# IDEA League

# Theoretical and practical limitations of IME-algorithms

**Francois Lux**

September 9, 2015

# Theoretical and practical limitations of IME-algorithms

MASTER OF SCIENCE THESIS

for the degree of Master of Science in Applied Geophysics at
Delft University of Technology
by

Francois Lux

September 9, 2015

Department of Geoscience & Engineering · Delft University of Technology

**Delft University of Technology**

IDEA LEAGUE
JOINT MASTER'S IN APPLIED GEOPHYSICS

Delft University of Technology, The Netherlands
ETH Zürich, Switzerland
RWTH Aachen, Germany

Dated: *September 9, 2015*

Supervisor(s):

_____

Dr. ir. D.J. Verschuur

_____

Prof. Dr. ir. C.P.A. Wapenaar

Committee Members:

_____

Dr. ir. D.J. Verschuur

_____

Prof. Dr. ir. C.P.A. Wapenaar

_____

Prof. Dr. J. van der Kruk

Dedicated to
Curiosity and Cheerfulness

# **Abstract**

Several internal multiple prediction methods have been proposed in the past and are still actively being researched. Of those, the Jakubowicz method is currently the most attractive and applied one, as it only uses surface data in a convolution and correlation process, and is therefore computationally relatively cheap. This allows its commercial application on 3D data. But due to several assumptions in the prediction process, it has to rely on adaptive subtraction to correct for potential errors. These errors are differentiated and investigated in more detail. Some of them derive from theoretical shortcomings, such as the incorrect implementation of transmission operators, or when neglecting the structural complexity of a 3-dimensional earth in case of 2D or 1D application. The others are introduced because of the unknown properties and simplifying assumptions that have to be made when working with field data. The most important issues are the unknown source characteristics, noise, and the required spatial sampling. These are also of significance for any other method based on the convolution of wavefields, such as Marchenko-imaging. Any advancement that is made regarding these problems, potentially leads to improved results in their application on field data.

# Acknowledgements

At last...
Nineteen years of education culminate in this scientific statement of mental proficiency that is called Master thesis. This seems like an appropriate time to look back and thank the people that enabled me to get to this memorable point in life when one turns from being the human equivalent of a money-, teachings- and fact-absorbing black hole to a more "exothermic", self-reliant and productive member of society.

Naturally I want to thank my parents for raising me lovingly, encouraging and enabling me to pursue education and life-experience away from "home", while being my financial and emotional safety-net. I also want to mention and thank my high school geography teacher, Mr. Gangloff, who taught me that complex problems never have an easy solution, and who sparked my interest for global issues and ultimately for the energy challenge.

If I would have to describe my journey of the past six years using a nautical metaphor, where my parents sponsored the ship and made sure that I had enough and a balanced variety of food on board, so that I would not suffer from scurvy, and Mr. Gangloff calibrated my compass guiding me towards my scientific profession of interest, I would have to thank my brother Christoph for encouraging me to go out of my comfort zone and providing the initial breeze setting this ship in motion with the following quote on a card for my 18th birthday:

*"A ship in harbor is safe, but that is not what ships are built for"*

But enough with the metaphors as I don't want anyone to get seasick!

With regard to this thesis, I first and foremost have to thank DMT Petrologic, and in particular Gerd Rybarczyk, for providing me with the opportunity to work on IME and to

Next I want to thank German Hoecht, who's experience of writing computer-algorithms for optimized processing of large-scale datasets helped me to appreciate and get an insight into all the small difficulties that one encounters when transforming a theory from paper to bytes. Additionally he assisted me with the 3D-IME tests.

For being kind enough and allowing me to use their field dataset in order to produce suitable visual examples used in chapter 5, I want to acknowledge Vermilion Energy Inc. It really helped me to get the "practical" into "Theoretical and practical limitations...".

Probably more people, than I would initially think of, have in some way or another contributed to this final thesis, so if you think you are one of them or if you know someone who might have, then please be assured that my gratitude extends to you/them.
Ahoy!

Delft University of Technology                                                      Francois Lux
September 9, 2015

# Table of Contents

# List of Figures

# Acronyms

**DUT** Delft University of Technology

**IME** Internal multiple elimination

**CFP** Common focus point

**SRME** Surface related multiple elimination

**EPSI** Estimation of Primaries by Sparse Inversion

**NMO** Normal move out

**LTI** Linear time-invariant

**CMP** Common mid-point

**SRMM** Surface related multiple modeling

**IMM** Interbed multiple modeling

**ISS** Inverse scattering (sub-)series

**S/N-ratio** Signal to noise ratio

**MME** Marchenko multiple elimination

**DMO** Dip move out

# Chapter 1

# Introduction

*"Energy is the single most important challenge facing humanity today."*

– Richard Smalley, *2004*

Energy plays a significant role in every major problem facing our ever increasing world population in the 21st century. The global energy demand is believed to be increasing by 30% in 2035 as of 2011. While fossil fuels relative share will decrease, its total demand will still increase [IEA, 2013].

Oil and Gas will, despite their bad CO2 footprint and prominent role in climate change, remain significant contributors in the years ahead, especially for transportation. In order to meet this demand in an era where "easy oil is over" the Oil & Gas industry faces the challenge of exploring for new oil and gas fields in more and more difficult settings such as deepwater subsalt- or Arctic-basins. This is only possible by developing improved methods and increasing technical proficiency, as was demonstrated with the recent shale oil/gas boom.

The most prominent method for the Oil & Gas Exploration, but also for exploration of geothermal resevoirs or potential atomic waste disposal locations, is the seismic method.

A source, most commonly placed near the surface, emits a seismic wavefield into the earth, propagating downwards till it reaches an inhomogeneity in the subsurface where part of its energy is reflected upwards again. The upcoming wavefield is recorded and sampled at appropriate intervals at the surface using receivers (hydrophones recording the pressure for the marine case, geophones recording ground velocity or acceleration for the land case). This is repeated for different source and receiver-positions so that a large area of the subsurface is illuminated.

This seismic dataset is then digitally processed in order to generate a structural map of the subsurface inhomogeneities. Most commonly only so called "primary reflections" are wanted for processing. Any other recorded signal is considered as noise and a considerable amount of effort is made in order to remove it.

Primary reflections are seismic waves that traveled into the earth and only bounced upwards once in the subsurface before being recorded. In the same sense, seismic waves that bounced downwards at least once before being recorded at the surface are called multiple reflections. These multiples cause problems for seismic processing and interpretation. Multiples, having traveled in a shallower part of the subsurface, compared to primaries with the same arrival time, often have lower normal-move-out(NMO)-velocities, which can lead to confusion during velocity-analysis. Also, most currently industry-used seismic imaging-algorithms operate under the assumption that the data only contains primaries.

For interpretation, not only structural information but also amplitude-related information is of interest. In the worst case structures caused by multiple reflection energy could be mistaken for valid prospects and drilled. But even if multiples are correctly identified they may interfere with primaries in the target-area and mask the amplitude-information.

When multiples and primaries can not be separated using seismic data transforms, multiple removal methods based on the wave-equation are needed. The most popular ones in the industry are "Surface related multiple elimination"(SRME) and "Internal multiple elimination"(IME), dependent on the multiple-type. These methods are based on adaptive subtraction of predicted multiple-wavefields generated by convolution and correlation of already recorded wavefields in the dataset.

Although a direct subtraction of the predicted multiple-wavefield would be desirable, in order to not involuntarily remove primary amplitude information, adaptive subtraction is required as the prediction is not kinetically and dynamically correct. These amplitude-, arrival-time-, and wavelet errors are the subject of this thesis and investigated with regard to the Jakubowicz-implementation of IME. The Jakubowicz approach for IME is particularly attractive as it is completely data-driven, and does not require wavefield-redatuming. The latter hinders the usage of common focus point(CFP)-IME for larger 3D-datasets in the light of current acquisition and hardware restrictions.

However most of the discussed errors are also relevant in SRME- and CFP-IME-implementations, and in fact a problem for any method based on wavefield-convolution, like Marchenko imaging. The goal of this thesis is to halt for a moment and take the time to consider all the different error-sources and their contribution to the total prediction-error, which we normally just disregard, hoping that the adaptive subtraction will take care of it.

First, I will summarize shortly the basic knowledge and characteristics of multiples and data-manipulation tools needed, in order to understand how their properties may contribute to the final prediction. Next, I will make a broad overview of different multiple-elimination algorithms and highlight their differences, before I explain the theory of SRME and IME in more detail, showing how they evolved from each other. In the main part of the thesis I will then separate and analyze the different error-sources with regards to whether they are due to simplifying assumptions in the theory, or due to the nature and many short-comings of real seismic data-sets when applying this theory in practice.

While theoretical errors are mostly caused due to not accounting for the 3D character of wavefields, in practice, noise, the unknown source-wavelet and the irregular and limited spatial sampling are the biggest concerns. Therefore, careful seismic survey planning and continuous research in the areas of noise-attenuation, wavelet-estimation and accurate trace-interpolation, are essential to further improve the result of the multiple prediction.

# Part I

# First Part

# 2

# Theoretical Basis

## 2-1 Multiples

### 2-1-1 What are multiples?

Seismic surveys exploit the physical phenomena that part of the energy of sound waves is reflected when a change in seismic impedance (the product of density and seismic velocity) is encountered. So, after emitting a down-going wavefield into the subsurface we can record an up-going wavefield for every impedance-contrast illuminated and resolved by our band-limited source-signal.

However, the laws of physics, rarely, if ever, concern and limit them-selves to what would be beneficial for us. This means that the same effect happening on the way down also happens on the way up, in the shallower parts of the subsurface. These wavefields, which experienced at least one down-ward reflection, before finding their way to the recording receivers, are defined as multiples. They usually are considered an undesired part of the total recorded wavefield.

### 2-1-2 How are multiples categorized?

For the purpose of this thesis, we only consider multiples with a period long enough so they can be recognized as separate events. It is also enough to categorize multiples by whether their shallowest down-ward reflection happened at the surface or not. If aforementioned condition is met they are called surface-related multiples (yellow and orange in figure 2-1). Multiples falling into the second category will from here on be referred to as internal multiples (green and blue in figure 2-2). This distinction is made because removing surface-related multiples has the distinct advantage of having recorded the wavefield on the multiple generating interface. For a more detailed discussion and differentiated categorization the reader is referred to [Verschuur, 2006a] and [O'Doherty and Anstey, 1971].
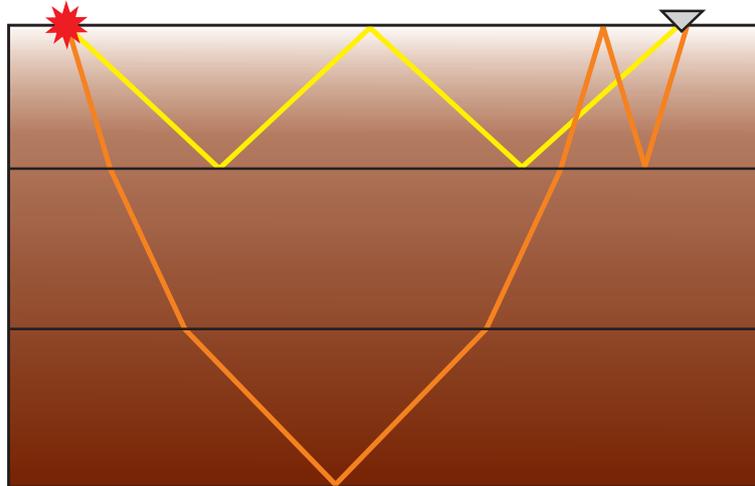
**Figure 2-1:** Surface-related multiples.
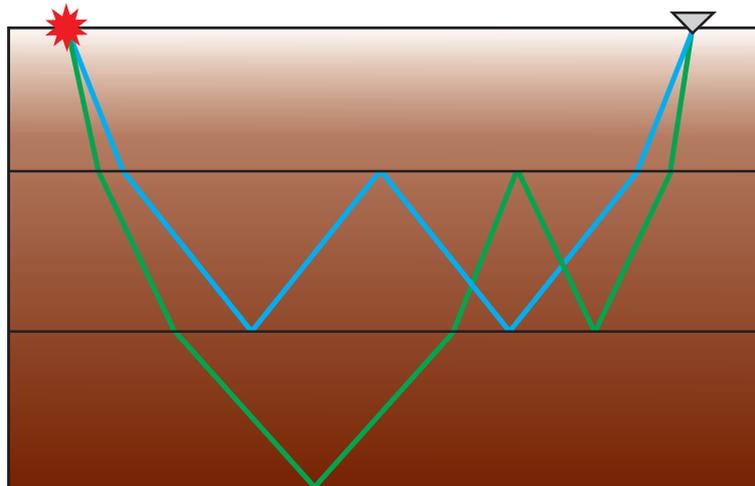


**Figure 2-2:** Internal multiples.
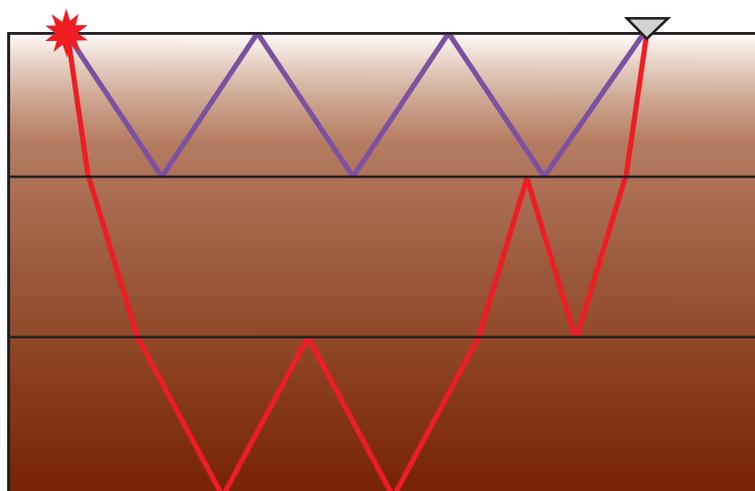


**Figure 2-3:** Higher-order multiples.

If the amplitude of primary reflections can be described with $R$ ($R$ being a representative value for the reflection-strength of the subsurface-boundaries), the amplitude of internal multiples is at least $R^3$ ($R^5, R^7$ and so on for higher order multiples [red and purple in figure 2-3]). Surface-related multiples are once again a special case as their multiple-generating boundary, the free surface, has a reflection coefficient of $-1$ [Chapman, 2004]. This means that due to their inherently low amplitude, multiples are mostly a problem in the presence of strong impedance-contrasts. Common examples for this are the free-surface, and the top of a salt-layer.

Perhaps counter-intuitively, as we normally affiliate a multiple to its down-ward reflection interface, the reflection coefficient of the up-ward reflecting interface is of higher importance, as it contributes twice to the final amplitude (see figure 2-2).

### 2-1-3 What are the characteristics of multiples?

In order to design effective tools for multiple-removal we first have to know what differentiates them from primaries and what other typical attributes they have.

In a simplified way multiples are primary reflections that made an additional round-trip through parts of the subsurface. They still have resemblance with the original reflection structure and therefore appear to be periodic repetitions of them. The subsurface structure repeatedly leaves an imprint on the multiple with each additional round-trip. This leads to an increasing dip and magnified lateral amplitude variation for higher order multiples.

Having traveled in the shallower part of the subsurface compared to primaries, arriving at a similar recording-time, multiples may exhibit a different move-out behavior [Ryu, 1982], often having a lower NMO-velocity. For the same reason, they may also have different, and therefore conflicting dip with primaries. As any signals arriving at the same recording time do, multiples may de- or con-structively interfere with primaries or other multiples.

### 2-1-4 How do we get rid of multiples?

The above analysis leads to two different categories of multiple removal methods based on the characteristics they try to exploit.

- Methods exploiting the different spatial behavior of multiples and primaries

  As multiples experienced a different path through the subsurface than primaries, resulting in a different dip and move-out, it is possible to separate them after resorting the data or transforming it into a different domain. Once separated they can be removed or at least weakened with the help of filter-methods. User interpretation for this step is needed though. Examples for such methods are, filtering after parabolic Radon-transformation [Hampson, 1986], or simply the stacking of data after move-out correction in the CMP-offset-domain. However, these methods have the fundamental weakness that they do not work when their underlying assumption, a larger dip or

move-out difference, is not met. These methods will not be further discussed in this thesis.

- Methods taking advantage of the predictability and periodicity of multiples

  These methods consist of a two step procedure: the prediction of a multiple wavefield and then subsequently its subtraction from the total wavefield. In its simplest form, the primaries are in some manner scaled and time-shifted in order to create the multiple prediction. Their underlying concept is always wavefield extrapolation. Different versions of these methods exist; from fully data-driven adaptations only using already recorded wavefields, to algorithms using a velocity model in order to calculate the Green's function for wavefield extrapolation, to inversion implementations. Their reliance on the physical law governing wave-propagation, the wave-equation, allows them to work even in environments where methods of the first category fail.

For over 60 years now [Robinson, 1954] the seismic-industry has been studying multiples, their proclaimed "enemy". Long enough to realize that with enough knowledge and proper tools it is possible to turn multiples into friends. Multiples, having traveled even twice through some parts of the subsurface, carry important information about reflection coefficients of shallow interfaces and can increase resolution of the final image. Several results have already shown, for how surface-related multiples contribute to the imaging of primaries [Berkhout and Verschuur, 1994][Guitton et al., 2002][Lu et al., 2013].

## 2-2   Fourier Transformation

Data-transforms in seismic processing are a common tool allowing the separation of events which are normally not isolated in a seismic record. Another advantage is that certain operations can be carried out more efficiently after transformation to a different domain.

One of these data-transforms is the Fourier-transformation. It is not only attractive for dip-filtering but also for processing seismic data one frequency-component at a time and simplifying complex mathematical operations like convolution and correlation. Its underlying assumption is that every periodic time-series can be represented using phase-shifted cosine and sine-waves of different frequency and amplitude. So a seismic trace, normally defined by real-valued time-samples, becomes, after temporal fourier-transformation, defined by frequencies of different phase and amplitude [Dudgeon and Mersereau, 1984].

These trigonometric functions are commonly represented in their complex exponential form using Euler's formula,

$$e^{ix} = \cos x + i \sin x \quad . \tag{2-1}$$

Note that any sample after Fourier transformation is now complex-valued.

The following notation is used throughout this thesis:

Functions in the space-time domain (x,y,z,t) are written using the lower case symbol. After transformation into the space-frequency domain (x,y,z,$\omega$) the corresponding upper case symbol is used.

Using the angular frequency,

$$\omega = 2\pi f \geq 0 \quad , \tag{2-2}$$

a time series h(x,y,z,t) can be transformed into the space-frequency-domain using the following equation:

$$H(x, y, z, \omega) = \int_{-\infty}^{\infty} h(x, y, z, t)e^{-j\omega t}dt \quad . \tag{2-3}$$

The inverse of the temporal Fourier transformation is given by:

$$h(x, y, z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(x, y, z, \omega)e^{j\omega t}d\omega \quad . \tag{2-4}$$

Seismic wavefields are not only sampled in time though. Using receivers with a set distance between each other, the recorded wavefield is also spatially sampled at the surface. Using the spatial Fourier-transformation, a function can also be transformed into the wavenumber domain, where $x$ and $y$ turn into the spatial frequencies $k_x$ and $k_y$. Functions in the wavenumber-frequency domain ($k_x$,$k_y$,z,$\omega$) are denoted by a tilde over the corresponding upper case symbol. So, a time series H(x,y,z,$\omega$) can be transformed into the the wavenumber-frequency domain using,

$$\tilde{H}(k_x, k_y, z, \omega) = \iint_{-\infty}^{\infty} H(x, y, z, \omega)e^{j(k_x x + k_y y)}dxdy \quad , \tag{2-5}$$

with the inverse of the double spatial Fourier transformation given by,

$$H(x, y, z, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \tilde{H}(k_x, k_y, z, \omega)e^{-j(k_x x + k_y y)}dk_x dk_y \quad . \tag{2-6}$$

Instead of doing it in two separate steps it is also possible to transform a time series h(x,y,z,t) directly into $\tilde{H}(k_x, k_y, z, \omega)$ using,

$$\tilde{H}(k_x, k_y, z, \omega) = \iiint_{-\infty}^{\infty} h(x, y, z, t)e^{-j(\omega t - k_x x - k_y y)}dxdydt \quad . \tag{2-7}$$

The inverse of the triple Fourier transformation then reads:

$$h(x, y, z, t) = \frac{1}{8\pi^3} \iiint_{-\infty}^{\infty} \tilde{H}(k_x, k_y, z, \omega)e^{j(\omega t - k_x x - k_y y)}dk_x dk_y d\omega \quad . \tag{2-8}$$

Using the following representation of a complex seismic trace,

$$H(x,y,z,\omega) = \int_{-\infty}^{\infty} h(x,y,z,t)\cos(\omega t)dt - i\int_{-\infty}^{\infty} h(t)\sin(\omega t)dt \equiv R(\omega) - iI(\omega) \quad , \quad (2\text{-}9)$$

the instantaneous amplitude or "reflection strength" is calculated using,

$$A(\omega) = \sqrt{R(\omega)^2 + I(\omega)^2} \quad , \tag{2-10}$$

and the instantaneous phase is given by

$$\Phi(\omega) = \tan^{-1}\frac{I(\omega)}{R(\omega)} \quad . \tag{2-11}$$

It becomes apparent that the amplitude is independent of the algebraic sign of the imaginary part. See [Taner, Koehler and Sheriff, 1979] for further analysis.

For the purpose of this thesis, understanding the relationship between amplitude spectrum and wavelet is of importance. Generally speaking, the broader the amplitude spectrum of a seismic trace is, the more confined the energy of the wavelet in time and, therefore, the higher its resolving power [Koefoed, 1981].

