{ijk}}{(\lvert x_{ijk}\rvert^2+\epsilon^2)}$
German	2	$\mathbb{R}$	$2\frac{x_{ijk}}{(\lvert x_{ijk}\rvert^2+\epsilon^2)^2}$

in equation C.0.5. More stable methodologies are recommended for such a case (see for example van den Berg and Friedlander [2011]).

Bibliography

Abma, R., and Kabir, N., 2006, 3D interpolation of irregular data with a POCS algorithm: Geophysics, **71**, no. 6, E91–E97.

Amundsen, L., 2001, Elimination of free-surface related multiples without need of the source wavelet: Geophysics, **66**, no. 1, 327–341.

Araujo, F. V., Weglein, A. B., Carvalho, P. M., and Stolt, R. H., 1994, Inverse scattering series for multiple attenuation: an example with surface and internal multiples: 64th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1039–1041.

Aster, R. C., Borchers, B., and Thurber, C. H., 2013, Parameter estimation and inverse problems: Elsevier.

Baardman, R. H., Verschuur, D. J., van Borselen, R. G., Frijlink, M. O., and Hegge, R. F., 2010, Estimation of primaries by sparse inversion using dual sensor data: 80th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 3468–3472.

Backus, M. M., 1959, Water reverberation - their nature and elimination: Geophysics, **24**, 233–261.

Berkhout, A. J., and Verschuur, D. J., 1994, Multiple technology, Part 2: Migration of multiple reflections: 64th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1497–1500.

Berkhout, A. J., and Verschuur, D. J., 1997, Estimation of multiple scattering by iterative inversion, part I: theoretical considerations: Geophysics, **62**, no. 5, 1586–1595.

Berkhout, A. J., and Verschuur, D. J., 2006, Focal transformation, an imaging concept for signal restoration and noise removal: Geophysics, **71**, no. 6, A55–A59.

Berkhout, A. J., and Verschuur, D. J., 2016, Enriched seismic imaging by using multiple scattering: The Leading Edge, **35**, no. 2, 128–133.

Berkhout, A. J., Ongkiehong, L., Volker, A. W. F., and Blacquiere, G., 2001, Comprehensive assessment of seismic acquisition geometries by focal beamspart i: Theoretical considerations: Geophysics, **66**, no. 3, 911–917.

Berkhout, A. J., Blacquiére, G., and Verschuur, D. J., 2009, The concept of double blending: Combining incoherent shooting with incoherent sensing: Geophysics, **74**, no. 4, A59–A62.

Berkhout, D. J. V. A. J., Davydenko, M., and Staal, S. R., 2015, Can we image everything with primaries?: 77th Ann. Internat. Mtg, EAGE, Expanded abstracts.

Berkhout, A. J., 1982, Seismic migration, imaging of acoustic energy by wave field extrapolation, A: theoretical aspects: Elsevier (second edition).

Berkhout, A. J., 1985, Seismic migration, imaging of acoustic energy by wave field extrapolation, A: theoretical aspects: Elsevier (third edition).

Berkhout, A. J., 2006, Seismic processing in the inverse data space: Geophysics, **71**, no. 4, A29–A33.

Berkhout, A. J., 2012, Combining full wavefield migration and full waveform inversion: Geophysics, **77**, no. 2, S43–S50.

Berryhill, J. R., and Kim, Y. C., 1986, Deep water peg legs and multiples emulation and suppression: Geophysics, **51**, no. 7, 2177–2184.

Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3d prestack imaging: Geophysics, , no. 63, 574–588.

Bisley, R., Moore, I., and Dragoset, H. W., 2005, Generalized 3D surface multiple prediction: PCT patent application publication, **WO 2005/103764**.

Bolondi, G., Loinger, E., and Rocca, F., 1982, Offset continuation of seismic sections: Geophys. Prosp., , no. 30, 813828.

Briggs, I. C., 1994, Machine contouring using minimum curvature: Geophysics, , no. 39, 3948.

Candes, E. J., and Recht, B., 2012, Exact matrix completion via convex optimization: Commun. ACM, , no. 55, 111119.

Candes, E., Demanet, L., Donoho, D., and Ying, L., 2006, Fast discrete curvelet transforms: Multiscale Model, **5**, no. 3, 861–899.

Chemingui, N., and Baumstein, A., 2000, Handling azimuth variations in multi-streamer marine surveys: 70th Ann. Internat. Mtg, SEG, Expanded abstracts.

Clapp, R. G., Biondi, B. L., Fomel, S. B., and Claerbout, J. F., 2007, Regularizing velocity estimation using geologic dip information: 68th Ann. Internat. mtg., Soc. Expl. Geophys., Expanded abstracts, 18511854.

Coates, R. T., and Weglein, A. B., 1996, Internal multiple attenuation using inverse scattering: Results from prestack 1 and 2-D acoustic and elastic synthetics: 66th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1522–1525.

Davydenko, M., and Verschuur, D. J., 2014, Full wavefield migration in three dimensions: 85rd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 3935–3940.

de Bruin, C. G., Wapenaar, C. P. A., and Berkhout, A. J., 1990, Angle-dependent reflectivity by means of prestack migration: Geophysics, **55**, no. 1, 1223–1234.

Deregowski, S. M., and Rocca, F., 1981, Geometrical optics and wave theory of constant offset sections in layered media: Geophys. Prosp., , no. 29, 374406.

Dragoset, W. H., and Jeričević, Z., 1998, Some remarks on surface multiple attenuation: Geophysics, **63**, no. 2, 772–789.

Dragoset, W. H., Moore, I., Yu, M., and Zhao, W., 2008, 3D general surface multiple prediction: An algorithm for all surveys: 78th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 2426–2430.

Dragoset, W. H., Verschuur, D. J., Moore, I., and Bisley, R., 2010, A perspective on 3D surface-related multiple elimination: Geophysics, **75**, no. 5, 75A245–75A261.

Duijndam, A. J. W., and Schonewille, M. A., 1999, Non-uniform fast fourier transform: Geophysics, **0**, no. 64, 539–551.

Fokkema, J. T. ., and van den Berg, P. M., 1993, Seismic applications of acoustic reciprocity: Elsevier Science Publishers B.V.

Gan, S., Chen, Y., Zhang, Y., and Jin, Z., 2015, De-aliased seismic data reconstruction using seislet-domain percentile thresholding: EAGE Conference and Exhibition, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Gandy, S., Recht, B., and Yamada, I., 2011, Tensor completion and low-n-rank tensor recovery via convex optimization: Inverse Problems, , no. 27, 025010.

Gensun, F., 1996, Whittakerkotelnikovshannon sampling theorem and aliasing error: Journal of Approximation Theory, **85**, no. 2, 115131.

Gisolf, A., and Verschuur, D. J., 2010, The principles of quantitative acoustical imaging: EAGE publications bv.

Guitton, A., and Verschuur, D. J., 2004, Adaptive subtraction of multiples using the l1-norm: Geophys. Prosp., **52**, 27–38.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: J. Can. Soc. Expl. Geophys., **22**, no. 1, 44–55.

Hargreaves, N., 2006, Surface multiple attenuation in shallow water and the construction of primaries from multiples: 76st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 2689–2693.

Hennenfent, G., Fenelon, L., and Herrmann, F. J., 2010, Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach: Geophysics, **75**, no. 6, WB203–WB210.

Herrmann, F. J., and Verschuur, D. J., 2008, Adaptive curvelet-domain primary-multiple separation: Geophysics, **73**, no. 3, A17–A21.

Herrmann, F. J., B oniger, U., and Verschuur, D. J., 2007, Nonlinear primary-multiple separation with directional curvelets: Geophys. J. Int., **170**, 781–799.

Herrmann, F. J., Wang, D., Hennenfent, G., and Moghaddam, P. P., 2008, Curvelet-based seismic data processing: A multiscale and nonlinear approach: Geophysics, **73**, no. 1, A1–A5.

Hjørungnes, A., 2011, Complex-valued matrix derivatives with applications in signal processing and communications: Cambridge University Press.