This is illustrated in figure 2-4, showing a seismic wavelet with a broad frequency-spectrum of 12 to 126 Hertz. Comparing the width of this wavelet with the width of the wavelet in figure 2-5, only having a narrow frequency-spectrum of 18 to 52 Hertz, it becomes apparent that the narrower the frequency-spectrum becomes the more the energy of a wavelet is smeared over several time-samples, if we consider its main peak. It should also be generally noted that zero-phase wavelets have the highest resolving capabilities for interpretation purposes, although the time length of the wavelet is phase-independent.
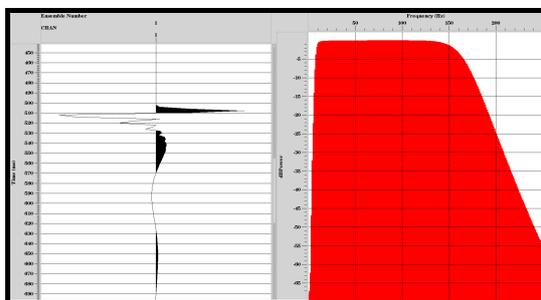


**Figure 2-4:** Wavelet with a broad frequency-bandwidth.
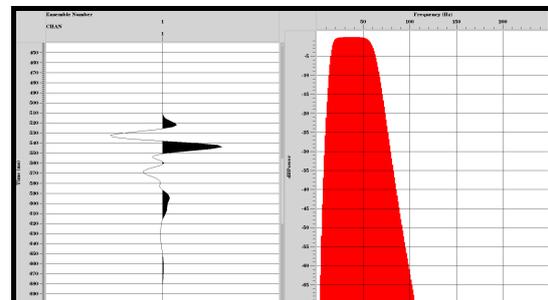


**Figure 2-5:** Wavelet with a narrow frequency-bandwidth.

## 2-3   Convolution and Correlation

Convolution is a mathematical operation useful when dealing with linear time-invariant (LTI) systems. The earth is such a system when interacting with sound waves. The band-limited signal emitted by the seismic source gets convolved with the earth's impulse response, and then, at the surface, convolved with the impulse response of the receiver and seismic recording device, resulting in the total seismic record. This is called the convolutional earth model:

$$X(t) = S(t) * G(t) * R(t) * A(t) \qquad (2\text{-}12)$$

where

$X(t) =$ seismic record
$S(t) =$ source signal
$G(t) =$ impulse response of the earth
$R(t) =$ impulse response of the receiver
$A(t) =$ impulse response of the recording-device

Technically speaking convolution means that a function is time-reversed and shifted along another signal by $\tau$, each time calculating the inner product. In a more descriptive manner this means that one function is "blended" or "folded" with another function, similar to leaving an imprint. The inverse operation, such as the removal of the source-signal from the seismic record, is known as deconvolution.

It is of importance to note that wavefield-extrapolation can be formulated as spatial convolution, and is used as such in the next chapter about the WRW-model and subsequently also for the formulation of SRME and IME.

### 2-3-1   Convolution

The convolution of two time signals $f(t)$ and $g(t)$ is defined by:

$$h_{fg}(t) = f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad . \qquad (2\text{-}13)$$

In the space-frequency domain convolution simplifies to a scalar multiplication,

$$H_{fg}(\omega) = F(\omega)G(\omega) \quad . \qquad (2\text{-}14)$$

Note that interchanging the two signals gives the same result.

### 2-3-2   Correlation

The correlation of two time signals $f(t)$ and $g(t)$,

$$h_{fg}(t) = \int_{-\infty}^{\infty} f(\tau)g(t + \tau)d\tau \quad , \qquad (2\text{-}15)$$

simplifies to a scalar multiplication of one time signal with the complex conjugate of the other time signal in the space-frequency domain,

$$H_{fg}(\omega) = F^*(\omega)G(\omega) \quad . \tag{2-16}$$

Contrary to convolution, interchanging the two correlated signals does not give the same result, but its time-reversed version.

To be more precise, convolution and correlation of two signals in the space-frequency domain is done by multiplying their amplitude spectrum and, for convolution: adding their phase-spectrum; for correlation: subtracting their phase-spectrum [Oppenheim, 1983]. This explains the effect that these two operations have on the resulting wavelet.

The frequency spectrum of seismic traces is always defined by a center frequency or frequency-"plateau", and by flanks, (see figure 2-5). Therefore multiplying two seismic amplitude spectra always results in steeper flanks and, therefore, a narrower bandwidth. This is illustrated using figure 2-6, showing one minimum phase-wavelet and its corresponding bandwidth. After convolution with itself (auto-convolution) the bandwidth is visibly narrower resulting in the creation of additional side-lobes in the time domain (see figure 2-7). This is even more pronounced after additionally correlating the result with the initial wavelet (see figure 2-9). Note that the phase of the initial wavelet was restored. An interesting fact regarding the phase-change of similar wavelets after correlation is that they always result in approximate zero-phase wavelets (figure 2-8).
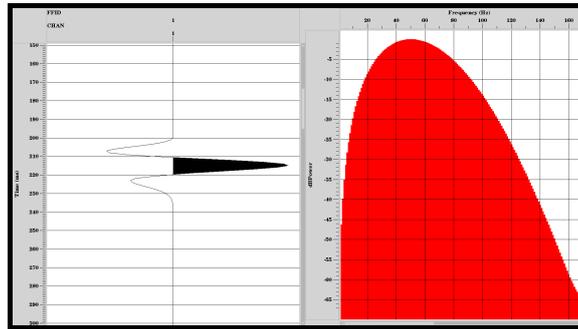
**Figure 2-6:** Minimum phase wavelet and its corresponding bandwidth.
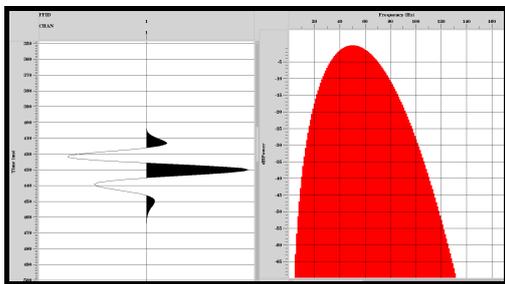


**Figure 2-7:** Minimum phase wavelet auto-convolved.
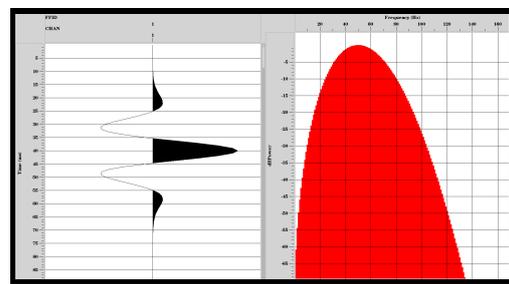


**Figure 2-8:** Minimum phase wavelet auto-correlated resulting in a zero-phase wavelet.
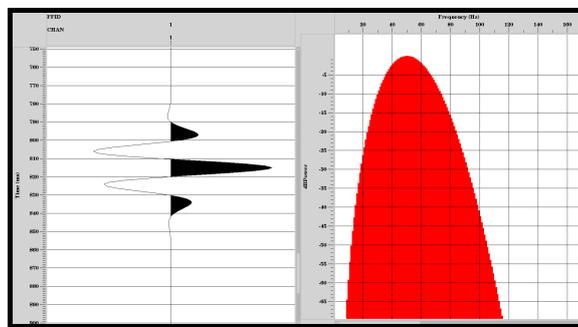


**Figure 2-9:** Minimum phase wavelet auto-convolved first and then correlated with itself.

## 2-4   WRW model

The WRW-model [Berkhout, 1982] is an attractive framework for the development and
derivation of seismic processing algorithms. Comparable to a set of Lego, it allows the con-
struction of a desired and more complex output, by combining individual pieces, describing a
single process or property. Each piece represents a discrete monochromatic vector or matrix
in the $(x, y, \omega)$ domain.

In its simplest form it conceptually formulates how a seismic signal enters the subsurface
and is altered before arriving back at the surface where it is subsequently recorded, resulting
in a primary wavefield measurement. For the purpose of this thesis the 2D-implementation
of this framework is considered, avoiding unnecessary multidimensional matrices, facilitating
intuitive understanding.

A 2D seismic experiment is conducted by placing a source at the surface and recording
the different signal-arrivals, caused by the interaction of emitted wavefield and medium, at
specific receiver stations (see figure 2-10). This is repeated for several shots, each time moving
the shot-location to a different surface station (see figure 2-11). Each shot-record results in
several recorded traces, containing time-sampled amplitudes (figure 2-12). All survey-samples
in the volume $[X_s, X_d, t]$ with $X_d$ representing different receiver-stations and $X_s$ representing
different shot-stations can be stored in a multidimensional data-matrix. Then, each col-
umn (indicated in blue in figure 2-13), represents one recorded trace for a specific $X_d - X_s$-pair.

For further processing the data-matrix is transformed into the $(X_s, X_d, \omega)$-domain using the
temporal Fourier-transformation (figure 2-14). Each sample is now a complex number con-
taining the amplitude and phase information for one frequency and trace. Therefore, each
horizontal data-slice (indicated in orange) is a monochromatic matrix, enabling the process-
ing of for each frequency-component individually. It is of importance to realize that matrix
multiplications, as done in the WRW-model, are the discrete equivalents of a spatial integrals,
representing convolution.
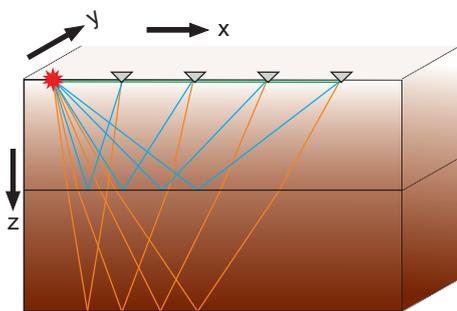


**Figure 2-10:** Seismic acquisition for one shot.
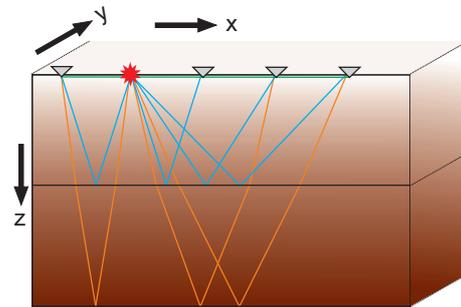


**Figure 2-11:** Seismic acquisition, shifted by one station, for the next shot.
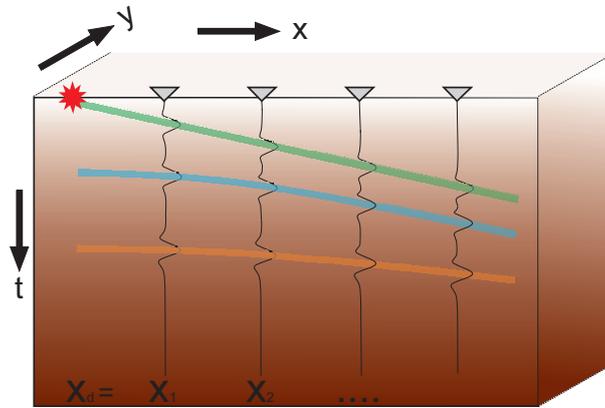
**Figure 2-12:** Recorded data for one shot.
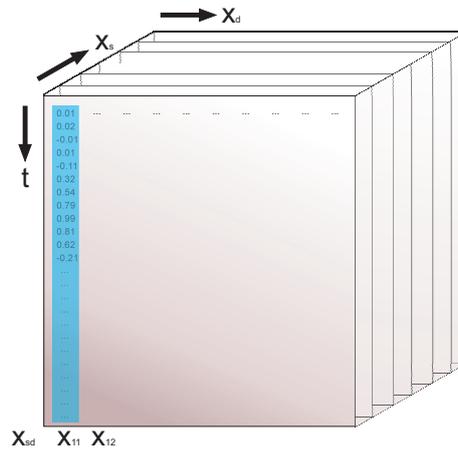


**Figure 2-13:** Datavolume in the time domain.
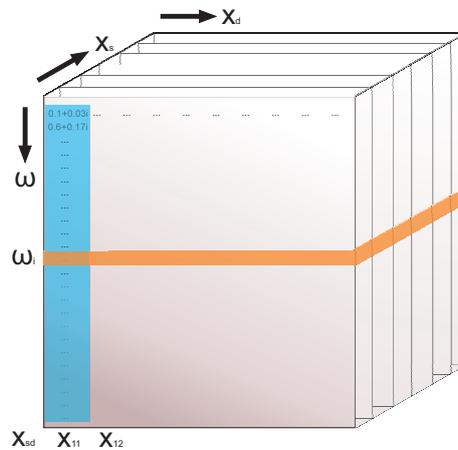


**Figure 2-14:** Datavolume transformed to the temporal-frequency domain.

### 2-4-1   Involved operators and notation convention

The following convention [Berkhout and Verschuur, 2005a] is used for the representation of different wavefields, properties and operators:

| Operator | Description |
|---|---|
| $P$ | Data-matrix; contains all seismic measurements for one frequency |
| $W$ | Wavefield extrapolation operators; contains angle-dependent propagation properties; a column represents an upgoing/downgoing impulse response |
| $S$ | Source matrix; contains angle-dependent source information |
| $D$ | Detector matrix; contains angle-dependent receiver information |
| $R$ | Reflection matrix; contains angle-dependent reflection properties for downwards propagating waves at one depth level; it contains full elastic scattering |
| $X$ | Transfer matrix; contains earth's impulse responses |
| $M$ | Data-matrix; contains all multiples |
| $A$ | Wavelet-deconvolution and scaling matrix; contains inverse of $S$ and $D$, and scales by $R(z_0, z_0)$ or $\hat{R}(z_n, z_n)$; usually estimated from the data using a least-squares filter |
| $F_{pr}$ | Prediction filter; multichannel and multirecord filter |
| $F_{ls}$ | Least-squares filter; minimizes the difference in the adpative subtraction |

| Notation | Description |
|---|---|
| $\hat{R}, \hat{X}$ | Scattering operator; transforming upwards propagating waves into downwards propagating waves |
| $P^+$ | The + denotes a downgoing wavefield |
| $P^-$ | The - denotes an upgoing wavefield |
| $P_s$ | The s denotes a column-vector related to the shot $X_s$; common shot gather |
| $P_d$ | The d denotes a row-vector related to the receiver $X_d$; common receiver gather |
| $P_{sd}$ | One trace related to the shot $X_s$ and receiver $X_d$ |
| $\Delta P$ | The $\Delta$ denotes that only primary reflection information is contained |
| $\{P\}_n$ | The $\{\}_n$ denotes that all multiples related to $z \leq z_n$ have been removed |
| $\{P\}_0$ | The $\{\}_0$ denotes that all surface-related multiples have been removed |
| $\{\delta M\}_n$ | This denotes the multiple wavefield related to one depth level $z_n$ |
| $\overline{P}(z_n, z_0)$ | The $^-$ denotes that all reflections (primaries and multiples) up to $z \leq z_n$ are muted |

### 2-4-2   Primary wavefield measurement

As the WRW-model describes how a wavefield is constructed one frequency-component at a time, the data-matrix has to be flipped and turned, as is illustrated in figure 2-15, so that the subscripts $s$ and $d$, denote a column and row respectively.

The primary wavefield measurement for one common shot gather (indicated in blue in Figure 2-15) is given by,
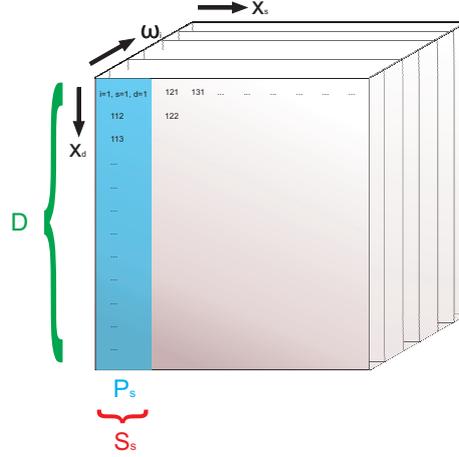
**Figure 2-15:** Datavolume in the frequency domain. In red and green, the individual dimensions of the source vector and receiver matrix are indicated.

$$\Delta P_s(z_0, z_0) = D(z_0)\Delta P_s^-(z_0, z_0) \quad , \tag{2-17}$$

indicating that the upgoing wavefield is affected by the receiver-properties. The involved upgoing reflected wavefield, for this shot gather, can be further unraveled to

$$\Delta P_s^-(z_0, z_0) = \sum_{m=1}^{\infty} W(z_0, z_m)R(z_m, z_m)S_s^+(z_m, z_0) \quad , \tag{2-18}$$

where the downgoing source wavefield at every depth-point $z_m$ is scaled by its reflectivity and then forward propagated to the surface as an upgoing wavefield. It originates from the source-signal emitted at shot $X_s$, which is forward propagated to every depth-level. This is describes in

$$S_s^+(z_m, z_0) = W(z_m, z_0)S_s(z_0) \quad . \tag{2-19}$$

For the land-case, it is necessary to acquire multicomponent data, in order to remove all wave conversions. Then, one matrix cell becomes a sub-vector, as is described by [Wapenaar and Berkhout, 1990]. It should also be noted, that in the equations, the source and receivers are supposedly placed directly at the surface $z_0$, which is not always the case for field data. Therefore, any travel-time delays have to be accounted for, and de-ghosting applied on the receiver-side, decomposing the wavefield into its upgoing part. As the involved matrices are in reference to a dipole response, source-side de-ghosting can be omitted in practice.

Using only the smallest "building blocks", the primary wavefield measurement, from source to receivers, for one shot-gather, can be represented with the following equation,

$$\Delta P_s(z_0, z_0) = D(z_0)\sum_{m=1}^{\infty} W(z_0, z_m)R(z_m, z_m)W(z_m, z_0)S_s(z_0) \quad . \tag{2-20}$$

This is visualized in figure 2-16 with the corresponding colors. In general, in order to be accurate, $W$ has to account for mode conversion, elastic attenuation and, in case of a complex subsurface, for multiple arrivals [Berkhout and Verschuur, 2005a].
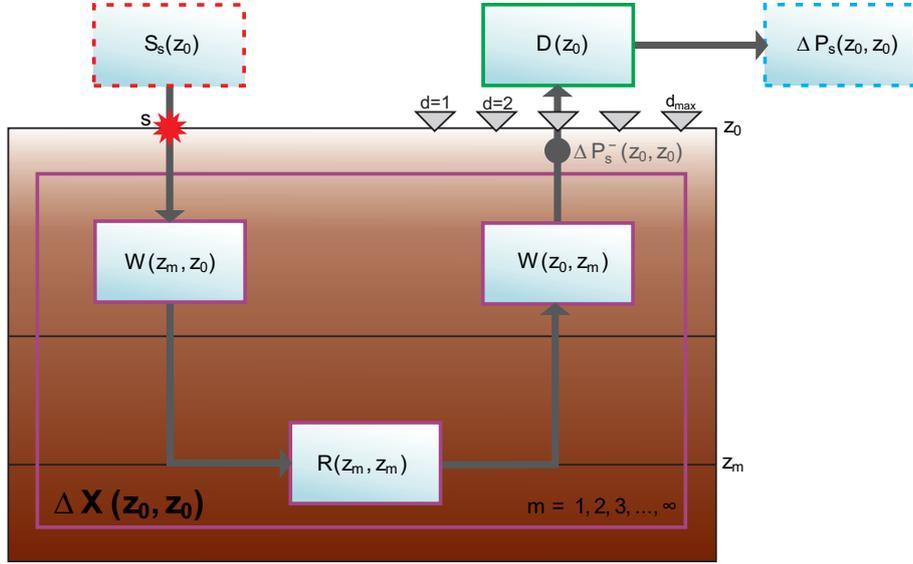


**Figure 2-16:** WRW-model, for one common shot record, resulting in a data-matrix column without multiples, showing the contribution from depth level $z_m$ only.

The whole subsurface-operator can be expressed using the multiple-free transfer-matrix $\Delta X$,

$$\Delta X(z_0, z_0) = \sum_{m=1}^{\infty} W(z_0, z_m) R(z_m, z_m) W(z_m, z_0) \quad . \tag{2-21}$$

Then equation 2-20 simplifies to,

$$\Delta P_s(z_0, z_0) = D(z_0)\Delta X(z_0, z_0)S_s(z_0) \quad . \tag{2-22}$$

### 2-4-3   Feedback model: Surface-related multiples

In oder to extend the model, to account for surface-related multiples, the upcoming wavefield has to be scaled, according to the free-surface reflectivity, and then fed back into the subsurface, adding a round-trip for each order of multiples. This is achieved using the so called feedback model illustrated in figure 2-17.

Therefore, equation 2-22 is extended by boundary operator $\hat{R}$, yielding $\hat{R}(z_0, z_0)P_s^-(z_0, z_0)$ as an additional downgoing wavefield at $z_0$:

$$P_s(z_0, z_0) = D(z_0)\Delta X(z_0, z_0)[S_s(z_0) + \hat{R}(z_0, z_0)P_s^-(z_0, z_0)] \quad . \tag{2-23}$$

Note that after introducing the internal multiples, $\Delta X(z_0, z_0)$ has to be changed to $\{X(z_0, z_0)\}_0$, in order to properly include them, as is done for SRME.
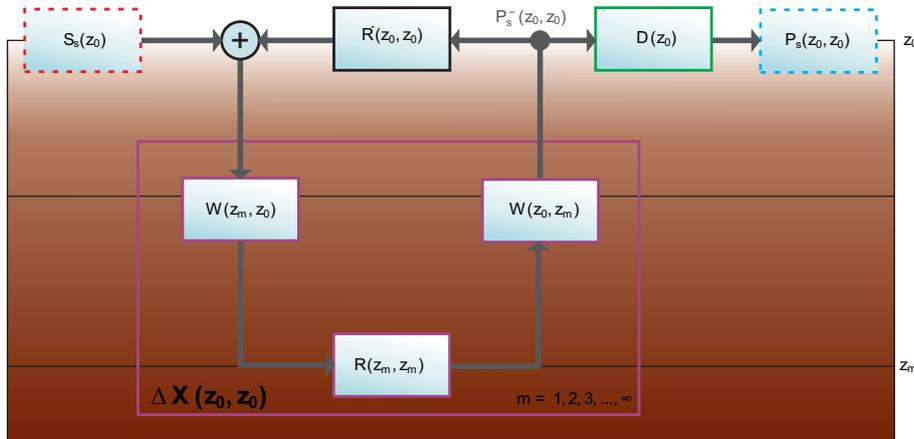
**Figure 2-17:** Feedback loop: WRW-model resulting in data with surface multiples.