Hugonnet, P., and Canadas, G., 1995, Aliasing in the parabolic radon transform: 65th Ann. Internat. Mtg, Expanded abstracts, 1366–1369.

Hung, B., Yang, K., Zhou, J., and Xia, Q. L., 2010, Shallow water demultiple: ASEG Internat. Mtg, Australian SEG, Expanded abstracts.

Kabir, M. M. N., and Verschuur, D. J., 1995, Restoration of missing offsets by parabolic Radon transform: Geophys. Prosp., **43**, no. 3, 347–368.

Kinneging, N. K., Budejicky, V., Wapenaar, C. P. A., and Berkhout, A. J., 1989, Efficient 2D and 3D shot record redatuming: Geophys. Prosp., **37**, no. 5, 493–530.

Knutsson, H., and Westin, C., 1993, Normalized and differential convolution: Conference on Computer Vision and Pattern Recognition, IEEE Computer Society.

Kolda, T., and Bader, B., 2009, Tensor decompositions and applications: SIAM Journal on Scientific Computing, , no. 51, 455500.

Kontakis, A., and Verschuur, D. J., 2006, Deblending via sparsity-constrained inversion in the focal domain: 76th EAGE Conference and Exhibition, Eur. Ass. of Expl. Geophys., Expanded abstracts, 10.3997/2214–4609.20141456.

Kreimer, N., and Sacchi, M. D., 2011, A tensor higher-order singular value decomposition (hosvd) for pre-stack simultaneous noise-reduction and interpolation: 81st Ann. Internat. mtg., Soc. Expl. Geophys., Expanded abstracts, 18511854.

Kumar, R., Aravkin, A. Y., Recht, H. M. B., and Herrmann, F., 2013, Seismic data interpolation and denoising using svd-free low-rank matrix factorization: 75th EAGE Conference and Exhibition, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Kumar, A., Blacquiere, G., Pedersen, M. W., and Goertz, A., 2015, Marine survey design and analysis using all multiples: 77th EAGE Conference and Exhibition, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Kurin, E., Denisov, M. S., and Lokshtanov, D., 2006, A method for 3D surface related multiple prediction in case of coarse sampling: 68th Ann. Internat. Mtg., Eur. Ass. of Expl. Geophys., Expanded abstracts, B032.

Kutscha, H., and Verschuur, D. J., 2012, Data reconstruction via sparse double focal transformation: An overview: IEEE, **29**, no. 4, 53–60.

Kutscha, H., Verschuur, D. J., and Berkhout, A. J., 2010, High-resolution double focal transformation and its application to data reconstruction: Internat. Geoph. Conf. and Exp., Soc. Expl. Geophys., Expanded abstracts, 3589–3593.

Lee, J., Recht, B., Salakhutdinov, R., and Srebro, N., 2010, Practical large-scale optimization for max-norm regularization: Neural Information Processing Systems.

Lin, T. Y., and Herrmann, F. J., 2009, Unified compressive sensing framework for simultaneous acquisition with primary estimation: 79th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 3113–3117.

Lin, T. Y., and Herrmann, F. J., 2013, Robust estimation of primaries by sparse inversion via one-norm minimization: Geophysics, **78**, no. 3, R133R150.

Lin, T. Y., and Herrmann, F. J., 2016, Estimation of primaries by sparse inversion with scattering-based multiple predictions for data with large gaps: geophysics, **81**, no. 3, V183–V197.

Liu, F., Kabir, M. M. N., and Verschuur, D. J., 1995, Seismic processing using the wavelet and the radon transform: J. Seis. Expl., **4**, 375–390.

Lokshtanov, D., 1999, Multiple suppression by data-consistent deconvolution: The Leading Edge, **18**, no. 1, 115–119.

Lopez, G. A., and Verschuur, D. J., 2013, 3D primary estimation by sparse inversion using the focal domain parameterization: 83th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 4172–4177.

Lopez, G. A., and Verschuur, D. J., 2014, Closed-loop SRME - a new direction in surface multiple removal algorithms: 76th Ann. Internat. Mtg., Eur. Ass. of Expl. Geophys., Expanded abstracts, 1.

Lopez, G. A., and Verschuur, D. J., 2015a, 3d focal closed-loop srme for shallow water: 77th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 4418–4422.

——— 2015b, Closed-loop surface-related multiple elimination and its application to simultaneous data reconstruction: Geophysics, **80**, no. 6, V189–V199.

Lu, S., Whitmore, D., Valenciano, A., and Chemingui, N., 2014, Enhanced subsurface illumination from separated wavefield imaging: First Break, **32**, 87–92.

Martin, C. D., Shafer, R., and LaRue, B., 2013, An order-p tensor factorization with applications in imaging: SIAM Journal on Scientific Computing, , no. 35, A474A490.

Nekut, A. G., and Verschuur, D. J., 1998, Minimum energy adaptive subtraction in surface-related multiple elimination: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1507–1510.

Pedregal, P., 2013, Introductin to optimization: Springer.

Robinson, E. A., 1954, Predictive decomposition of time series with applications to seismic exploration: Ph.D. thesis, MIT.

Schreier, P. J., 2010, Statistical signal processing of complex-valued data: Cambridge University Press.

Soni, A. K., and Verschuur, D. J., 2015, Full wavefield migration to image salt-flanks using turning waves in walkaway vsp data: 77th Ann. Internat. Mtg, EAGE, Expanded abstracts.

Spagnolini, U., and Opreni, S., 1996, 3d shot continuation operator: 66th Ann. Internat. Mtg, Expanded abstracts.

Spitz, S., 1991, Seismic trace interpolation in the f-x domain: Geophysics, , no. 56, 785–794.

Staal, S. R., and Verschuur, D. J., 2013, Joint migration inversion, imaging including all multiples with automatic velocity update: 75th Ann. Internat. Mtg, EAGE, Expanded abstracts.

Sun, W., and Wang, P., 2010, Model-based water-layer demultiple: ASEG Internat. Mtg, Australian SEG, Expanded abstracts.

Tarantola, A., 1987, Inverse problem theory, methods for data fitting and model parameter estimation: Elsevier Science Publ. Co., Inc.

Thorson, J. R., and Claerbout, J. F., 1985, Velocity-stack and slant-stack inversion: Geophysics, **50**, no. 12, 2727–2741.

Trad, D. O., Ulrych, T., and Sacchi, M., 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386–399.

Trad, D., 2000, Implementations and applications of the high resolution radon transform: Ph.D. thesis, University of British Columbia.

van Borselen, R. G., Fokkema, J. T., and van den Berg, P. M., 1994, A modified conjugate gradient method for minimizing the energy in the removal of free-surface-related wavefields: J. Seis. Expl., **4**, 351–364.

van Dedem, E. J., and Verschuur, D. J., 1998, 3D surface-related multiple elimination and interpolation: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1321–1324.

van den Berg, E., and Friedlander, M. P., 2008, Probing the pareto frontier for basis pursuit solutions: SIAM J. Sci. Comput., **31**, no. 2, 890–912.

van den Berg, E., and Friedlander, M. P., 2011, Sparse optimization with least-squares constraints: SIAM Journal on Optimization, **21**, no. 1201.

van Groenestijn, G. J. A., and Verschuur, D. J., 2008, Towards a new approach for primary estimation: 78th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 2487–2491.

van Groenestijn, G. J. A., and Verschuur, D. J., 2009a, Estimating primaries by sparse inversion and application to near-offset data reconstruction: Geophysics, **74**, A23–A28.

——— 2009b, Estimation of primaries and near offsets by sparse inversion: Marine data applications: Geophysics, **74**, R119–R128.

van Groenestijn, G. J. A., and Verschuur, D. J., 2010, Estimation of primaries by sparse inversion from passive seismic data: Geophysics, **75**, no. 4, SA61–SA69.

van Groenestijn, G. J. A., and Verschuur, D. J., 2011, Using surface multiples to estimate primaries by sparse inversion from blended data: Geoph. Prosp., **59**, no. 1, 10–23.

van Groenestijn, G. J. A., 2010, Estimation of primaries and multiples by sparse inversion: Ph.D. thesis, Delft University of Technology.