## 2-4-4 Feedback model: Internal multiples

In a similar way, the WRW-feeback model can be adjusted to represent internal multiple generation (see figure 2-18). Each reflection boundary $z_n$ is then treated as the new surface for the feedback loop, again taking into consideration every depth-point $z_m < z_n$. As it is no longer situated at the surface though, the individual wavefields have to be extrapolated between the surface $z_0$ and depth-level $z_n$.
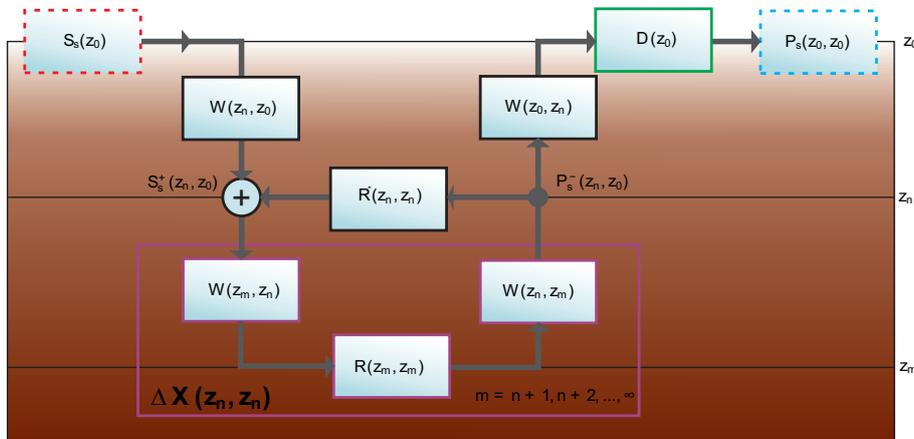


**Figure 2-18:** Feedback loop: WRW model resulting in data with internal multiples due to downward scattering at $z_n$.

## 2-4-5 Layer formulation

So far, the feedback model has only been expressed using specific downward scattering boundaries. It is therefore referred to as the boundary formulation. An alternative way to approach multiple-generation, is to not place depth-level $z_n$ on a boundary, but instead, to put it in-between boundaries. This is referred to as the layer formulation. The downward scattering operator $\hat{R}(z_n, z_n)$ is then replaced by the downward scattering operator $\Delta \hat{X}(z_n, z_n)$. It contains the impulse responses of the complete overburden, transforming and scaling upward

propagating wavefields at level $z_n$, into downward propagating wavefields, including all possible downward scattering within a certain layer above level $z_n$. This is displayed in figure 2-19. Note that scattering happens at many depth levels $z \leq z_n$.
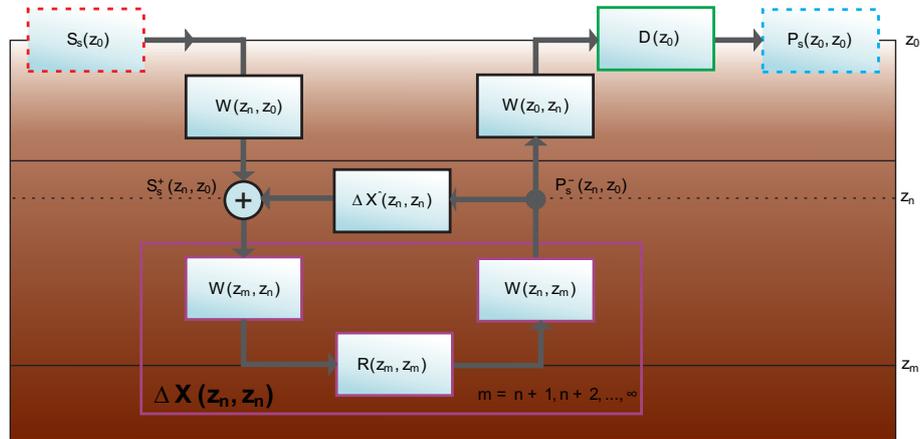


**Figure 2-19:** Layer formulation: WRW model resulting in data with internal multiples due to downward scattering within the complete layer above $z_n$.

# 3

# Multiple Elimination Algorithms

Multiple suppression methods, exploiting their periodic character, have been around for a long time, starting with the design of filters for the attenuation of water layer reverberations, using predictive deconvolution [Robinson, 1954]. By amplitude-scaling and time-shifting each single trace independently, it is loosely based on the wave equation. The simplicity of its application, though, leads to problems when the subsurface is laterally variant.

Since then, several more sophisticated algorithms have been developed, in order to remove or at least attenuate multiples. While varying in their approach, and therefore having different prerequisites, they all have in common that they try to create an accurate prediction of the multiple wavefield, before subsequently subtracting it from the recorded data.

This chapter goes into more detail on the derivation of the Jakubowicz-implementation of IME, as it is currently the only feasible algorithm for large 3D-datasets in the industry, before giving a short description of other popular and promising algorithms, while highlighting their different advantages and weaknesses.

## 3-1  2D-SRME

The development of SRME is based on the thesis-work of [Verschuur, 1991]. It was later published as a paper as "Adaptive surface-related multiple elimination" [Verschuur et al., 1992]. SRME is a data-driven multiple removal method, convolving already present reflections in the recorded data in order to construct surface-related multiples. By deriving it from the WRW-model, it is possible to illustrate the potential error-sources, making adaptive subtraction a requirement [Verschuur and Berkhout, 1997a].

### 3-1-1 SRME-derivation based on the WRW-model

Following the notation used in chapter 2-4 the expression for removing all surface-related multiples is

$$\{P(z_0, z_0)\}_0 = P(z_0, z_0) - \{\delta M(z_0, z_0)\}_0 \quad . \tag{3-1}$$

Based on equation 2-23, derived from the feedback model for surface-related multiples, and after having included the internal multiples, $\{\delta M(z_0, z_0)\}_0$ is given by

$$\{\delta M_s(z_0, z_0)\}_0 = D(z_0)\{X(z_0, z_0)\}_0 \hat{R}(z_0, z_0) P_s^-(z_0, z_0) \quad . \tag{3-2}$$

Implementing equations 2-18 and 2-19 for $P_s^-(z_0, z_0)$, this can be expanded to

$$\{\delta M_s(z_0, z_0)\}_0 = D(z_0)\{X(z_0, z_0)\}_0 \hat{R}(z_0, z_0) X(z_0, z_0) S_s(z_0, z_0) \quad . \tag{3-3}$$

By adding an additional source and receiver-term, $S$ and $D$, together with their corresponding inverses $S^-$ and $D^-$, it is possible, using the expressions given by equation 2-22, to represent surface multiples as a matrix-multiplication of two different data-matrices, resulting in

$$\{\delta M_s(z_0, z_0)\}_0 = \{P(z_0, z_0)\}_0 A(z_0, z_0) P_s(z_0, z_0) \quad , \tag{3-4}$$

where $A$ contains the source- and receiver-deconvolution-term, as well as the reflection information of the surface,

$$A(z_0, z_0) = S^{-1}(z_0)\hat{R}(z_0, z_0)D^{-1}(z_0) \quad . \tag{3-5}$$

The left expression of equation 3-4, is commonly referred to as the prediction filter, as it is not readily available in the recorded data. This is represented by

$$\{\delta M_s(z_0, z_0)\}_0 = F_{pr}(z_0, z_0) P_s(z_0, z_0) \quad , \tag{3-6}$$

with $F_{pr}$ being,

$$F_{pr}(z_0, z_0) = \{P(z_0, z_0)\}_0 A(z_0, z_0) \quad . \tag{3-7}$$

In the same fashion as equation 3-4, the multiple-prediction for one single trace, related to shot $X_s$ and receiver $X_d$ is given by

$$\{\delta M_{sd}(z_0, z_0)\}_0 = \{P_d(z_0, z_0)\}_0 A(z_0, z_0) P_s(z_0, z_0) \quad . \tag{3-8}$$

Essentially, this represents the spatial convolution of a surface-multiple free common-receiver gather with a common-shot gather, as is further illustrated in figure 3-1, using the corresponding colors.
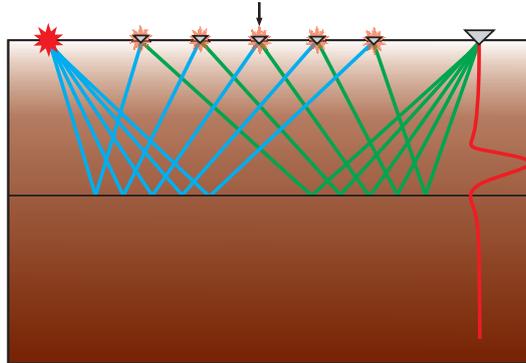
**Figure 3-1:** Prediction of surface-related multiples by convolving traces from a common-shot with those of a common-receiver gather.

By breaking it down into individual steps, the general approach can be explained in more detail. The multiple wavefield for a 2D seismic survey, can be constructed in the following way:

- For each source-receiver pair,

    1. for each grid point on the surface with a coinciding shot- and receiver-location,

        (a) the trace, resulting from a shot at the source position and a receiver at the grid-point, is deconvolved using the angle-dependent receiver properties at the grid-point;

        (b) the surface-multiple free trace, resulting from a shot at the grid-point recorded at the receiver-position, is deconvolved using the angle-dependent source-properties at the grid point;

        (c) both traces are convolved with each other resulting in one trace of the multiple contribution gather, located at the grid-point;

        (d) the trace is scaled using the angle-dependent reflection coefficient at the grid point.

    2. all traces in the multiple contribution gather are stacked horizontally, in direction of the acquisition-line.

As mentioned in this algorithm, each convolution, of traces related to the surface grid point, results in one trace contributing to the final multiple-trace. All grid point traces for one source-receiver pair therefore represent the so called multiple contribution gather. It is important to note that not all traces contribute constructively to the final multiple trace.

In general, the multiple contribution gather contains curved events. Only the apex of such an event and the traces in its vicinity are added constructively. All other wavelets cancel each other due to the general nature of band-limited seismic acquisition: The integral of a wavelet is approximately zero. This concept is illustrated in figure 3-2. Another important fact is that not only the reflection coefficients of the apex-traces influence the amplitude of the final multiple event, but also the curvature of the event. Events with small curvature result in
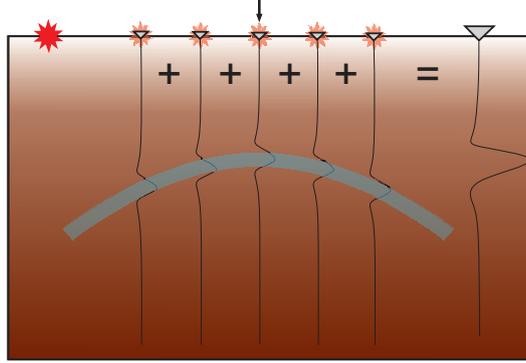
**Figure 3-2:** Multiple contribution gather depicting how the curved event constructs the multiple prediction trace after summation.

larger amplitudes, after stacking, than events with strong curvature.

As only the apex-traces of the multiple-contribution-events are of importance for the final multiple trace, it might be acceptable to limit the aperture of grid-points along the surface, in order to reduce computation cost. Note that the spatial location of the apex represents the coordinate of the reflection point at the surface (indicated by the vertical arrow in figure 3-1). However, the location of the apex depends on the complexity of the subsurface, as for dipping reflectors, it is no longer at the midpoint. This is illustrated in figures 3-3 and 3-4. Interestingly, it is not required to know where the apex of the event is located, because as long as it is present within the multiple contribution gather, the correct arrival time and amplitude, will be mapped to the correct trace.

For densely sampled, noise-free, reflection data, correctly processed by the derived algorithm, straightforward subtraction of the multiple prediction trace from the recorded trace,

$$\{P_{sd}(z_0, z_0)\}_0 = P_{sd}(z_0, z_0) - \{\delta M_{sd}(z_0, z_0)\}_0, \tag{3-9}$$

would probably give an acceptable result. In practice however, we omit several of the afore-mentioned steps, due to two unknowns.

### 3-1-2   Adaptive subtraction due to unknown $A$

Generally no exact, angle-dependent, knowledge of the different elements of $A$, source wavelet, receiver properties and reflection coefficient of the multiple generator, is available. Therefore $A$ has to be estimated from the data. This is done by reformulating the subtraction-process, as an adaptive subtraction,

$$\{P_{sd}(z_0, z_0)\}_0 = P_{sd}(z_0, z_0) - F_{ls_d}(z_0, z_0)\{\delta M_{sd}(z_0, z_0)\}_0 \quad , \tag{3-10}$$

using a least-squares filter $F_{ls_d}(z_0, z_0)$ [Verschuur and Berkhout, 1997a]. This assumes that the total energy of the seismic trace after subtraction is minimum. To achieve this, the filter corrects the amplitude and phase of the multiple prediction so that it fits the data. Relating
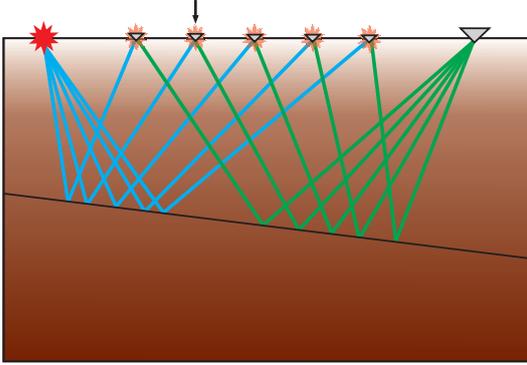
**Figure 3-3:** SRME for dipping reflectors in the in-line-direction.
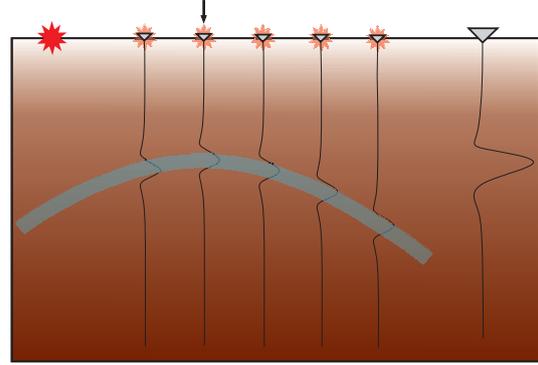


**Figure 3-4:** Multiple contribution gather for a 2D-subsurface. The apex is longer located directly between source and receiver.

this to the individual elements of $A$, this approximately means that the amplitude is scaled, to represent the reflection coefficient, and the wavelet reshaped in order to account for the source- and receiver-deconvolution.

But the adaptive subtraction does not differentiate between different error sources, and inaccuracies due to noise, unknown $A$, bad amplitude-balancing, omitted de-ghosting etc. are all handled in one single step. Most often, it is simply assumed that adaptive subtraction will be able to handle one single error source, disregarding the fact, that the total prediction error is a lot more complex, convoluted and unsystematic.

### 3-1-3    Iterative implementation due to unknown $\{P_d(z_0, z_0)\}_0$

Another approximation has to be done regarding $\{P_d(z_0, z_0)\}_0$, as surface-multiple free receiver gathers are commonly not available. Already having them, would defeat the purpose of having to do SRME. Fortunately, as long as the correct primary reflections are present in the input-data, the result will normally contain less multiples. Therefore SRME can be implicitly formulated in an iterative fashion as,

$$\{P^{i+1}(z_0, z_0)\}_0 = P(z_0, z_0) - \{P^i(z_0, z_0)\}_0 A(z_0, z_0) P(z_0, z_0) \quad , \tag{3-11}$$

meaning that the multiple attenuated output of SRME is used as input for another iteration [Verschuur and Berkhout, 1997a]. For the first iteration, commonly the total data is used as an approximation,

$$\{P^0(z_0, z_0)\}_0 = P(z_0, z_0) \quad . \tag{3-12}$$

Generally three iterations for SRME and one for IME are enough to achieve an acceptable result [Berkhout and Verschuur, 2005a].

## 3-2  1D-SRME (CMP-implementation)

Conceptualized at a time when pre-stack migration was still regarded as an uncommon processing step, due to its computational cost, 2D SRME had similar limitations. Therefore a simplified application of SRME was proposed, suited for approximate (locally) 1D subsurfaces [Verschuur, 1991].

Based on the observation that for a 1D medium, a CMP-gather represents a shot gather, 1D SRME uses CMP-gathers as input. For a true 1D medium, all shot records are essentially equal, and can be representend by one single CMP-gather. Even though this is rarely, if ever, the case, approximating CMP-gathers as shot gathers has its advantages, potentially outweighing the resulting inaccuracies due to the incorrect assumption.

By transforming CMP gathers to the wavenumber-frequency domain, the matrix multiplications of 2D SRME simplify to scalar multiplications. This considerably speeds up computation. However, re-sorting shot gathers to CMP-gathers, has the negative side-effect of CMP-gathers having only half the spatial sampling, potentially leading to a spatially aliased multiple prediction.

In order to avoid this, CMP-gathers can be merged together, not only to groups of two, but even a lot higher, essentially avoiding the problem of spatial aliasing [Kelamis and Verschuur, 2000]. These so called supergathers, also have the significant advantage of allowing strong noise attenuation filters. This is especially important for land-data, which is normally characterized by a bad signal-to-noise (S/N) ratio. As SRME is completely data-driven, using noisy data for the convolutions results in a poor multiple prediction. Examples of successful applications of CMP-SRME and CMP-IME can be found at [Alá'i et al., 2003] and [Alá'i and Verschuur, 2006] respectively.

## 3-3  3D-SRME

When the subsurface becomes structurally too complex, and can no longer be described as a 2D-medium, 2D SRME is no longer able to correctly predict the arrival time and amplitude of the multiple wavefield, as the apex of the event, in the multiple contribution gather, representing the location where the multiple reflects downward at the surface, is no longer located under any grid-point on the acquisition-line (illustrated in figure 3-5). This is generally the case for out-of-plane point diffractors or reflectors dipping in the cross-line direction.

Apart from requiring the sampling of the recording wavefield in the y-direction, in order to sample the 3D-wavefield, extension of 2D to 3D SRME is quite straight-forward as is illustrated in figure 3-6. The only differences are, that the aperture determining which grid-points are used for convolution, is now extended to the y-direction (see figure 3-7). The result are more complex 3D-events, curved in all azimuths. Therefore horizontal summation has to be done twice, in orthogonal directions, commonly the x- and y-direction, in order to retrieve the multiple prediction trace.
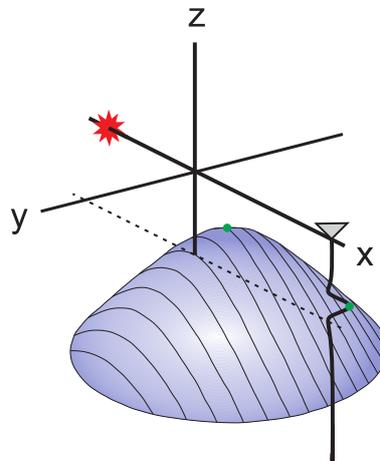
**Figure 3-5:** 3D multiple contribution gather illustrating a hyperboloid and the multiple trace resulting from stacking in x- and y-direction. The apex-contribution is indicated in green.
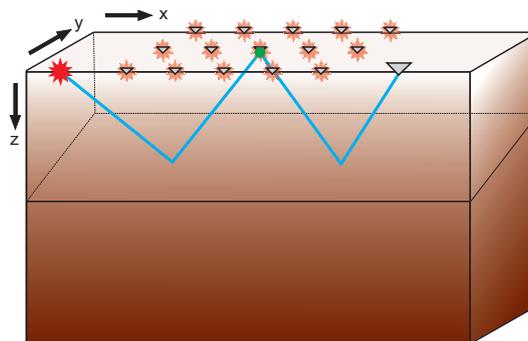


**Figure 3-6:** 3D SRME with 3D acquisition, where the multiple reflection point is located up-dip in the cross-line direction (green circle).
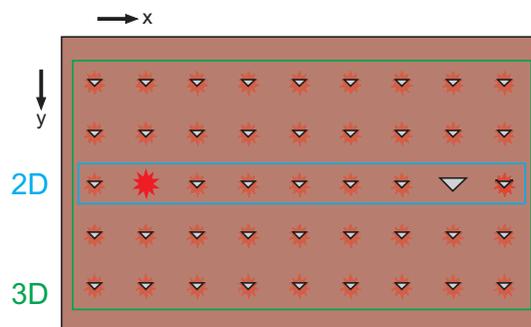


**Figure 3-7:** Different apertures in 2D and 3D-SRME applications.

## 3-4 Extension to Internal Multiples

Multiples are always related to a down-ward scattering reflection interface. For surface-related multiples, this is the free-surface; luckily, the same interface where the seismic measurements are conducted. However, for internal multiples, this interface is located in the subsurface, where no direct measurements were made. Therefore, SRME can conceptually be extended, for internal multiples, in a straight forward fashion, by treating any subsurface reflector as the measurement surface. This can be achieved by re-datuming the sources and receivers, to said reflector. As [Verschuur and Berkhout, 1997a] pointed out, this turns out to not be necessary. Only one side, the receivers for the common-shot-gather, and the sources for the common-receiver-gather, need to be re-datumed. This configuration is visualized in figure 3-8 where the grid-points on the surface, such as visible in figure 3-1, are replaced by grid-points on the chosen boundary.
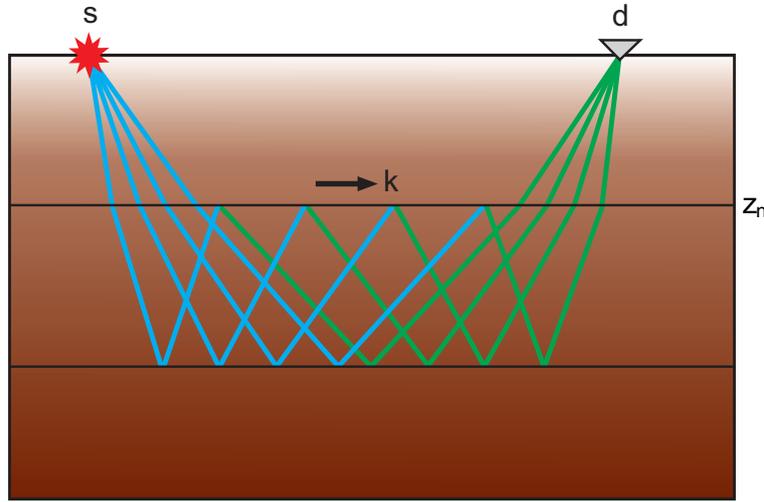


**Figure 3-8:** Half-redatumed common-shot- and common-receiver-gather related to several subsurface grid-points at depth-level $z_n$.

Redatuming to depth-boundary $z_n$ can be done by using a velocity model in order to retrieve the necessary Green's functions. Therefore, using an erroneous velocity model will, if not corrected, create an erroneous multiple prediction.

Following from the WRW-model for internal multiples, the multiple wavefield related to a subsurface boundary $z_n$ can be retrieved in the following fashion,

$$\{\delta M_{sd}(z_0, z_0)\}_n = \{\overline{P}_d(z_0, z_n)\}_n A(z_n, z_n)\{\overline{P}_s(z_n, z_0)\}_{n-1} \quad , \qquad (3\text{-}13)$$

where $A(z_n, z_n)$ contains the downward scattering operator $\hat{R}(z_n, z_n)$. Broken down into words, this means that the multiple-prediction, related to a downward-scattering boundary $z_n$, for one single source-receiver pair, is calculated by the spatial convolution of, a half-redatumed shot-gather, containing no multiples related to boundaries above $z_n$ and muted up to and including $z_n$, with a half-redatumed receiver-gather, containing no multiples related to boundaries above and including, $z_n$, muted up to and including $z_n$. Each convolution result

is scaled by the angle-dependent reflection coefficient of boundary $z_n$, and deconvolved for source and receiver-characteristics in a similar way as done for SRME.

As was the case for SRME, $A$ is normally not known, and compensated for by adaptive subtraction. Commonly, also no common-receiver gather, containing no multiples related to $z_n$, is available. Therefore, implementation of IME is done in an iterative fashion, implicitly formulated as,

$$\{P^i(z_0, z_0)\}_n = \{P(z_0, z_0)\}_{n-1} - \{\overline{P}^i(z_0, z_n)\}_n A(z_n, z_n)\{\overline{P}(z_n, z_0)\}_{n-1} \quad , \qquad (3\text{-}14)$$

where the output is half-redatumed and muted, before subsequently being used as input for the next iteration. Commonly,

$$\{\overline{P}^0(z_0, z_n)\}_n = \{\overline{P}(z_0, z_n)\}_{n-1} \quad , \qquad (3\text{-}15)$$

is used for the first iteration.

Note that multiple prediction is only done for one single boundary $z_n$, and requires the multiple-removal of all boundaries above $z_n$ beforehand. Therefore, retrieving completely multiple free data, has to be done in a top-down fashion, for every single reflection interface. As commonly only a few strong reflectors contribute to the multiple problem, applying IME for even one single boundary can produce the desired result.

### 3-4-1 CFP-boundary-IME

So far, the algorithm uses data organized in common-shot- and common-receiver-gathers. But it can also be organized using so called common-focus-point-(CFP)-gathers [Thorbecke, 1997], where the data is resorted for one focus point on the chosen boundary with its many corresponding sources and receivers on the surface, as is illustrated in figure 3-9. Then, instead of calculating the partial contributions for one shot record, the partial contributions for one CFP gather are retrieved. Reorganizing the data into CFP gathers has several advantages [Berkhout and Verschuur, 2005a]:

- The necessary muting of the involved gathers becomes straight forward, as the focusing operators used for re-datuming, can be time-reversed and then used as mute lines.

- For boundary-related IME, the focusing operator can be updated, in order to correct for errors in the velocity model, so that boundary-related CFP-IME is still a completely data-driven approach (see also [Berkhout and Verschuur, 2005a]).

### 3-4-2 CFP-layer-IME

Continuously updating the depth model for each chosen boundary is not really desirable for IME. Therefore, similar to the layer formulation of the WRW model, IME can also be applied in the same fashion, by using focus points on a chosen depth-level $z_n$, normally lying in-between reflectors. Then, instead of predicting only multiples related to one boundary, all
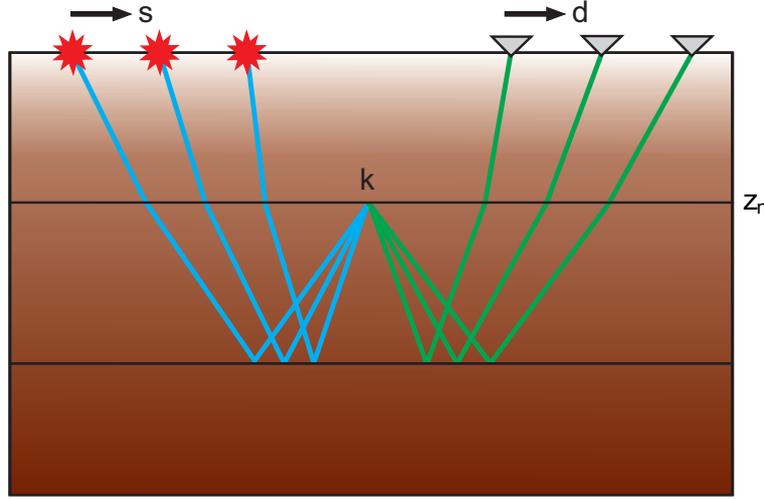
**Figure 3-9:** Internal multiple prediction reformulated for CFP-gathers.

multiples, for which the corresponding ray-path crosses the chosen layer at least four times, are predicted. This is illustrated in figure 3-10. Boundary operator $\hat{R}(z_n, z_n)$ has to be replaced by layer operator $\hat{X}(z_n, z_n)$, representing the reflectivity of the whole overburden, or a part of it. For the data-driven approach, the unknown $\hat{X}(z_n, z_n)$, is instead replaced by wavefield $\Delta Q(z_n, z_n)$ (indicated in red in figure 3-10); a wavefield containing only primaries, resulting from a source on $z_n$, at focal point $k$, illuminating the overburden and being recorded on $z_n$, at focal point $l$. Therefore, $\Delta Q(z_n, z_n)$ is a double-focused gather, that contains no multiples from layers above $z_n$, has been time-reversed, and who's anti-causal part has been removed by muting [Berkhout and Verschuur, 2005b]. It is often bounded between two levels and rarely represents the whole overburden. The whole approach can be represented more explicitly as

$$[M_{sd}(z_0, z_0)]_n = -\sum_{k,l} \{\overline{P}_{ld}(z_0, z_n)\}_n \Delta Q_{kl}(z_n, z_n) \{\overline{P}_{sk}(z_n, z_0)\}_{n-1} |S(z_0)|^{-2} |D(z_0)|^{-2} \quad ,$$

(3-16)

where $[M_{sd}(z_0, z_0)]_n$ denotes the whole family of multiples generated by the chosen layer $n$, for one source-receiver-pair. As three wavefields are combined in order to predict the multiple wavefield, deconvolution of source and receiver-characteristics has to be done twice, but only for the amplitude spectrum. The phase spectrum is already correct, as was explained in chapter 2-3. Note that the chosen implementation of the deconvolution terms implies that the directivity of $D$ and $S$ is omitted, and that they can therefore be represented by scaled diagonal matrices. Also note, that only a subset of all multiples are predicted, but even placing only one layer between two strong reflector-packages, can already produce the desired result.
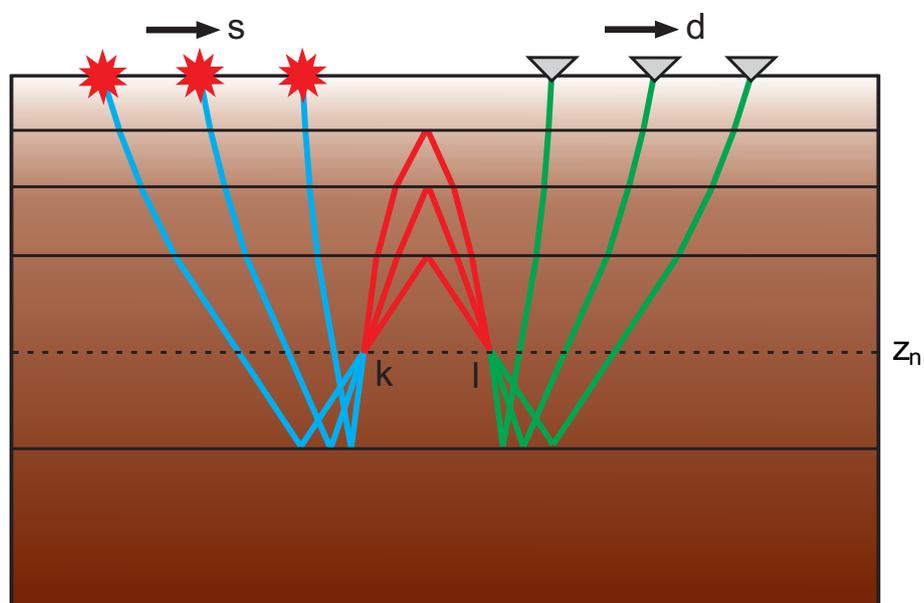
**Figure 3-10:** Internal multiple prediction, for the layer-approach, reformulated for CFP-gathers (in blue and green) and double-focused gathers (in red).

## 3-5   Jakubowicz Method

Partially redatuming the wavefields to boundary $z_n$ is not the only approach, by which SRME can be extended for the internal multiple case. [Jakubowicz, 1998] realized that, similar to the implementation of SRME, it is possible to create the multiple wavefield by combining three already recorded wavefields without the need of back-propagation. This becomes intuitively apparent when the corresponding ray-paths, involved in the generation of an internal multiple, are fully drawn. This is illustrated in figure 3-11, where it can be seen, that the ray-path of the internal multiples, related to depth level $z_n$, for one source-receiver pair, share part of the ray-paths of already recorded primary reflections, indicated in blue and green. The part of the ray-path that is not involved in the construction of the internal multiples, is visualized in red. Luckily, it is also already contained in the recorded wavefield, as the primary reflections of reflection interface $z_n$.
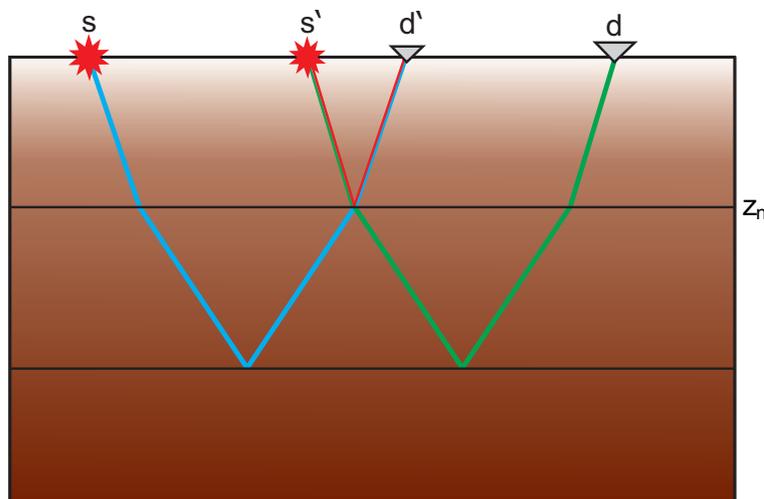


**Figure 3-11:** Illustration of the different ray-paths involved in the Jakubowicz-approach.

Therefore, it can be removed by correlating the result of the convolution of the blue and green trace, with the primary reflection of the boundary $z_n$, indicated by the red path. This essentially represents back-propagation towards a point on the boundary $z_n$. Correlation is essentially the same as convolution with the complex-conjugate trace. Figure 3-12 illustrates how the traces are convolved in order to construct the internal multiple trace. As was already shown in chapter 2-2, complex-conjugation of a trace only affects the phase-spectrum and not the amplitude-spectrum.



**Figure 3-12:** Illustration how the individual ray-paths are convolved with each other in order to retrieve the multiple ray-path.

In a similar way as it is done in SRME, construction of the multiple prediction trace is achieved by repeating the convolution-process for each surface-grid-point pair $s'd'$. This essentially means convolving one trace of a common-shot gather, with each trace of a common-receiver gather, each time applying back-propagation by convolving the result with the time-reversed primary reflection of $z_n$ for the corresponding surface-grid-point pair $s'd'$. Lateral summation, of all convolution-ouputs, results in one trace of the multiple contribution gather. Repeating the process for each trace in the common-shot gather fills it up. Summing the traces laterally again, gives the multiple prediction trace for the shot-receiver pair $sd$. This spatial convolution process using matrix-multiplication, is described, as

$$[M_{sd}(z_0, z_0)]_n = -\{\overline{P}_d(z_0, z_n)\}_n |S(z_0)|^{-1}|D(z_0)|^{-1}\overline{\underline{P}}^H(z_n, z_n)|S(z_0)|^{-1}|D(z_0)|^{-1}\{\overline{P}_s(z_n, z_0)\}_{n-1} \quad,$$
$$(3\text{-}17)$$

where the superscript $^H$ indicates the Hermitian conjugate of the matrix. The following data is needed for this approach:

- A common-shot-gather $P_s(z_n, z_0)$, muted for all primary reflections related to $z \leq z_n$, and containing no multiples related to $z < z_n$.

- The time-reversed primary reflection wavefield $P(z_n, z_n)$ related to layer $z_n$, achieved by muting all other events.

- A common-receiver-gather $P_d(z_0, z_n)$, muted for all primary reflections related to $z \leq z_n$, and containing no multiples related to $z \leq z_n$.

- The source and receiver amplitude-spectrum properties of each surface-grid point.

Repeating this algorithm for each source-receiver pair $sd$, results in the multiple wavefield related to one subsurface-boundary $z_n$.

Jakubowicz-IME faces many similar complications as CFP-IME, and handles them in the same way:

- Source and receiver-deconvolution for the amplitude spectrum, has to be applied twice, as three wavefields are involved in the prediction-process, where one is in correlation mode.

- Source and receiver properties are often not known, and are therefore compensated for by adaptive subtraction.

- The actual implementation is done in an iterative fashion, as the common-receiver gathers initially still contain multiples related to boundary $z_n$.

- Complete multiple removal has to be done in a top-down approach, for each subsurface reflection boundary.

## 3-6   Wave equation based multiple suppression methods

### 3-6-1   Model vs. Data-driven algorithms

One way to differentiate multiple suppression algorithms, is to separate them based on their required input. Model-driven methods require additional input-information, such as a reflectivity-, structure- or velocity-model. The quality of their multiple prediction, depends on the accuracy of said input-models. Often, only the most important reflections or velocity-trends are included in so called "macro-models". Creating and utilizing more localized and detailed models, often comes at the expense of an increased computation-time, an important current limitation for this group of algorithms. Basing the prediction-algorithm on a model though, has the distinct benefit of being less restricted on the model-sampling, often only being limited by the receiver- and not the source-spacing. Usually, model-driven methods rely on wavefield extrapolation of the measurements through a certain layer in order to simulate an extra roundtrip. First-order multiples are created from primaries, second-order multiples from first-order multiples, and so on (see [Berryhill and Kim, 1986] or [Wiggins, 1988]).

Data-driven methods have the advantage of not requiring any a priori information on the subsurface. They utilize the fact that the available recorded wavefields, having traveled through the subsurface, already contain all the necessary information. This, being their biggest advantage, is also their shortcoming: as recorded wavefields are commonly not sampled densely or regularized enough, especially in the cross-line direction, the resulting predicted multiples will suffer from aliasing effects. Therefore, interpolation of missing traces is often a requirement.

### 3-6-2   Surface-related and internal multiple modeling

The so-called method surface-related multiple modeling (SRMM), and its extension, internal multiple modeling (IMM), is based on the work of [Pica et al., 2005]. It combines the model- and data-driven approach. First, the recorded data is migrated in the pre-stack domain. Either, de-migration algorithms, are then used to create a more densely sampled data-set, effectively eliminating any constraints on the acquisition geometry. The newly sampled data-set is then used as input for 3D-SRME.

Or, choosing a different approach, the migrated section is assumed to accurately represent the subsurface reflectivity, and using a velocity macro-model the data is then forward propagated, based on the one-way wave equation, to each depth level, scaled by the estimated reflectivity, and then forward propagated to the surface, one shot record at a time.

Being able to re-sample the recorded wavefield, is a considerable strength of this approach. For further information, the reader is referred to [Pica et al., 2005] for SRMM, [Pica et al., 2008] for IMM, and [Yang et al., 2012] for shallow water application.

### 3-6-3   Multiple attenuation based on the inverse scattering series

Internal multiples differ from primaries, in that they were scattered more than once in the subsurface. This allows the application of inverse scattering sub-series-algorithms in order to predict them ([Weglein et al., 1997]). Multiples, are the result of upwards reflected wavefields, scattered downwards, in a shallower part of the subsurface, before being scattered upwards again. This so called "lower-higher-lower" relationship is used as a constraint, for the selection of 3 (or 5, or 7...) different imaged data points on a migrated section. The contribution of all possible combinations of data points is then summed, resulting in the multiple wavefield, after de-migration. Therefore all internal multiples (including converted phases [Coates and Weglein, 1996]) of one given order are predicted, and not only the ones related to a specific boundary.

Despite only being based on an approximation, their arrival time is correct. However, their amplitudes are slightly too low (typically around 80% - 95%) [Verschuur, 2006a]. As the prediction is directly subtracted, this means that methods based on inverse-scattering are normally described as multiple attenuation algorithms. Continuous research, in the last decade, led to the complete removal of a subset of internal multiples ([Ramírez et al., 2008]), and also to algorithms addressing surface-related multiples ([Li and Hu, 2009]).

Contrary to other methods, all internal multiples, of one order, are handled at the same time without the need for a-priori information or user interpretation. However, despite all these advantages, ISS-based algorithms are hindered by their large computational cost.

### 3-6-4   EPSI

Estimation of Primaries by Sparse Inversion (EPSI), was introduced by [van Groenestijn and Verschuur, 2009]. Being based on the WRW-model, it has been generalized in the recent years to also include internal multiples, effectively reformulating and combining SRME and IME [Ypma and Verschuur, 2013]. However, current implementations are only formulated for the 2D-case.

The primary impulse responses are parameterized and used in a large-scale inversion process, aiming to iteratively create a primary-wavefield estimation, which, combined with the associated predicted multiples, converge with the recorded data. This is done in a top-down approach for each interface. Utilizing a minimization function, which will theoretically, eventually go to zero, adaptive subtraction is avoided.

EPSI, can also be extended for internal multiples; then it requires subsurface knowledge and user input, by dividing the data into several layers, and also requiring an estimate of the source-wavelet, which is then used for the whole dataset.

As it also relies on wavefield-convolution in order to predict multiples, near-offset data is crucial. However, the general approach of EPSI is to extract information contained in the

multiples in order to estimate primaries. Each iteration of EPSI takes twice the computational time of SRME or IME. In order to arrive at an acceptably convergence, EPSI requires several tens of iterations. Therefore, computational cost is currently a considerable limitation for the application of EPSI, and will only be worse for any 3D implementation.

### 3-6-5   Marchenko imaging

Marchenko imaging is a newly re-opened field of research in the seismic industry by [Broggini et al., 2012], based on the Marchenko equations. It allows the retrieval of the up- and down-going wavefield, at any focal point in the subsurface using an iterative scheme. The only requirements are, surface-multiple-free seismic-reflection data, recorded at the surface, and an initial estimate of the direct wavefield from said surface to the chosen focal point.

The recorded data is then cross-correlated with the direct wavefield, which essentially represents back-propagation to the focal point. Reverse-time migration (RTM) follows a similar process, but stops at this point. Its result is therefore degraded by spurious events, due to remaining internal multiples, not being focused correctly. Marchenko imaging improves upon this method using auto-focusing. It no longer considers internal multiples as noise, but instead values its contribution to the imaged section.

Auto-focusing is applied by truncating the result of the cross-correlation at arrival times earlier than the direct arrival. Signals arriving before the truncation-line are then used to update the focusing operator (the initial direct wavefield) for the next iteration. This focusing operator aims at focusing primary reflections, as well as all orders of internal multiples at the focus point. Eventually, by updating it with new events, and applying enough iterations so that the amplitudes converge, spurious events are canceled. While each even iteration updates the downgoing wavefield, and each odd iteration updates the upgoing wavefield, the seismic image in the end is obtained using a multidimensional deconvolution of these separate wavefields [Wapenaar et al., 2014].

Applying Marchenko imaging to all subsurface-points will result in a multiple free migrated section. This is however computationally quite expensive, and still being researched as truncations are not, yet, well defined for the 2-dimensional or 3-dimensional case [Wapenaar et al., 2014]. In order to produce correct amplitudes and achieve good results, this method needs an accurate source wavelet, background velocity model, and a true-amplitude reflection response, as well as requiring de-ghosting, and noise attenuation before-hand.

For an illustrative description of Marchenko imaging the reader is referred to [van der Neut et al., 2015].

### 3-6-6  Multiple elimination based on Marchenko imaging

As applying Marchenko imaging to all points in the subsurface is quite expensive, [Meles et al., 2014] proposed a multiple removal method, only using a few focal points, in order to create a multiple prediction. I will henceforth refer to it as Marchenko multiple elimination (MME). It is based on the removal of the direct waves from the autofocused Greens functions. Primaries, being of scattering order one, have one part of the up/down-going wavefield that reaches the focal point as a direct wave. Multiples arrive as reflections from both sides. Based on this simple observation, it is possible, by removing the direct waves, to essentially remove the primaries.