Vasmel, M., Robertsson, J. O. A., and Amundsen, L., 2016, A model-independent finite-difference method for removal of free-surface generated multiples: Geophysics, **81**, no. 2, T79T90.

Verschuur, D. J., and Berkhout, A. J., 1997, Estimation of multiple scattering by iterative inversion, part II: practical aspects and examples: Geophysics, **62**, no. 5, 1596–1611.

Verschuur, D. J., and Berkhout, A. J., 2005, Removal of internal multiples with the common-focus-point (CFP) approach: Part 2 - Application strategies and data examples: Geophysics, **70**, no. 3, V61–V72.

Verschuur, D. J., and Berkhout, A. J., 2006, Multiple removal in the inverse data space: 68th Ann. Internat. Mtg., Eur. Ass. of Geosc. and Eng., Expanded abstracts, D040.

Verschuur, D. J., and Berkhout, A. J., 2015, From removing to using multiples in closed-loop imaging: The Leading Edge, **34**, no. 7, 744–759.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 9, 1166–1177.

Verschuur, D. J., 1991, Surface-related multiple elimination: an inversion approach: Ph.D. thesis, Delft University of Technology.

Verschuur, D. J., 2006, Seismic multiple removal techniques - past, present and future: EAGE Publications BV.

Verschuur, D. J., 2013, Estimation of primaries by sparse inversion incuding the ghost: 79th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 4094–4100.

Vrolijk, J., and Verschuur, D. J., 2015, Integrated deghosting and multiple removal: field data experience: 77th Ann. Internat. Mtg, SEG Expanded abstracts.

Vrolijk, J. W., and Verschuur, D. J., 2016, Integrated receiver deghosting and closed-loop surface multiple elimination: DELPHI Volume XXVII, Chapter 4, pages 53–66.

Wapenaar, C. P. A., and Berkhout, A. J., 1989, Elastic wave field extrapolation: redatuming of single- and multi-component seismic data: Elsevier Science Publ. Co., Inc.

Wapenaar, C. P. A., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E., and Snieder, R., 2014, Marchenko imaging: Geophysics, **79**, no. 3, WA39–WA57.

Weglein, A. B., Gasparotto, F. A., Carvalho, P. M., and Stolt, R. H., 1997, An inverse scattering series method for attenuating multiples in seismic reflection data: Geophysics, **62**, 1975–1989.

Wiggins, J. W., 1988, Attenuation of complex water-bottom multiples by wave-equation-based prediction and subtraction: Geophysics, **53**, no. 12, 1527–1539.

Ypma, F. H. C., and Verschuur, D. J., 2013, Estimating primaries by sparse inversion, a generalized approach: Geophys. Prosp., **61**, 94–108.

Zwartjes, P. M., and Gisolf, A., 2006, Fourier reconstruction of marine-streamer data in four spatial coordinates: Geophysics, **71**, no. 6, V171–V186.

CV

Gabriel Lopez was born in Bogota-Colombia in 1987, he received his diploma in Physics from the University of Los Andes in early 2010. Before finishing his BSc he started his MSc studies in Physics at the same university, from which he graduated one year later in 2011. He worked in the university as assistant teacher, graduate assistant and research associate between 2009 and 2011. His MSc thesis entitled "Source scaling of intermediate-depth earthquakes in the Bucaramanga Nest" was awarded and fully financed by the universitys faculty of science. The same research was also awarded by the AGU and the university of Los Andes faculty of science for presentation in the 2010 AGU meeting. In late 2011 Gabriel joined Delft University of Technology and the DELPHI consortium as a Ph.D. candidate in the department of Imaging Physics, under the supervision of dr. ir. D. J. Verschuur. After 5 years of research within DELPHI he received his Ph.D. diploma for his work on surface multiple estimation.