Implementation of MME is done by retrieving the up- and down-going wavefield, and removing the direct waves, at focal points along an arbitrary level in the subsurface. This level is not required to be horizontal. The multiple prediction is constructed by integrating over the chosen boundary. The main-contribution comes from points of stationary phase, essentially where the correct multiple ray-path crosses the boundary, and their vicinity. Therefore only those multiples are predicted, for which the integration boundary lies between the reflecting interfaces. As higher order multiples cross the boundary several times, they have more stationary points.

While theoretically their exact phases are predicted, inaccuracies in the auto-focused up- and down-going wavefields affect the result. Relative amplitudes are also wrong, as they are predicted too high for higher order multiples [Meles et al., 2014]. Other contributing error-sources are, inaccurate source-signature deconvolution, noise, attenuation and variations in source/receiver coupling. Therefore adaptive subtraction makes MME potentially more robust.

## 3-7  Similarities between MME, CFP-IME and Jakubowicz-IME

Multiple elimination after [Meles et al., 2014], has many similarities to the CFP-layer implementation of IME. Both of them only predict multiples for which the raypath crosses the chosen level at least four times. While Jakubowicz-IME and CFP-boundary-IME, only predict multiples related to one subsurface-boundary, all these methods convolute or correlate wavefields, and then integrate over many contributions, in order to arrive at the multiple prediction.

The difference, though, is how they retrieve the needed wavefields for the summation. Marchenko imaging retrieves the up- and down-going wavefield at all points on the boundary. But MME has to first eliminate the direct arrivals from the down-going wavefield before the multiple prediction. CFP-IME and Jakubowicz-IME achieve the same, by imposing restrictions on their input data. In the Jakubowicz-method, requiring, for the shot-gather, that all multiples, generated above the considered multiple generator, are already removed, and muting its associated primary reflection, is essentially the same as removing the downgoing wavefield for points on the boundary. By applying reciprocity, for the case

that receiver-properties and source-properties have been removed or are equal, the same can be said for the common receiver-gather. This would mean that for every point on the multiple generator the total upgoing wavefield is known, and that there is no down-going wavefield. As scaling the up-going wavefield of the shot-gather with the reflection-strength of the multiple generator, represents tranformation into a down-going wavefield, the multiple prediction gather can then constructed by convolution of the corresponding wavefields, with opposing propagation-direction, at each point on the boundary.

The Jakubowicz approach also exchanges the redatuming, based on correlation with the direct wave-estimates, by instead correlating with the primary reflection of the multiple generator. Potential amplitude inaccuracies due to an incorrect direct wave-estimate, are therefore traded, with amplitude errors due to incorrect transmission operators, which will be discussed later in the thesis.

As all of these methods, are based on the convolution of wavefields, they are all prone to prediction errors caused by: spatial under-sampling, incorrect or omitted deghosting, amplitude balancing, attenuation-compensation, and noisy data. Therefore, in the end, they all rely on adaptive subtraction.

# Part II

# Second Part

# 4

# Theoretical limitations

This chapter will discuss the theoretical shortcomings of multiple prediction methods based on wavefield-convolution, in more detail. Special focus will be on the Jakubowicz-implementation of IME, due to its popularity. However, as most methods use the same processes and have similarities, most of the mentioned error-sources are also applicable to them. In general, prediction errors for multiple events can be classified into three categories:

- Arrival-time errors of the multiple event

- Amplitude-errors of the multiple event

- Phase- and amplitude-spectrum errors in the wavelet of the multiple event

Note that they are interconnected to some extent, and that most error-sources create all three of them.

## 4-1 Absolute and relative amplitude errors

Producing absolutely correct amplitudes would require the convolved wavefields, to be true unit-valued representations of the subsurface-impulse response. However, this is not realistic for field data [Verschuur et al., 1992]. But it is also not necessary, due to adaptive subtraction. Therefore, it is important to realize that only correct relative amplitudes need to be retrieved during the multiple prediction for one multiple generating boundary $z_n$. This means, that the amplitudes have to be relatively correct,

- between multiples of different order.

- spatially, between different traces.

- between multiples generated from different up-wards scattering boundaries at $z > z_n$.

Then, adaptive subtraction would only need to apply a scalar amplitude correction factor.

## 4-2   Neglecting the structural subsurface dimension

The goal of the seismic method is to image the three dimensional subsurface reflectivity, defined by density and seismic velocity. In order to achieve this, a seismic wavefield is sampled at the surface, and then inverted using digitally implemented algorithms. However, to accommodate our limitations in computation power, theoretical understanding, and spatial sampling, algorithms are often based on simplified theories, making the assumption that the subsurface parameters are locally somewhat invariant, by replacing the 3D character of the earth by a 2D or even 1D approximation. But, in the case of SRME and IME, even when these assumptions are met, they do not correctly reproduce amplitude of the multiple wavefield. Even worse, when the assumptions are not met, additional arrival time errors are introduced and the amplitude- and phase-errors become larger.

### 4-2-1   1D vs. 3D

Using a synthetic dataset, representing 3D marine acquisition with 6, 2km-long, streamers separated by 100m, along 12, 4km long, sail-lines separated by 150m, the shortcomings of a 1D-IME algorithm on 3D subsurface data are illustrated. The model consists of four reflectors, three of them horizontal, with the other one having a 12° dip in in-line and cross-line direction. The resulting primary wavefield, and its corresponding multiple wavefield, were modeled using 3D ray-tracing. Figure 4-1 shows a stacked section, containing only the primary reflections, while figure 4-2 has also the internal multiples included.
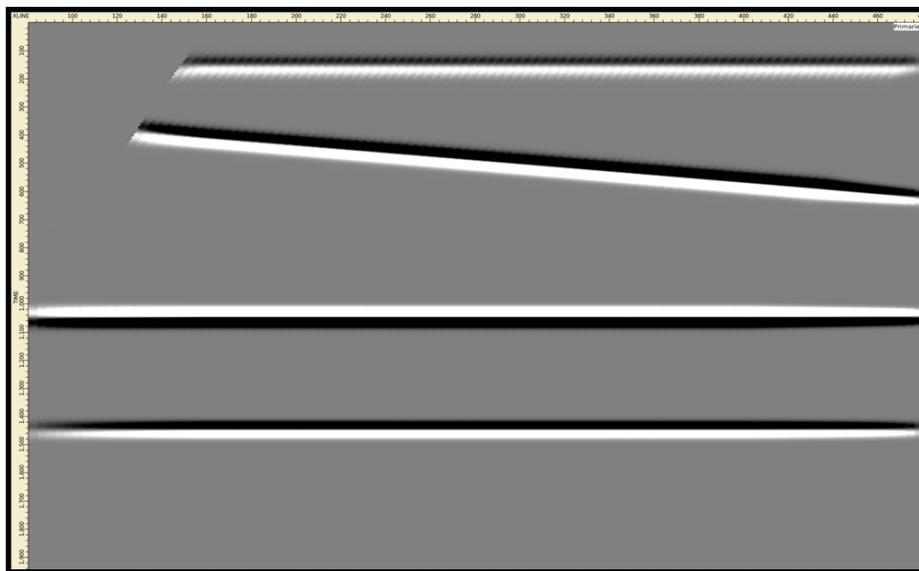


**Figure 4-1:** Synthetic data from a four-reflector medium, containing only primaries.

The 1D-IME algorithm used, is a layer-related CMP-implementation, resorting the data into CMP-gathers and treating them each separately, based on the work of the DELPHI-consortium at TU Delft. As a reference, a 3D-IME algorithm based on the Jakubowicz approach, currently being developed by DMT Petrologic, was also applied to the data, using

**Figure 4-2:** Synthetic data containing primaries and multiples.

a 25m x 25m surface-grid. Note, that due its specific implementation and the usage of ray-tracing for modeling, no amplitude-observations were made. For the 1D-algorithm, a layer between the first and second reflector was chosen, while for the 3D-algorithm the first reflector was chosen as the down-ward scatterer. This means that both algorithms only predict the multiples related to the first reflector, but not the multiples created between the 2nd and 3rd reflector, dipping in the opposite direction. Figure 4-3 shows the multiple prediction, before adaptive subtraction, of the 1D-algorithm, while figure 4-4 shows the result of the 3D-algorithm.

It becomes very apparent that the 1D algorithm cannot handle the dipping reflector and only vaguely predicts events resembling the multiples, while the 3D algorithm taking account of the subsurface-complexity is able to adequately predict them. Note that the specific version of the 3D-algorithm which was used for this test still produces artifacts, by still containing weak imprints of some primaries.
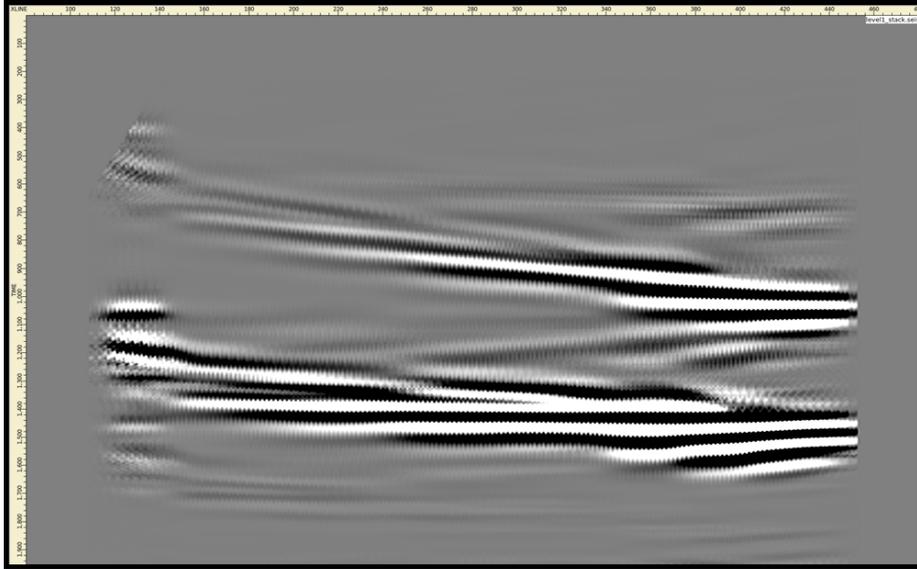
**Figure 4-3:** Synthetic dataset: Multiple prediction of the 1D-CMP-layer-IME-algorithm, for the first layer.
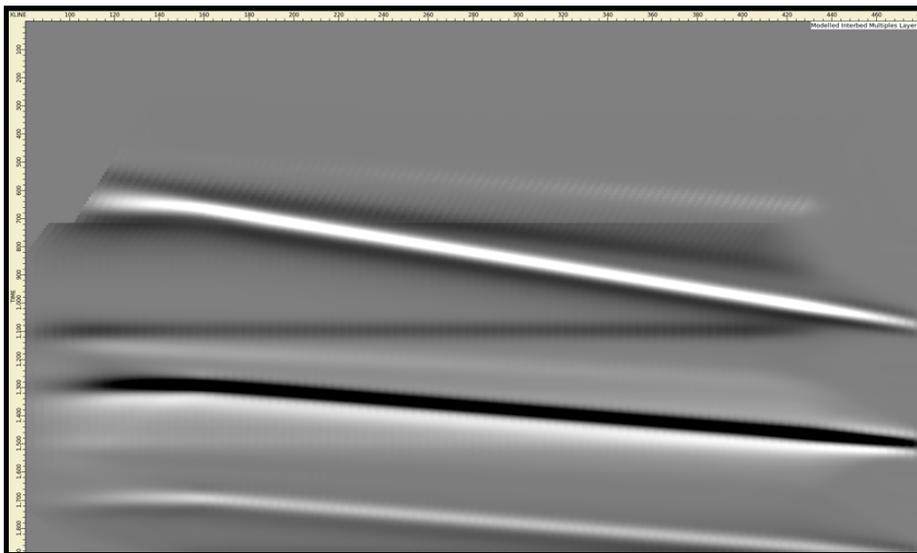


**Figure 4-4:** Synthetic dataset: Multiple prediction of the 3D-Jakubowicz-IME-algorithm, for the first layer.

### 4-2-2    2D vs. 3D

As the 1D-algorithm operates on differently sorted data, it is not surprising that the multiple prediction for a complex subsurface is so poor. The 2D and 3D algorithm, on the other hand, are quite similar and only differ in the spatial sampling of the wavefield, and therefore multiple contribution gather. The prediction error comes from the fact that the location, where the physical ray-path of the multiple reflects down-wards at the surface or subsurface boundary (the apex-location of the curved event in the multiple contribution gather), is not present in the recorded data. This is illustrated in figure 4-5 for a surface-related multiple. Figure 4-6 illustrates the curved event in the 3D multiple contribution gather resulting from such a multiple. The multiple prediction of the 2D-algorithm is based on the constructive summation of this gather by convolving the shot- and receiver-gathers recorded on the acquisition-line (indicated in green). In red, the line that would be necessary for a 3D-acquisition, in order to predict the multiple event with the correct arrival time, is indicated. It is of interest that the local curvature of the green and red line differs in most cases.
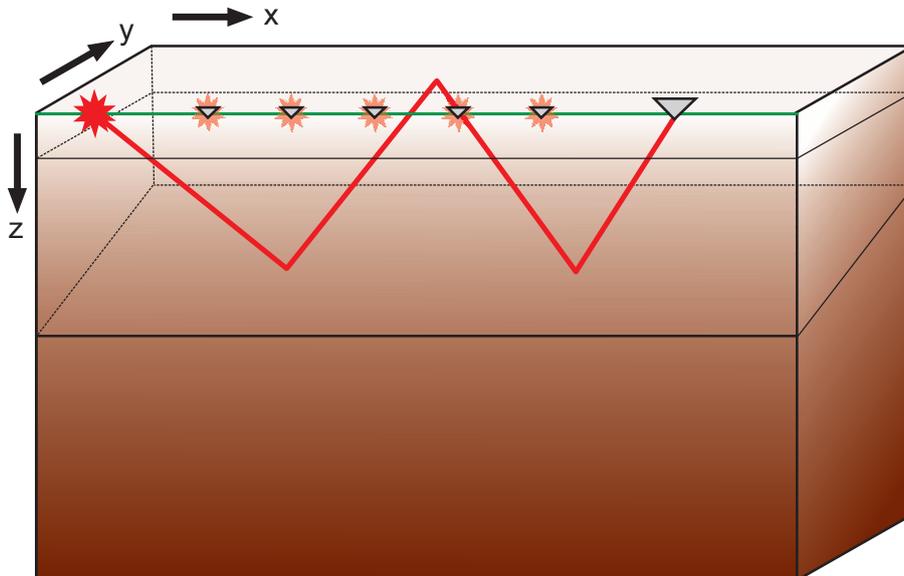


**Figure 4-5:** Physical ray path of a multiple recorded in a 2D-survey.

A more detailed look at the individual ray-paths of the primaries, creating the apex of the curved events, reveals that they traveled with different angles and through different reflection-interface-locations in the subsurface (see figure 4-7). Note, that for educational purposes, the simple case of a surface-related multiple and a very weak overburden reflector, not influencing the ray-angle, were chosen. But the same observations hold true for internal multiples, and stronger overburden impedance-contrasts.

Based on these observations, the following potential prediction errors are created, by using 2D-SRME/IME-algorithms, when the assumption, that the earth does not vary in cross-line direction, is invalid:

- Arrival-time error due to the lower-placed apex of the 2D curved event with regard to the real apex of the 3D curved event.
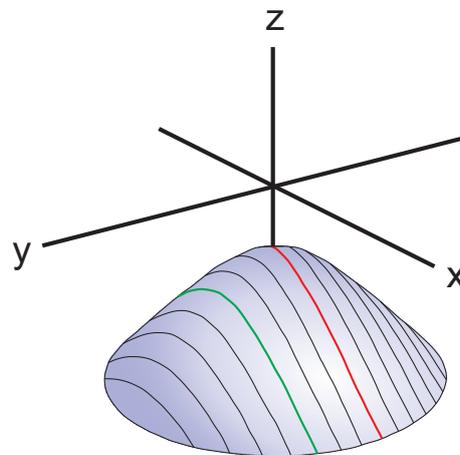
**Figure 4-6:** 3D multiple contribution gather: 3D curved event with individual lines showing the corresponding 2D curved events equivalent to 2D acquisition.
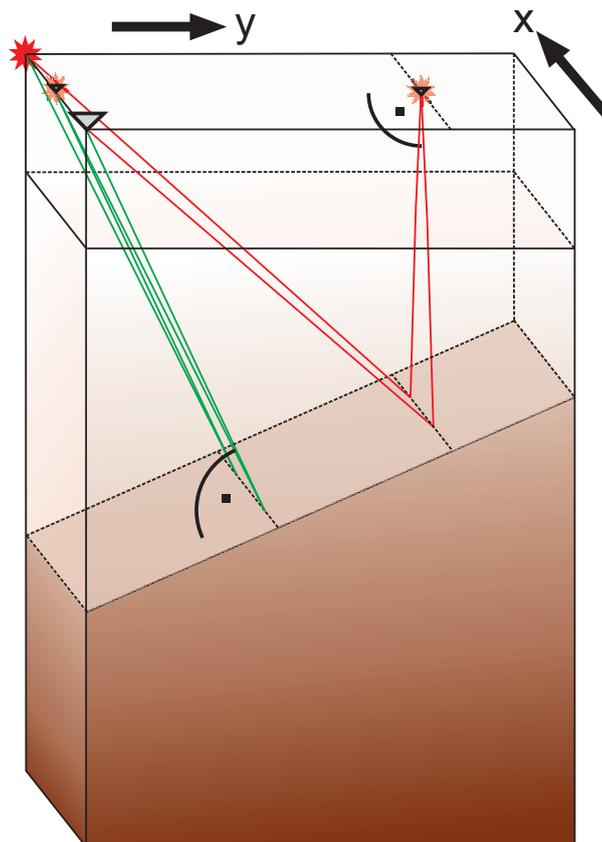


**Figure 4-7:** Different ray-paths of the physically correct multiple (red) and the multiple predicted by 2D SRME (green).

- Amplitude error because the contributing apex-traces encountered different transmission/reflection coefficients at each reflector in the subsurface.

- Phase and amplitude error because the contributing apex-traces experienced different source- and receiver-angles.

- Phase and amplitude error because of a locally different curvature at the apex.

It should also be noted, that even if there is no dip in the cross-line direction, that the 2D-algorithms will not predict the amplitudes correctly, because near-apex traces, from all the other azimuths, are not included in the summation. Normally this is accounted for by rescaling the data so that the point source becomes a line source. However, common approaches only work for 1D media and create additional artifacts in other media [Wapenaar et al., 1992]. In practice, commonly a pre-scaling factor of $\sqrt{t}$ is used (Verschuur 2015: personal communication).

## 4-3    Jakubowicz-IME: approximation of transmission effects

The Jakubowicz-approach uses three different already recorded wavefields in order to predict the multiple wavefield. Or broken down into smaller parts: Three different ray-paths are used to approximate the multiple ray-path. Each of the individual recorded ray-paths, representing for example a primary reflection, has been scaled by the encountered transmission and reflection operators. So, while combining the three ray-paths gives the correct arrival time, this is not true for the amplitude, due to incorrect transmission/reflection operators.

For the acoustic case the transmission and reflection operators in the wavenumber-frequency domain, $\tilde{R}$ and $\tilde{T}$, are introduced by observing how down- and up-going wavefields react when they encounter an impedance contrast [Wapenaar, 1989] (figures 4-8 and 4-9).
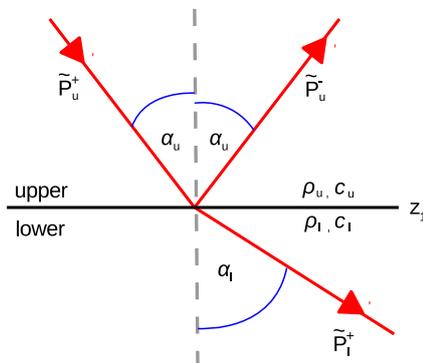


**Figure 4-8:** Down-going wavefield encountering an interface.
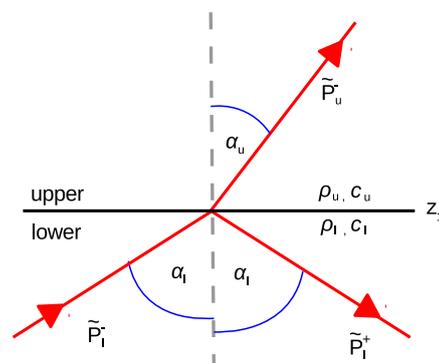
**Figure 4-9:** Up-going wavefield encountering an interface.

Adding to the previously used notation, the subscripts " $_u$" and " $_l$" denote the upper and lower layer respectively. Also, differing from the previous notation, $\tilde{R}^+$ and $\tilde{R}^-$ represent the reflection operators for down-going, and up-going wavefields respectively. The following holds true for a wavefield propagating downwards, encountering an impedance contrast:

$$\tilde{P}_u^-(z_1) = \tilde{R}^+(z_1)\tilde{P}_u^+(z_1) \quad . \tag{4-1}$$

$$\tilde{P}_l^+(z_1) = \tilde{T}^+(z_1)\tilde{P}_u^+(z_1) \quad . \tag{4-2}$$

In a similar way, for wavefields propagating upwards, the following holds true:

$$\tilde{P}_l^+(z_1) = \tilde{R}^-(z_1)\tilde{P}_l^-(z_1) \quad . \tag{4-3}$$

$$\tilde{P}_u^-(z_1) = \tilde{T}^-(z_1)\tilde{P}_l^-(z_1) \quad . \tag{4-4}$$

The reflection and transmission operators used in these equations are given by the properties of the upper and lower layer as

$$\tilde{R}^+(z_1) = \frac{\rho_l k_{z,u} - \rho_u k_{z,l}}{\rho_l k_{z,u} + \rho_u k_{z,l}} \quad , \tag{4-5}$$

and

$$\tilde{T}^+(z_1) = \frac{2\rho_l k_{z,u}}{\rho_l k_{z,u} + \rho_u k_{z,l}} = 1 + \tilde{R}^+(z_1) \quad , \tag{4-6}$$

for the downward propagating wavefield; And as

$$\tilde{R}^-(z_1) = \frac{\rho_u k_{z,l} - \rho_l k_{z,u}}{\rho_u k_{z,l} + \rho_l k_{z,u}} = -\tilde{R}^+(z_1) \quad , \tag{4-7}$$

and

$$\tilde{T}^-(z_1) = \frac{2\rho_u k_{z,l}}{\rho_u k_{z,l} + \rho_l k_{z,u}} = 1 + \tilde{R}^-(z_1) = 1 - \tilde{R}^+(z_1) \quad , \tag{4-8}$$

for the upward propagating wavefield. Note the indicated possibility of expressing any of these operators using the other operators.
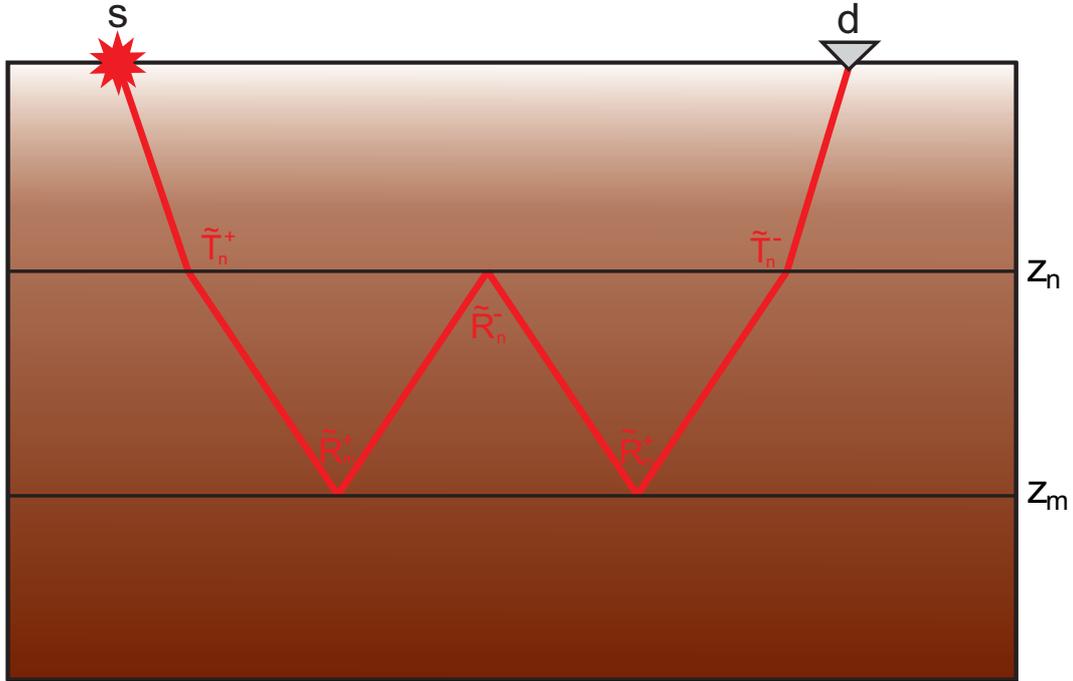
**Figure 4-10:** Physical ray-path of the internal multiple generated between horizontal reflectors at $z_n$ and $z_m$.

## 4-3-1   Transmission/Reflection operators for one horizontal reflector

Assuming a subsurface model with only one single horizontal boundary $z_n$, that is shallower than the horizontal boundary $z_m$, the associated ray-path of the internal multiple related to both layers, and recorded using a source at $s$ and a receiver at $d$, is illustrated in figure 4-10. The encountered transmission and reflection operators are also indicated.

Using the surface-grid-point-pair $s'd'$, having its midpoint at the horizontal location of the down-ward bounce of the multiple ray-path, the three ray-paths related to the primary reflections used in the Jakubowicz-approximation, are also investigated for their encountered transmission and reflection operators. This is shown in figure 4-11, using different colors. The same color-code is used in the following equations, allowing the assignment of the individual operators to their corresponding ray-path. The wavefields, for each ray, arriving at the surface can be implicitly expressed, omitting propagating and source/receiver effects, by operators acting on the emitted down-going wavefield,

$$P(sd) \sim \tilde{T}^+(z_n)\tilde{R}^+(z_m)\tilde{R}^-(z_n)\tilde{R}^+(z_m)\tilde{T}^-(z_n) \tag{4-9}$$

$$P(sd') \sim \tilde{T}^+(z_n)\tilde{R}^+(z_m)\tilde{T}^-(z_n) \tag{4-10}$$

$$P(s'd') \sim \tilde{R}^+(z_n) \tag{4-11}$$

$$P(s'd) \sim \tilde{T}^+(z_n)\tilde{R}^+(z_m)\tilde{T}^-(z_n) \quad . \tag{4-12}$$

In the Jakubowicz-approach, using these wavefields, the following approximation is done,
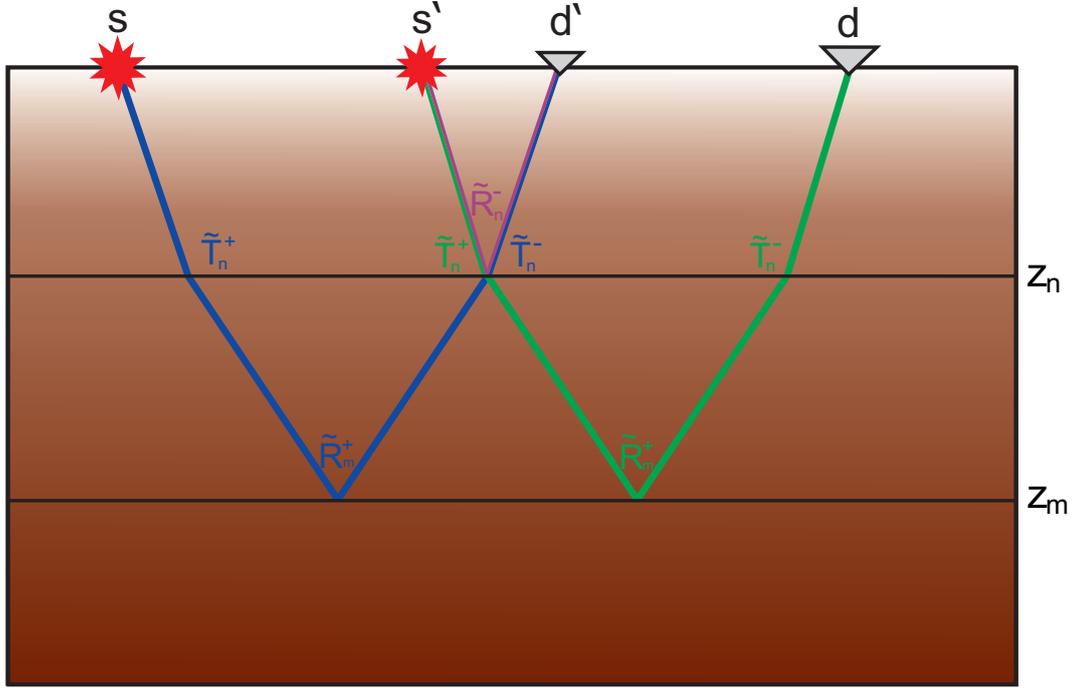
**Figure 4-11:** Physical ray-paths used for the approximation of an internal multiple ray-path related to horizontal reflectors at $z_n$ and $z_m$.

$$P(sd) \sim -P(sd') \cdot [P(s'd')]^* \cdot P(s'd) \quad . \tag{4-13}$$

Combining these expressions and eliminating the operators which are correctly accounted for, yields

$$\tilde{R}^-(z_n) \sim -\tilde{T}^-(z_n)[\tilde{R}^+(z_n)]^*\tilde{T}^+(z_n) \quad , \tag{4-14}$$

meaning that the down-ward scattering reflection operator $\tilde{R}^-(z_n)$ is approximated by minus the up-ward scattering reflection operator $-\tilde{R}^+(z_n)$, and that additional transmission operators are still present. Note that all operators are related to one focal point on the multiple generating boundary. Equation 4-14 can be further simplified to

$$1 \sim \tilde{T}^-(z_n)\tilde{T}^+(z_n) \quad , \tag{4-15}$$

as the remaining additional operators, included by using the Jakubowicz-approximation. In terms of the reflection strength of the multiple generating boundary $z_n$, the amplitude error can be expressed as

$$(1 - (\tilde{R}^+(z_n))^2) \quad . \tag{4-16}$$

This means, that, the stronger the reflection coefficient, the lower the predicted amplitude, compared to the real amplitude. For a laterally varying reflection strength this would mean that the prediction-error acts as a lateral smoothing-operator.

## 4-3-2   Transmission/Reflection operators for several horizontal reflectors

A similar thought-experiment can be done using a subsurface model with two weak interfaces on each side of reflector $z_n$, to investigate if they are properly included in the Jakubowicz approximation. This is illustrated in figure 4-12.
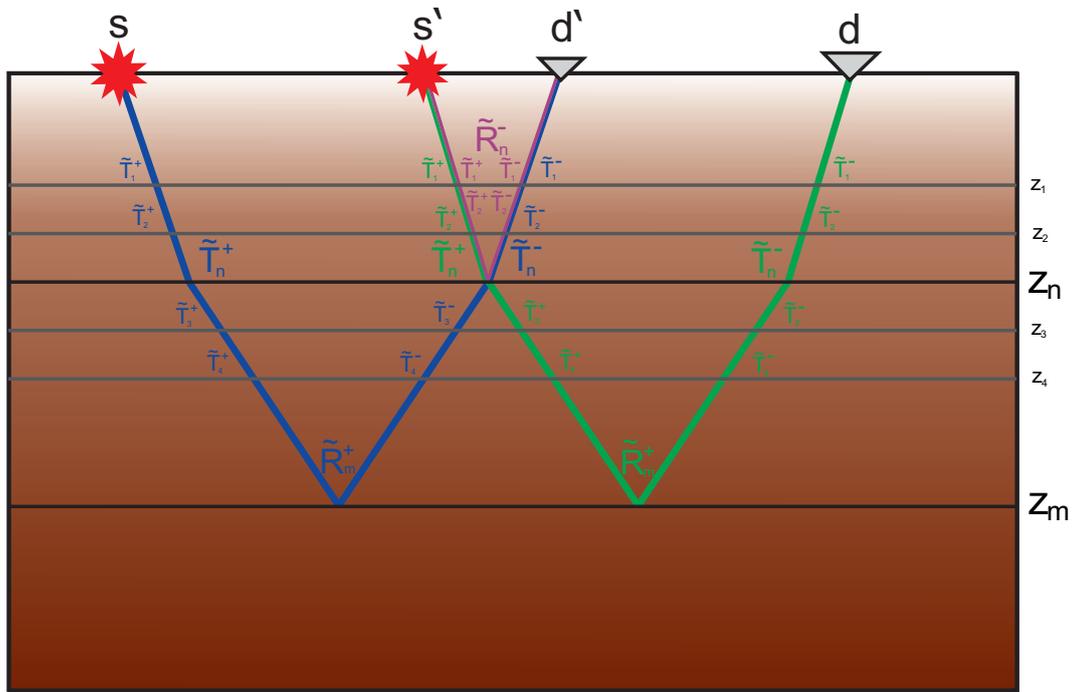


**Figure 4-12:** Transmission and reflection operators encountered by the physical ray-paths used for the approximation of an internal multiple ray-path related to horizontal reflectors at $z_n$ and $z_m$. The subsurface contains several other, weaker, interfaces.

In the same fashion, the upcoming wavefields can be expressed as

$$P(sd) \sim \tilde{T}_o^+ \tilde{T}^+(z_n) \tilde{T}_u^+ \tilde{R}^+(z_m) \tilde{T}_u^- \tilde{R}^-(z_n) \tilde{T}_u^+ \tilde{R}^+(z_m) \tilde{T}_u^- \tilde{T}^-(z_n) \tilde{T}_o^- \tag{4-17}$$

$$P(sd') \sim \tilde{T}_o^+ \tilde{T}^+(z_n) \tilde{T}_u^+ \tilde{R}^+(z_m) \tilde{T}_u^- \qquad\qquad \tilde{T}^-(z_n) \tilde{T}_o^- \tag{4-18}$$

$$P(s'd) \sim \tilde{T}_o^+ \qquad \tilde{R}^+(z_n) \qquad\qquad\qquad\qquad \tilde{T}_o^- \tag{4-19}$$

$$P(s'd') \sim \tilde{T}_o^+ \tilde{T}^+(z_n) \tilde{T}_u^+ \tilde{R}^+(z_m) \qquad\qquad \tilde{T}_u^- \tilde{T}^-(z_n) \tilde{T}_o^- \quad, \tag{4-20}$$

with the transmission operators of the overburden and underburden given by

$$\tilde{T}_o^- = \tilde{T}_1^- \tilde{T}_2^- \tag{4-21}$$

$$\tilde{T}_o^+ = \tilde{T}_1^+ \tilde{T}_2^+ \tag{4-22}$$

$$\tilde{T}_u^+ = \tilde{T}_3^+ \tilde{T}_4^+ \tag{4-23}$$

$$\tilde{T}_u^- = \tilde{T}_3^- \tilde{T}_4^- \quad . \tag{4-24}$$

Similar to equation 4-14 the following expression is derived,

$$1 \sim \tilde{T}^-(z_n)\tilde{T}^+(z_n)\tilde{T}_o^{2+}\tilde{T}_o^{2-} \quad . \tag{4-25}$$

This shows that the while the transmission operators of the underburden are properly included, this is not the case for the overburden. In fact, this is true for all forms of attenuation (spherical divergence, elastic attenuation etc.). As the total ray-path used for Jakubowicz-IME is longer than the actual ray-path, the multiple prediction is subject to additional attenuation. Therefore, it would be necessary to correct all the attenuation losses before-hand, resulting in actual representations of the subsurface-reflectivity.

In conclusion, neglecting wave-conversion (acoustic case), the following prediction errors are included in the Jakubowicz approach for IME:

- Amplitude error due to additional transmission-losses for the reflection-boundary $z_n$

- Amplitude error due to the additional quadratic transmission-losses of the total over-burden

# 5

# Practical limitations

## 5-1 Required Pre-processing

Raw field data differs quite a bit from our idea of suitable input-wavefields for multiple prediction. Therefore, it has to undergo many corrections beforehand. Failing to do so will lead to a poor multiple prediction. The following pre-processing steps are necessary before SRME or IME can be successfully applied:

- **Angle-dependent source- and receiver-deconvolution**

  Convolving a wavelet several times with itself results in a broader wavelet. The original wavelet may not be reconstructible using adaptive subtraction. This is discussed in more detail in section 5-3.

- **Amplitude balancing between different shot-records, and between each trace in one shot record**

  The recorded wavefields are supposed to be a representation of the subsurface reflectivity. Their relative amplitudes have to reflect that. Therefore, amplitude balancing is necessary to put them into relation. This is often necessary for land seismic data where the coupling of the source and receivers can vary from location to location. SRME and IME assumes a constant coupling across all sources and receivers. Note that the traces contributing to the multiple prediction trace are not the same as the ones used for the subtraction.

- **Elimination of non-reflection events such as noise or ground-roll**

  In general, using noisy data as input will result in a noisy multiple prediction, or even artificial events in the presence of strong-amplitude noise. This is discussed in more detail in section 5-8.

- **Muting the direct wave and reflections generated above the multiple generating boundary**

  SRME and IME assume the wavefield is upgoing at the surface. The direct wave should therefore be removed. For IME also all events from above the multiple generating boundary need to be removed. As they are of similar strength as the other involved primary reflections, they will produce artificial events, non-existent in the data, if included in the convolutions.

- **Interpolation of missing offsets and trace-regularisation**

  This is necessary in order to avoid spatial aliasing and artifacts in the multiple prediction, and is discussed further in section 5-9.

- **Deghosting**

  Because SRME and IME require the upgoing wavefield as input, the ghost (a downward component) needs to be removed. Depending on their periodicity, they will cause phase- and amplitude-errors, or even separate events, if included in the convolutions.

- **Multiple elimination for multiple generators above the chosen boundary $z_n$**

  If multiples from shallower layers are still present in the data then a subset of additional multiples will be predicted, which causes problems if IME is applied for more than one level, as these multiples are potentially predicted twice. This is illustrated in section 5-7.

- **Attenuation compensation**

  Partly because of attenuation, the recorded wavefields are not really representations of the subsurface reflectivity. The predicted amplitudes will be lower, if not compensated for. This was already discussed in section 4-3.

Note that not all of these processing steps can be perfectly applied for field data, because often the necessary information is missing. Therefore, we have to rely on adaptive subtraction.

## 5-2   High computational cost

One downside of 3D IME that was mentioned before was its high computational cost. Computational cost is quite different from the other complications discussed in this thesis, as it is the only limitation that eventually disappears with time. And even though Moore's law already slowed down, it can be expected that 3D IME-algorithms will be standardly implemented in most seismic processing suits in less than a decade. The same happened with 3D SRME, which was once thought to be too computationally expensive, and has now started to be considered as a common processing tool. This is especially of interest for the other multiple elimination methods mentioned in this thesis which are currently way too expensive to be of interest for the industry (3D ISS, 3D EPSI, 3D-CFP-IME, 3D Marchenko imaging).

Even for Jakubowicz-IME, extending the 2D-algorithm to a 3D-algorithm adds quite a considerable amount of computation time, as it depends on the number of grid-points involved. Assuming an aperture of $N$ grid-points in the in-line-direction, the computation time is of the order $N^2$. Using the same aperture in cross-line-direction would result in a computation time of factor $N^4$, a quadratic increase from the 2D case. However, using the same aperture is only necessary if the subsurface-complexity is expected to be the same in all directions (such as in the presence of a salt-dome). For other cases it might be acceptable to reduce the aperture in cross-line direction, in order to save computational time (figure 5-1).
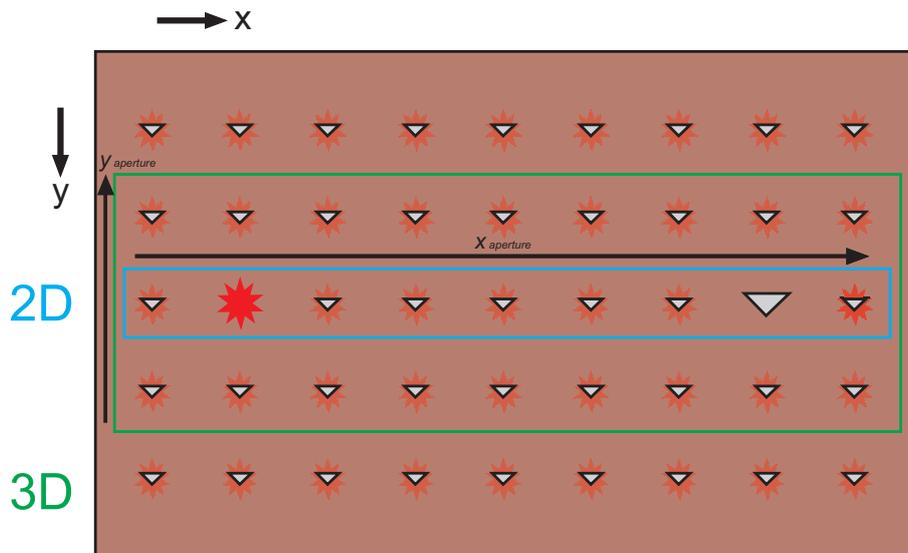


**Figure 5-1:** 3D-aperture depending on expected subsurface complexity.

It is however currently not known how to predict the necessary apertures before-hand and deriving a relationship between reflector-depth, -dip and apex-location would be helpful. In order to reduce the computation time for 3D SRME, the analysis is often only done for a few grid-points, and interpolated in-between (Verschuur 2015, personal communication).

## 5-3 Wavelet-distortion due to convolution and correlation

Section 2-3 illustrated how a wavelet changes in width after convolving it several times with other band-limited events. Normally this can be corrected for by adaptive subtraction, by correcting the phase and restoring the band-width. However, as events are seldom isolated in the seismic section, broadening the wavelets can cause their side-lobes to interfere with each other. This is illustrated in figure 5-2, showing a trace with four seismic events, from actual field data. After convolution with the time-reversed wavelet of a primary, acting as the multiple generating boundary, and an additional convolution with one single primary to simulate the 3 different multiples related these events, figure 5-3 is obtained. This rudimentary simulates the approach used in Jakubowicz-IME.

As is visible, the already close events are now interfering with each other, and are no longer distinguishable. It is important to note that the side-lobes of a wavelet are part of the
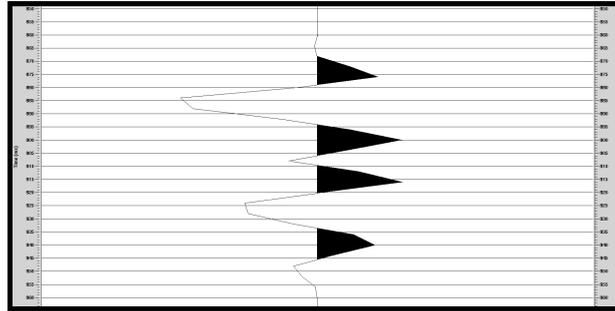
**Figure 5-2:** Four events on a seismic trace, representing primary reflections.
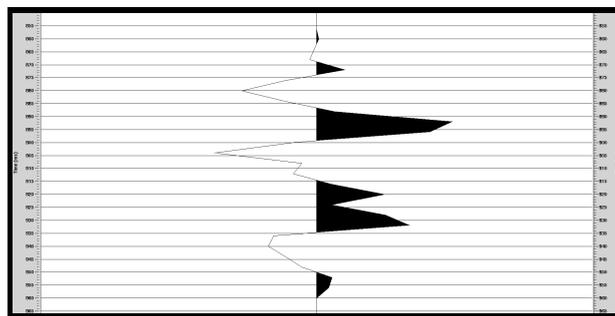


**Figure 5-3:** Four events after correlation and convolution with one single event.

amplitude information. However, due to the strong interference, it may happen that only the main-lobes are fitted to the seismic data during the adaptive subtraction, and therefore only part of the amplitude information used. Omitting source and receiver-deconvolution, can therefore potentially lead to amplitude errors.