List of publications

Journal/conference publications

- G. A. Lopez and D. J. Verschuur, 2016, Closed-Loop SRME and its application to simultaneous data reconstruction, Geophysics, SEG
- G. A. Lopez and D. J. Verschuur, 2015, 3D Closed-Loop SRME for shallow water, SEG, Expanded Abstracts
- G. A. Lopez and D. J. Verschuur, 2015, Closed-Loop SRME with data reconstruction, EAGE, Expanded Abstracts
- G. A. Lopez and D. J. Verschuur, 2014, Closed-Loops SRME and its application to data reconstruction, SEG, Expanded Abstracts
- G. A. Lopez and D. J. Verschuur, 2014, Closed Loop SRME New direction in surface multiple attenuation, EAGE, Expanded Abstracts
- G. A. Lopez and D. J. Verschuur, 2013, Estimation of Primaries By Sparse Inversion Parameterization Via the Focal Domain, EAGE, Expanded Abstracts
- G. A. Lopez and D. J. Verschuur, 2013, 3D primary estimation by sparse inversion using the focal domain parameterization, SEG, Expanded Abstracts
- G. A. Lopez and D. J. Verschuur, 2012, SRME and estimation of primaries by sparse inversion: a hybrid approach, SEG, Expanded Abstracts
- G. A. Prieto , G. C. Beroza, S. A. Barrett, G. A. Lopez, M. Florez, 2012, Earthquake nests as natural laboratories for the study of intermediate-depth earthquake mechanics, Tectonophysics

Book chapters

- G. A. Lopez and D. J. Verschuur, 2015, 3D Focal Closed-Loop SRME, DELPHI book, M&I project, Den Haag
- G. A. Lopez and D. J. Verschuur, 2015, Focal Closed-Loop SRME, DELPHI book, M&I project, Den Haag
- G. A. Lopez and D. J. Verschuur, 2015, Closed-Loop SRME, DELPHI book, M&I project, Den Haag,
- G. A. Lopez and D. J. Verschuur, 2014, Closed-Loop SRME for coarsely sampled data, DELPHI book, M&I, Den Haag
- G. A. Lopez and D. J. Verschuur, 2014, Closed-Loop SRME, DELPHI book, M&I project, Den Haag
- G. A. Lopez and D. J. Verschuur, 2013, Towards 3D EPSI, DELPHI book, M&I project, Den Haag
- G. A. Lopez and D. J. Verschuur, 2013, The SRME+ method, Multiple removal and structural imaging project, M&I project, Den Haag
- D. J. Verschuur and G. A. Lopez, 2012, Estimation of primaries by sparse inversion: shallow water applications andoutlook for 3D, M&I project, Den Haag

Selected posters and presentations

- G. A. Prieto, M. Florez, G. A. Lopez, S. Barrett and G. C. Beroza, 2012, Earthquake source scaling, stress drops and seismic efficiency of intermediate-depth earthquakes, EGU, Expanded Abstracts
- G. A. Prieto, G. A. Lopez, S. Barrett and G. C. Beroza, 2011, Earthquake source scaling, stress drops and radiated seismic energies of intermediate depth earthquakes, AGU, Expanded Abstracts
- G. A. Lopez and G. A. Prieto, 2010, Source scaling of intermediate-depth earthquakes in the Bucaramanga Nest, AGU, Expanded Abstracts

Acknowledgments

I've learned a lot during my Ph.D. both as a person and as a scientist. But offcourse, much of this learning I owe to particular individuals, to whom I'm particularly grateful.

My right-hand, guidance and apparently source of unlimited knowledge is, and has always been, my supervisor dr. Eric Verschuur. I want to thank him specially for believing in me and for being always the friendly hard-working supervisor that any student eagers to have. My admiration extends to Prof. Berkhout, the leader of our team, who has kept his name as one of the most influent scientists of our field. Thanks also to dr. Gerrit Blacquire and dr. Dries Gisolf for all the interesting discussions during the meetings.

I visited Shell Rijswick for an internship in early 2015. Special thanks to dr. Koos de Vos and dr. Fons ten Kroode for a very insightful stay at Shell.