## 5-4   Large offset muting

For IME purposes, the wavefields involved in the multiple-prediction process need to be muted from reflections related to the chosen level and from reflectors above. At large offsets, for high velocity contrasts and for close interfaces, it can happen that the recorded reflections interfere. This is shown in figure 5-4, but an even better example can be seen in [Berkhout and Verschuur, 2005a]. Incorrect muting causes either the creation of artifacts or no multiple-prediction at the larger offsets. The proposed approaches for dealing with such a case are either least-squares subtraction of the focusing operator, in the case of CFP-algorithms, or the double back-propagation to in-between the reflections, allowing for easier muting before forward propagating the wavefield to the surface again [Verschuur, 2006a].
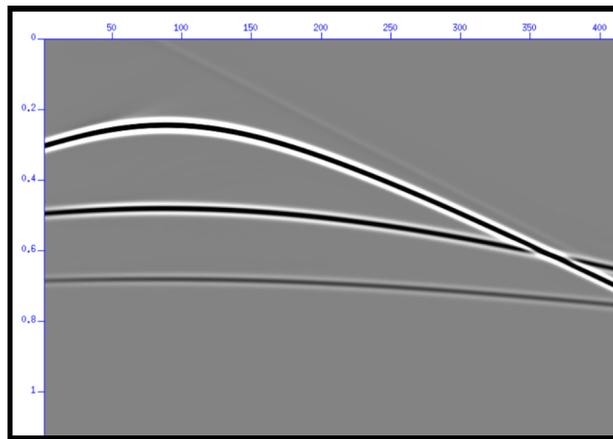
**Figure 5-4:** Shot gather showing two primary events which start to interfere with each other at larger offsets.

## 5-5    Identification of multiple generating boundaries

As multiples have a different spatial behavior than their corresponding primaries, it is often easy to spot them, especially when the dips are conflicting. Using their distinct resemblance with other reflections, and by judging their periodicity, it is often possible to correctly guess the multiple generating boundary. However, this is not the case for horizontally layered media. For such a subsurface, it is quite hard to distinguish multiples from primaries, and then to identify their multiple generating boundary. Especially when many strong reflectors are present in the subsurface it is hardly possible to say where multiples in a specific target-area originate from. This is also because all the different internal multiples combine into a dispersive curtain [Verschuur, 2006a].

Such a case is illustrated in figure 5-7, showing data from the Parisian onshore basin (courtesy of Vermillion Energy). Using well-logs, a wavelet derived from the nearby seismic section, and an elastic modeling software, the synthetic seismic section containing only primaries, and the section containing primaries with their corresponding multiples, have been modeled (figure 5-5). The stacked seismic section in the vicinity of the well, zoomed into an area of interest, is displayed next to them, in figure 5-6. It is visible that the modeled seismic sections, fit the recorded data pretty well, apart from a multiple at 1350 ms, which is only faintly visible. This indicates that at least one multiple generating boundary, causing such a multiple, is not present in the well-logs, and must therefore be earlier than 850 ms. Note that the event at 1480 ms shows a reflection event created by the interference of a primary with a multiple.

Using the already mentioned 1D-CFP-layer-IME algorithm the multiples are also predicted using the recorded wavefields, instead of well-log-information. However, it is not apparent where to put the required level, as it is not possible to identify the multiple generating boundaries causing the multiples in the area of interest. As the seismic section contains many stronger reflection events, it is expected that several levels are needed, to predict all the interfering multiples. One possible approach would be to separate the seismic section into packages of strong reflectors and to chose one layer, lying between such packages, in the shallower area in order to predict multiples with longer periodicity, and to chose a second level, closer to the area of interest, also between stronger reflectors, to predict multiples of shorter period. Following this approach one level, at 930 ms, and one at 1220 ms, was chosen (see figure 5-7). The multiple predictions before adaptive subtraction are presented in figure 5-8 and 5-9.

As expected both levels predict different multiples that arrive at similar times, such as at 1450 ms, but also multiples at different times. The 930 ms level seems to be chosen shallow enough to predict the strong multiple at 1350 ms, while the 1220 ms level better predicts the faintly interfering multiples between 1300 ms and 1430 ms. Combining these two levels could potentially achieve a better result.
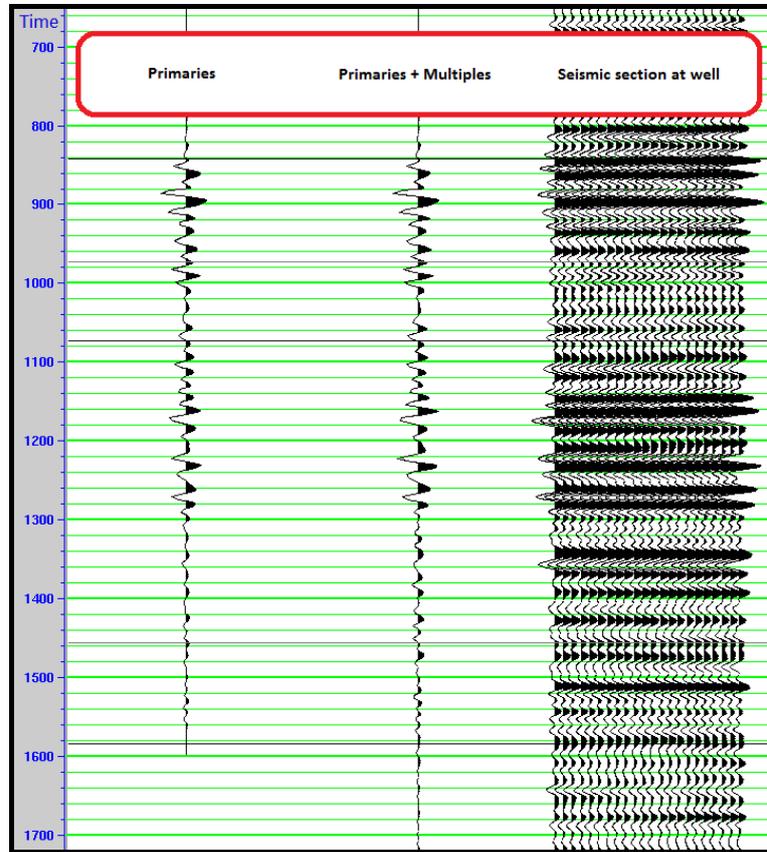
**Figure 5-5:** Modeled primaries and multiples next to the seismic section.
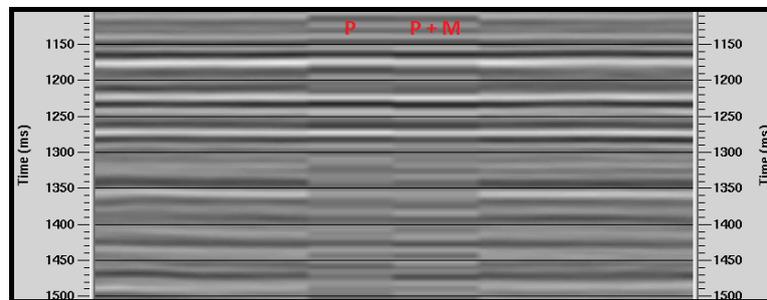


**Figure 5-6:** Stacked seismic section before multiple removal in the vicinity of the modeled primaries and multiples.
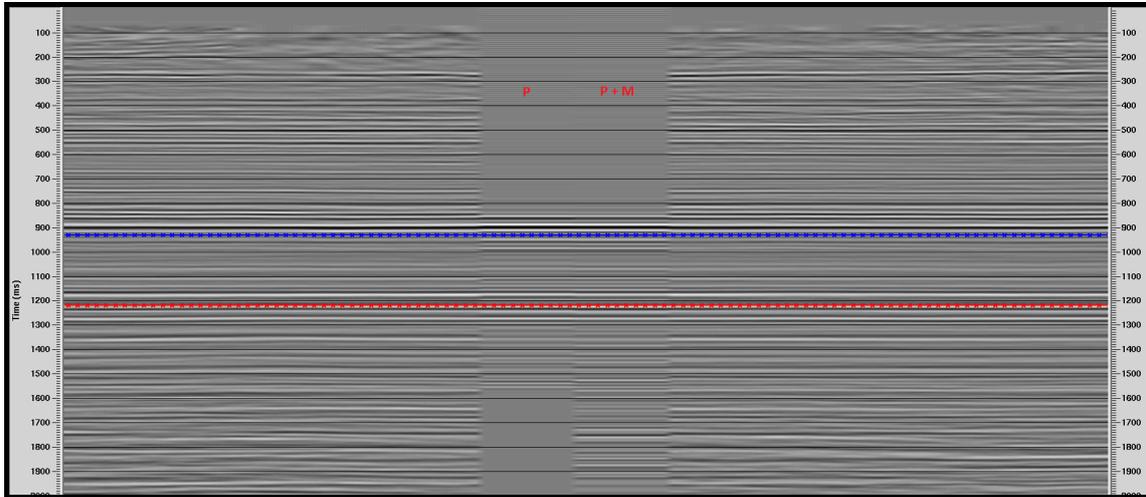
**Figure 5-7:** Larger part of the seismic section, with the chosen IME-layers indicated in color (930 ms in blue, and 1220 ms in red). Note the inter-bedded modeled synthetics.
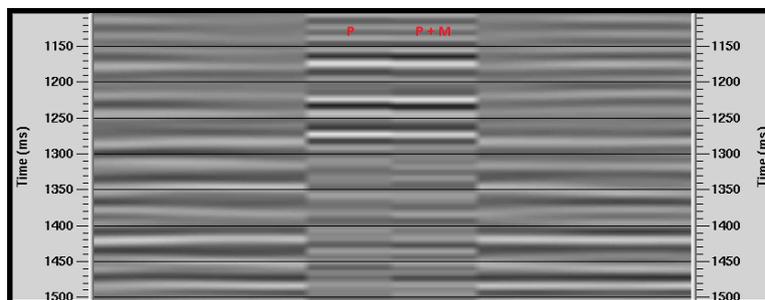


**Figure 5-8:** 1D-IME multiple prediction using a level at 930 ms.
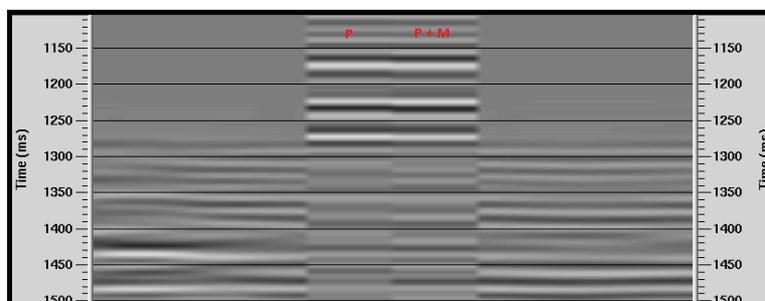


**Figure 5-9:** 1D-IME multiple prediction using a level at 1220 ms.

## 5-6   Adaptive Subtraction

All the different error sources discussed in this thesis, make adaptive subtraction a requirement, not only for IME, but also for other wave-equation based multiple prediction methods. In the case of Jakubowicz-IME it has to correct for the general amplitude errors, for potentially small arrival-time errors, and for the altered wavelet-amplitude spectrums, due to the neglected source- and receiver-deconvolutions.

Many different subtraction methods exist, such as, subtraction based on pattern recognition [Spitz, 1999], multi-gather least-squares subtraction , least mean squared error (LMS) adaptive [Dragoset, 1995], and standard least-squares subtraction on single gathers [Verschuur and Berkhout, 1997b]. Additionally, adaptive subtraction can be applied in many different domains, such as in the common offset domain, in the linear Radon domain, in the common offset vector domain [Retailleau et al., 2012], or simply using shot-gathers. Each method has its distinctive advantages and potential. For example, choosing a common-offset or common-angle domain, allows the filters to better honor the angle-dependency of the source and receiver signature [Verschuur, 2006b].

However, the focus of this chapter will be on the most commonly used subtraction method, applied on shot gathers, based on least-squares filtering. Standard least-squares filtering aims at minimizing the sum of the squares of the difference between two signals [Levinson, 1947]. This objective-function restriction is called the L2-norm.  A good approach is to apply adaptive subtraction using a two-step process [Verschuur and Berkhout, 1997b]. For the first step, the complete data is used in a global window for the derivation of the adaptive filter, aimed to act as a deconvolution filter, taking care of the overall source signature. It should be noted though that this deconvolution filter is also influenced by other error-sources. The second step consists of using smaller, overlapping, local windows, in time and space, taking care of the small deviations caused by noise, ghosts and the angle-dependency of the source-signature. Tapering and blending together the adaptively modified multiple predictions of each individual window gives the total predicted wavefield used for the subtraction.

Several potential complications are associated with this approach. In the presence of relative amplitude errors, using larger local windows, causes adaptive subtraction to always be a compromise, and to leave residual multiple energy. In a perfect world with no noise and no primary-multiple interference, choosing a very small local window, preferably only in time, could potentially lead to perfect subtraction as the filters can correctly account for the angle-dependent deconvolution. As this is not the case though, we have to rely on larger windows to give us more robust and stable results. But even then, the minimum energy assumption is not valid for multiples interfering with primaries, or even simply in the presence of strong primaries near weak multiples. Applying the L2-norm will then potentially remove primary energy.

A different approach, would be to use the L1-norm for the objective function. It evaluates the sum of the absolute values of the difference between two signals instead of the sum of the squares.  This makes it potentially more suited for the aforementioned cases and

better preserves the primary amplitudes. However, as least-squares filtering using the L1-norm requires a non-linear inversion procedure, typically implemented as a conjugate gradient method [Guitton and Verschuur, 2004], its computation time is around 10 times slower than the L2-norm application. For field data, the L2-norm least squares subtraction is typically still considered to be the most robust method [Abma et al., 2005].

These aspects of adaptive subtraction are investigated for the field dataset from the Parisian basin, using the least-squares subtraction-algorithms implemented by the DELPHI-consortium. A small local window size of 80 ms and 5 traces (figure 5-10), and a larger one using 400 ms and 30 traces (figure 5-11) were tested for the L2-norm application. The results after subtraction are evaluated for the same zoomed-in area of interest as in the previous chapter, using the modeled synthetics as a reference.
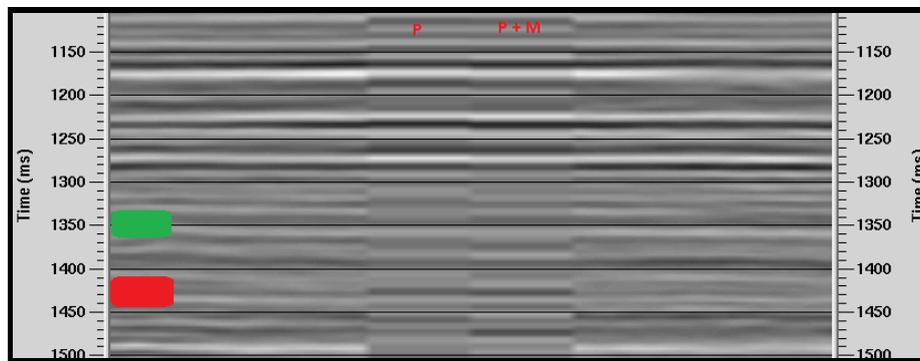


**Figure 5-10:** Adaptive subtraction result for a local window of 80 ms and 5 traces (level 930) using the L2-norm.
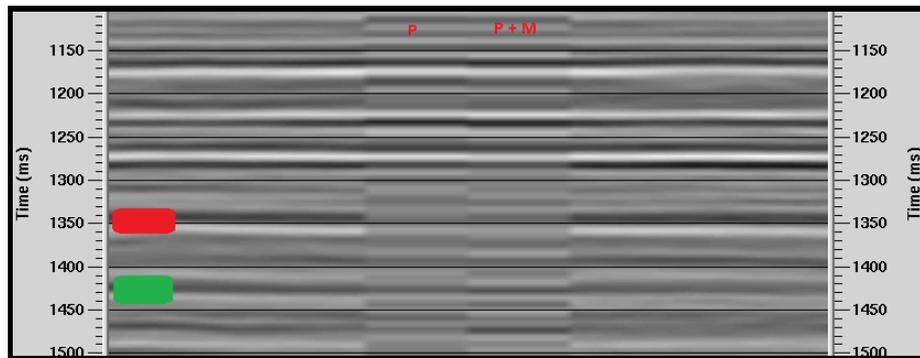


**Figure 5-11:** Adaptive subtraction result for a local window of 400ms and 30 traces (level 930) using the L2-norm.

Indicated in green, the multiple event at 1350 ms (not present in the synthetics), has been correctly removed using the smaller window size. However, indicated in red, also part of the primary energy of the primary-multiple interference at around 1430 ms has been removed. The opposite can be seen for the larger window size, where the multiple at 1350 ms is still rather strong after the subtraction (indicated in red), but where the multiple at 1430 ms has been reduced without harming the primary event (indicated in green). This matches the previous description that small window sizes aggressively reduce even non-multiple energy, while larger window-sizes only result in a compromise.

Using the smaller local window size, the adaptive subtraction is repeated with the L1-norm. The result is shown in figure 5-12, where it is indicated in green, that the L1-norm leaves slightly more energy of the primary at 1350 ms, but where the lower part of the primary event at around 1180 ms has been altered negatively (indicated in red).
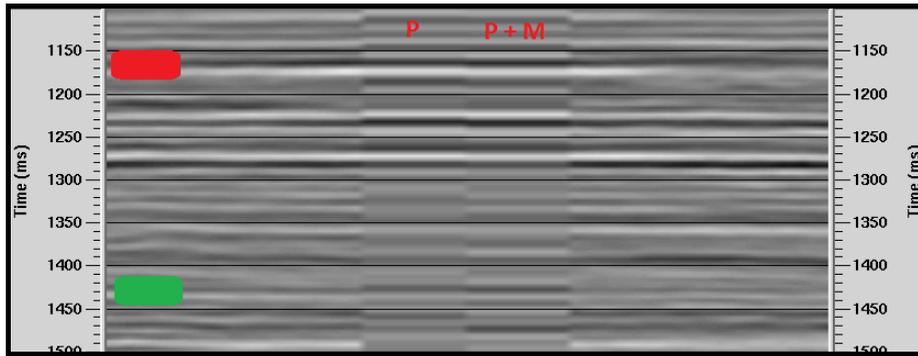


**Figure 5-12:** Adaptive subtraction result for a local window of 80 ms and 5 traces (level 930) using the L1-norm.

## 5-7 Iterative IME: approximation of multiple free data

Predicting the internal multiples related to one specific boundary requires receiver gathers which are free of them. This paradoxical situation is solved by reformulating the multiple prediction process in an iterative fashion. Gathers still containing all internal multiples related to the specific layer, are used for the first iteration, assuming that after adaptive subtraction the result contains less multiples and is therefore suited as input for the next iteration. However, violating the multiple-free assumption obviously produces errors.

In order to investigate these errors, it is important to understand which events predict which multiples in Jakubowicz-IME. First-order internal multiples are predicted by the convolution and correlation of primary events. Therefore, they are not affected, and predicted correctly. Second-order internal multiples are generated by the combination of a first-order internal multiple with two primary reflections. One of these is the multiple generating boundary, used for the correlation. The same process is repeated for each higher order.

This is illustrated using a simple subsurface model containing only two primaries, displayed in figure 5-13. On the left the stacked section, containing only the primaries, is shown, while on the right the corresponding multiples are included. Note that each higher-order multiple exhibits an increasing dip. The prediction process is illustrated in figure 5-14 using colors assigned to each contributing event. The multiple generating boundary is always involved and therefore not colored. Utilizing intuitive subtractive color-mixing it is made visible which events are combined to create the individual multiples. The specific color-mixing rules are illustrated on the lower left.
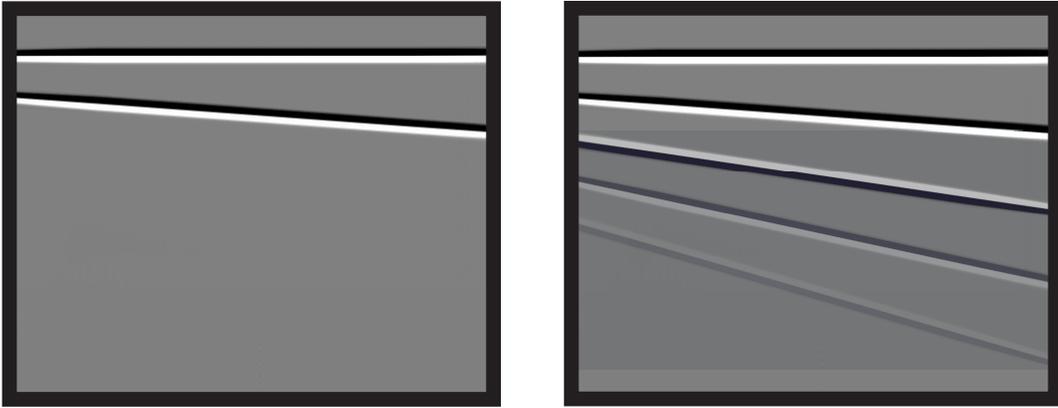
**Figure 5-13:** Stacked section of a subsurface model consisting of two primaries (left). The corresponding multiples are included on the right.
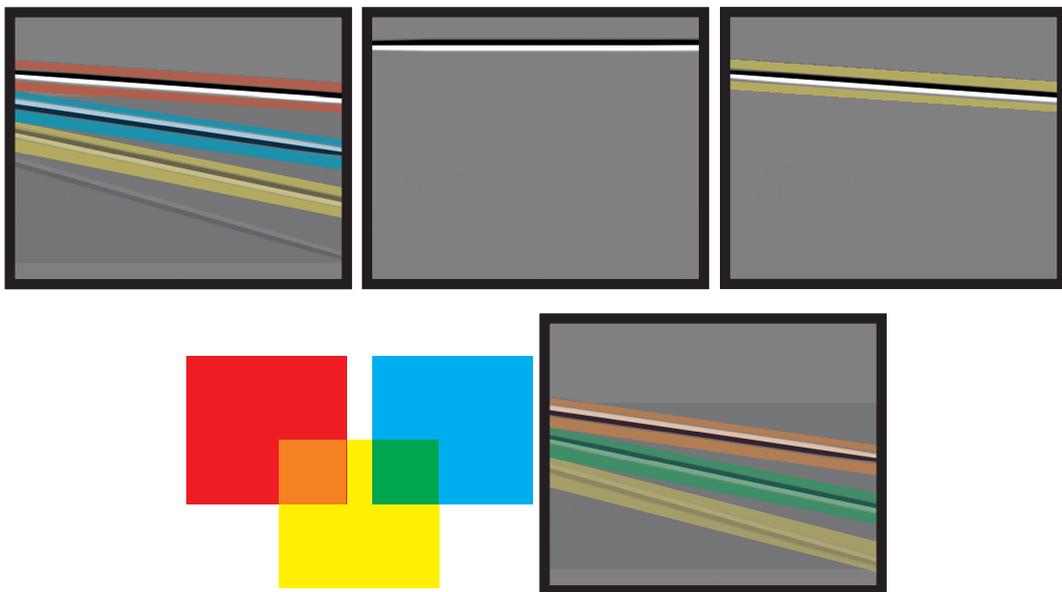


**Figure 5-14:** Schematic illustration showing which events contribute to the predicted multiples in the optimal Jakubowicz-approach.