During late 2015 I had the privilege to visit the Leiden research center of PGS, there I met an amazing an enthusiastic team of experts from whom I learned a lot. Thanks to dr. Gert-Jan van Groenestijn and dr. Roald van Borselen for directing my research at PGS, and offcourse thanks also to Rolf, Martijn, Christina, Dorit and Sixue for an amazing (an also fun!) experience in PGS. You guys thought me that eating canned soup every day is much better than having to cook. I will keep a can to remember you by. Please remember me by my "slow-goals".

I spent my day-to-day life among my dear TUDelft mates. Special thanks to Mikhail (for the sleep-overs), Hannes (for being my "amigo"), Khalid (for teaching me some Dutch), Xander (for all the very insightful discussions), Apostolos (for coping with me every day), Matteo (for the evening "expressos"), Abdulrahman (for teaching me about other cultures), Siddarth (for his nice life-philosophy) and Amarjeet (he

knows why!). And off-course to all the ex-Delphi mates Peter, Araz, Alok, Tomo and Panos from whom I've learned a lot. Thanks also to all the new people: Aayush, Runhai, Aparajita, Bouchaib, Hussain, Shan, Shogo, Ranjani, Prabu and Ewoud for all the new energy they bring to the consortium. Thanks also to Farid, Ana, Jan-Willem, Jewoo, Mikhail B., Niels and Yimin. The list wouldn't be complete without our technical support staff: Edo and Henry. Thank you all and please forgive me if I have unintentionally left some one aside.

Ninguno de mis logros habría sido posible sin el apoyo constante de mi amada familia, ellos han sido siempre el componente más importante de mi vida. Mis padres Luis y Julia me hicieron quien soy hoy, con su amor, instrucción y apoyo me pude desarrollar completamente. Aún hoy, ellos continúan siendo mi principal inspiración y motivación para continuar adelante en mi vida. Sólo quiero que estén orgullosos de mí. Ellos me dieron lo mejor de sí mismos, me dieron alas para volar alto. No tengo más sino recuerdos felices de mi niñez y juventud. Los amo profundamente.

Mis hermanos David, Juliana y Angélica caminaron a mi lado el camino de la vida, del cual atesoro los mejores recuerdos. Hemos compartido mucha aventuras juntos, en muchos países diferentes. Ellos son mis consejeros y mi apoyo permanente. Tienen mi amor incondicional. Agradezco especialmente a mi hermana Angélica, quien ha sido escogida por la naturaleza como mi leal compañia durante los últimos cinco años. Estoy bendecido de tenerla a mi lado.

Mi tía Sara y mi abuela Julia tuvieron un rol importante en mi vida. Tengo un lugar especial en mi corazón para ellas. Gracias a Carlos, Armando, Diógenes, Nancy, Gonzalo, Gonzalo D., Jacqueline, Catalina, Leonor e Israel. Siempre cargaré un poco de ellos en mí.

Special thanks to Remko (for his work in this document), Brett, Carlos P., Maarten, Andres, Paola, Arjen, Coby, Wietze, Henrieke, Artur and Sandra (they know why).

Now, time comes for me to continue with the next chapter of my life. Who knows what mysteries await in the unassailable future? Which Feynmann-path will provide a formal solution to the universal equation of evolution $\delta S = 0$?

A mystery. For even the wisest cannot tell.

Gabriel López A

Front and Back Cover Plots

The front and back cover plots in this thesis are some examples of some of the fractal art-work I developed as a hobby during my Ph.D. years in Delft University.

The fractals shown in the front and back covers depict the convergence set over the complex plane of a couple of trigonometric generating functions, which become chaotic after a large amount of auto-compositions. The outline of the figures is formally known as the Julia set and its compliment is known as the Fatou set. I generated these fractals myself after some research on their functional expressions.

These fractals represent complex and chaotic dynamics normally obtained in the evolution of dynamical systems. These solutions typically arise in interacting regimes in which the equations of motion of the system become non-linear. Systems with these properties exist everywhere in Nature, from the triggering mechanism of earthquakes till the patterns of climate dynamics.