The same process is repeated for the iterative IME-approach, where both gathers contain multiples. This is illustrated in figure 5-15. Note that the color-mixing rules were adjusted to include the additive mixing of red (the primary reflection) with green (the second order multiples), producing the color yellow (the third order multiples). The same is achieved by combining the first order multiples (initially yellow). Contrary to the first example, this illustrates that the third order multiple is not only predicted by the combination of the primary in the receiver gather with the second order multiple in the shot gather, but also by the combination of the first order multiples, or by combining the second order multiple in the receiver gather with the primary in the shot gather. Therefore, the third order multiple is predicted three times. In a similar way the second order multiple is predicted twice, as is illustrated in figures 5-17 and 5-16 in terms of involved ray-paths.
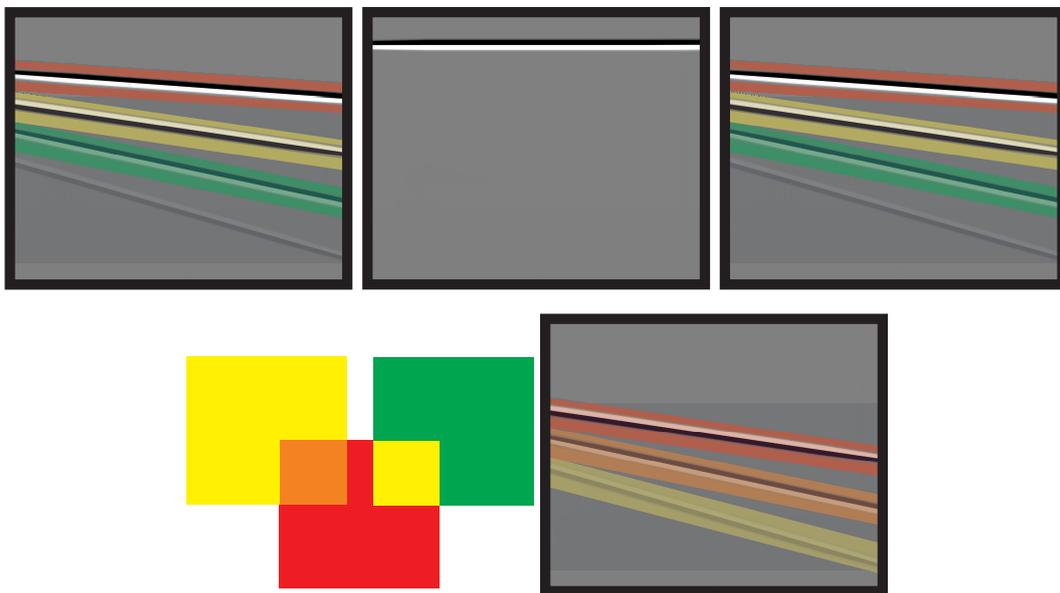


**Figure 5-15:** Schematic illustration showing which events contribute to the predicted multiples in the adjusted first iteration of the Jakubowicz-approach.

In conclusion, the first iteration results in higher-order multiples being predicted with too high amplitudes. With each iteration they converge closer to their correct value. The resulting relative amplitude error between the different multiple orders potentially leads to a compromise during the adaptive subtraction, leaving residual first-order multiple energy, even though they were predicted correctly.

The derived equations for multiple prediction also require that no multiples related to interfaces above the chosen boundary are remaining in the involved wavefields. As was pointed out, multiples of a given order generate multiples of the next-higher order. Multiples, non-related to the chosen boundary, are no exception to that rule. However, they generate a family of higher-order multiples, related to shallower reflection interfaces. Such a case is illustrated in figure 5-18. While this may be a nice addition if IME is only supposed to be done for one level, it will lead to problems if IME is repeated for several levels, as some multiples are potentially predicted twice. This is avoided by applying IME consecutively in a top-down approach.
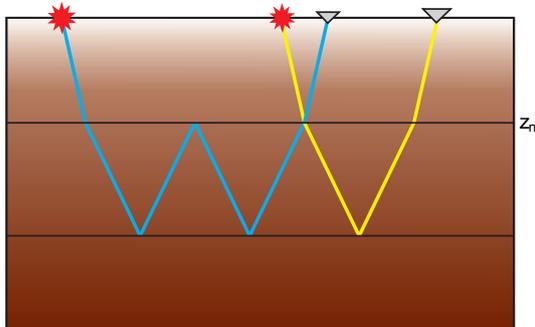
**Figure 5-16:** Physical ray-path illustration of a first-order multiple in the common-shot-gather contributing to the generation of a second-order multiple.
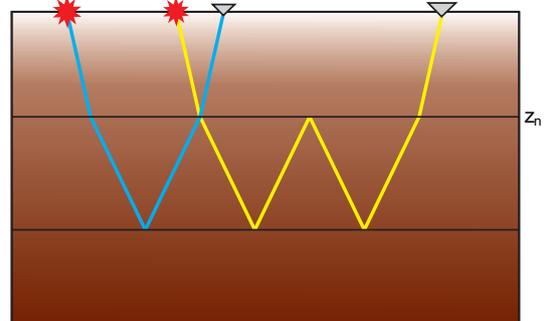


**Figure 5-17:** Physical ray-path illustration of a first-order multiple in the common-receiver-gather contributing to the generation of the same second-order multiple.
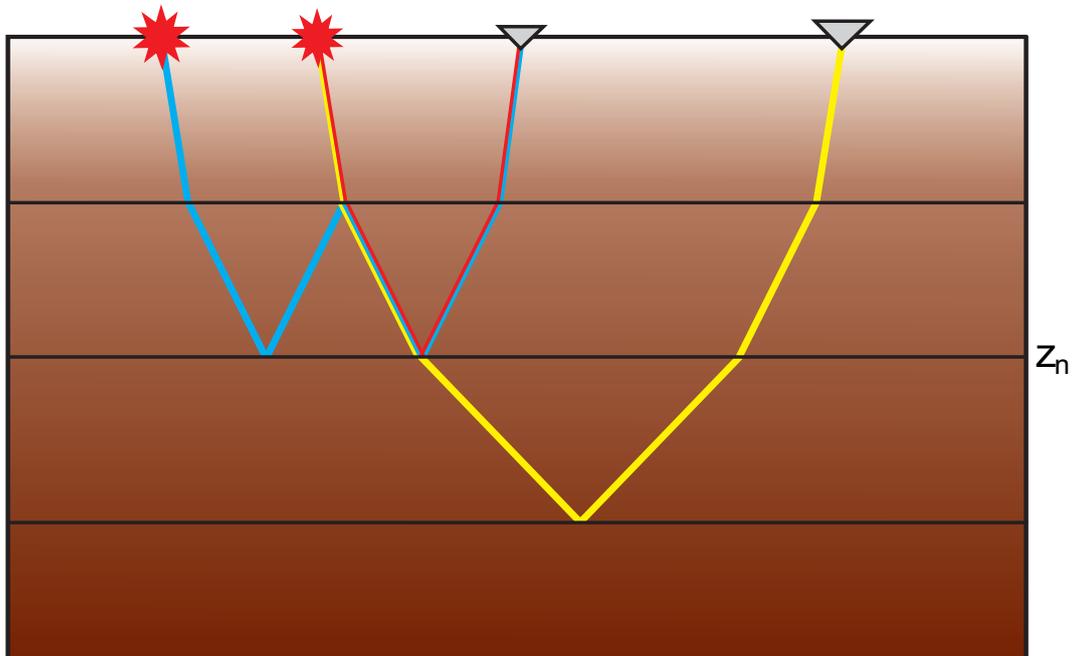


**Figure 5-18:** Physical ray-paths of events generating a multiple related to boundary $z < z_n$, as a result of not applying IME in a top-down approach.

## 5-8   Noise

Probably the biggest difference between synthetic data, generated for the purpose of showing that a method theoretically works, and actual field data, is the presence of noise. As Jakubowicz-IME approximates the earth impulse response using recorded surface data, it requires reflections events with strong amplitudes relative to the background noise. The characteristic of a data-driven prediction method is that noisy input-data will also result in a noisy prediction.

Different forms of noise can be seen in seismic recordings. For the purpose of this analysis it is enough to divide them into two categories; continuous background noise, and stronger amplitude bursts. Generally, noise adds ambiguity during both steps, the prediction and the subtraction, as all three wavefields used for convolution, and the traces from which the prediction is subtracted, contain noise [Van Groenestijn, 2010]. This is implicitly expressed, omitting the deconvolution terms and muting annotation, as

$$\{P\}_n = (P + N) - (-(P_d + N)(P + N)^H(P_s + N)) \quad . \tag{5-1}$$

Background noise has generally low amplitudes, and will remain as such after the convolution process. It will normally not interfere constructively during the summations, as it is supposedly random, although this would theoretically be possible. The effect of background noise is more relevant during the adaptive subtraction process, as it masks the local wavelet properties. Additionally, as the energy of the predicted multiple events is more spread out, due to the two wavelet-convolutions, it can happen that the weaker side-lobes are also masked by noise, potentially losing their amplitude-information, if the amplitude spectrum is not recovered.

Sudden amplitude spikes, on the other hand, have a more pronounced effect on the prediction. A single spike on one trace leaves an imprint on all convolution-results it is involved with. And as the summation process does not involve any normalization, the spike will, first, transfer artifacts to the multiple contribution gather, and then the predicted multiple trace. However, as there are no other coherent spikes adding constructively, in a similar way as the near-apex traces of curved events, the resulting amplitude is normally lower than any predicted multiple. Laterally coherent noise though, can produce similarly strong artifacts [Van Groenestijn, 2010].

The effect of noise on the multiple prediction has been tested by adding background noise and three larger noise-spikes to each trace of the synthetic dataset. Figure 5-19 shows the original shot-gather, while figure 5-20 shows the same shot gather with the added noise. After applying the 3D Jakubowicz-algorithm to predict the internal multiples related to the shallowest boundary, it appears that the added noise was not enough to create larger prediction errors. While the background noise can definitely be seen (figure 5-22), the individual multiples are still correctly reproduced (figure 5-21). This is probably also ture, because the added background noise is perfectly random, and only single spikes were added, instead of coherent noise events such as ground-roll. In practice, noise is never completely random and will not

interfere destructively to the same extent. However, the weaker second-order multiples, are no longer resolved so clearly and therefore affected by the noise. Note that the noise-free synthetic result still contains the weak artifacts from the primary reflections, as mentioned in section 4-2-1.
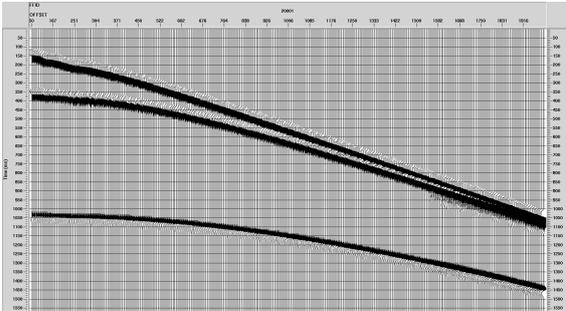


**Figure 5-19:** Synthetic shot-gather containing no noise.
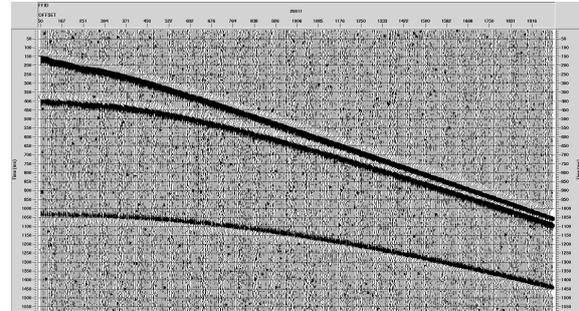
**Figure 5-20:** Synthetic shot-gather with added background noise and three spiky noise events per trace.
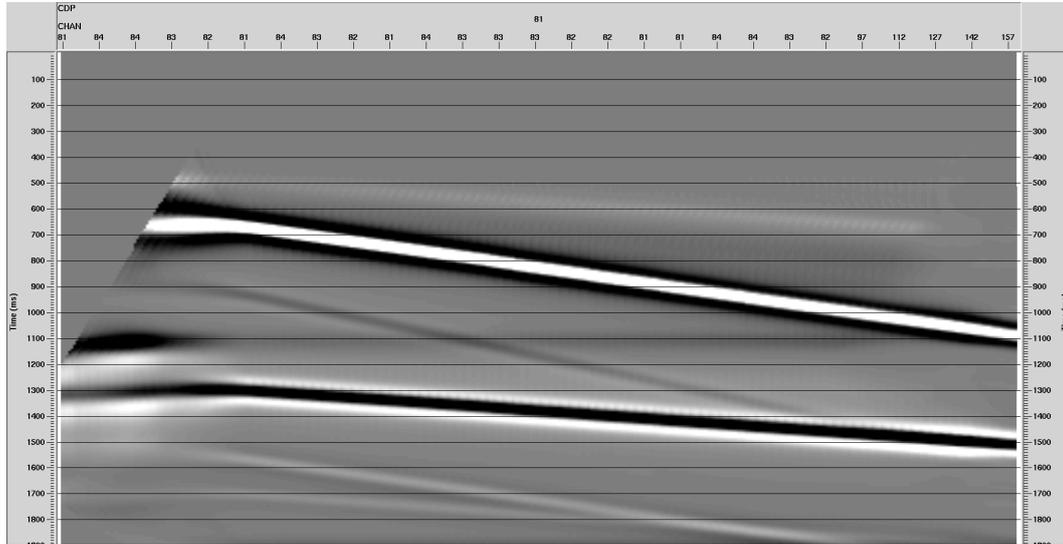
**Figure 5-21:** Multiple prediction after 3D-IME on synthetic data containing no noise.
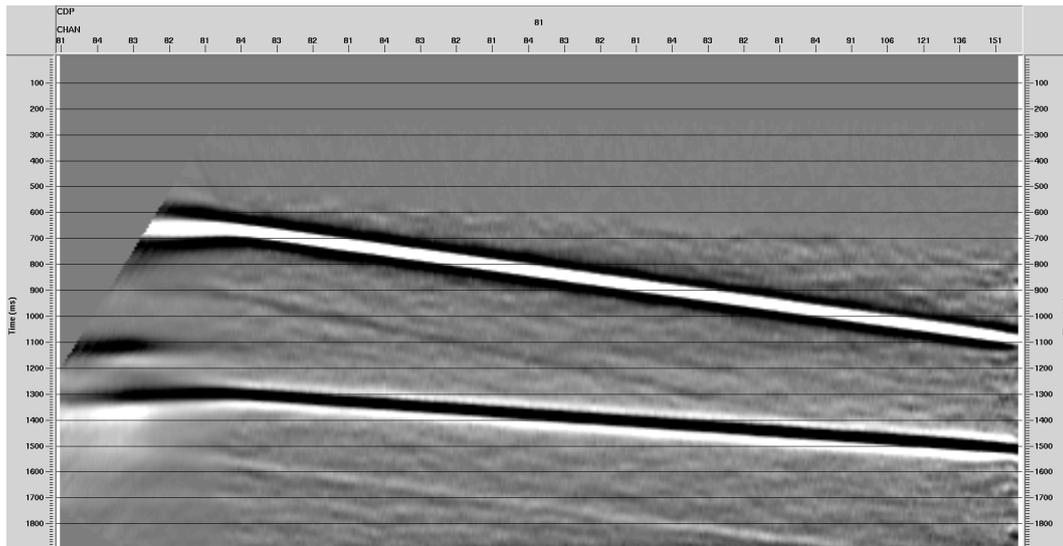


**Figure 5-22:** Multiple prediction after 3D-IME on synthetic data with added noise.

## 5-9    Spatial sampling

The biggest advantage of data-driven multiple elimination methods is that no subsurface information is required, because the recorded wavefields are used as the multiple prediction operators. However, this also means that these wavefields have to be appropriately sampled. This is especially a problem for 3D-applications, as sampling in the cross-line direction is normally a lot coarser than in the in-line direction.

The theory for Jakubowicz-IME implies the presence of regularly sampled grid-points on the surface. On such a grid-point, a source and a receiver have to coincide. However, often the shot-stations are sampled coarser than the receiver-stations. This means that no contribution from a grid-point containing only a receiver is retrieved. Such an example is schematically depicted in figures 5-23 and 5-24. The implication of these illustrations is that several errors arise from in-appropriately sampled surface-grid-points:

- Too coarse sampling of surface-grid-points can lead to spatial aliasing of the predicted multiple events resulting in phase and amplitude errors.

- Artifacts are created in the presence of larger sampling gaps, where the individual wavelets, on the slopes of the curved events, do not cancel each other. This results in amplitude errors as the adaptive subtraction tries to account for these artifacts. Such larger gaps commonly occur in the following instances:

    - At the aperture-edges
    - At near-offset gaps
    - In the cross-line direction

- If the apex of a curved event in the multiple contribution gather is not included, the corresponding multiple event is predicted with a later arrival time.

Interestingly, as the Jakubowicz-approach involves a double summation process, these errors can occur twice; during the prediction of the multiple contribution gather traces, and subsequently during the prediction of the multiple trace.

### 5-9-1    The importance of near offset traces

Near-offset traces are of special importance for the correct estimation and attenuation of multiples [Dragoset and Jericevic, 1998]. Figure 5-25 illustrates that for a subsurface containing only horizontally layered interfaces, the primaries needed for the multiple-prediction of the source-receiver-pair with the smallest offset, have an even smaller offset. Generally, this also extends to moderately complex subsurfaces. Multiple-predictions of near-offset traces are more important for the stacked image than those of far-offsets. This is because, after NMO-correction, multiples are normally not completely flat but exhibit a residual move-out (indicated in green for a NMO-corrected shot-gather, using a velocity model for primaries). Note that the illustration only holds true for a 1D-subsurface, where shot-gathers and CMP-gathers can be interchanged. Stacking of only partially corrected multiples in a CMP-gather
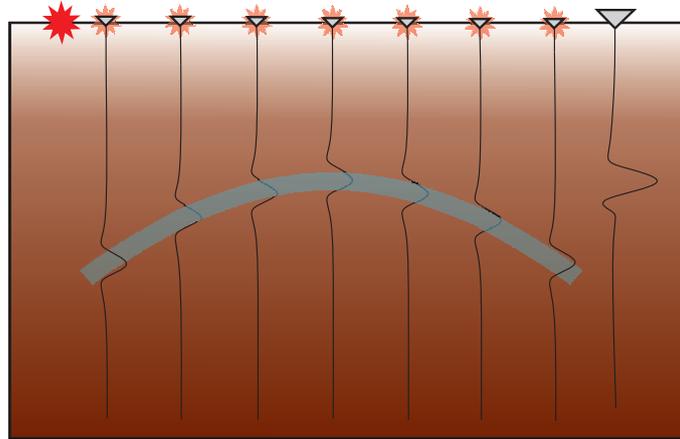
**Figure 5-23:** Generation of the multiple prediction from multiple-contribution-gather-traces based on densely sampled surface-grid-points with coinciding source- and receiver-stations. On the right hand side, the summation of all contributions is depicted, showing only a contribution from the apex area.
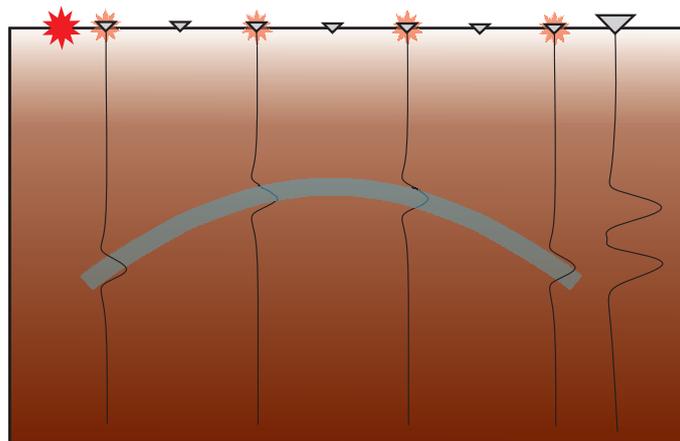


**Figure 5-24:** Generation of the multiple prediction from multiple-contribution-gather-traces based on coarsely sampled surface-grid-points. Only every second surface-grid-point has coinciding source- and receiver-stations, resulting in aliasing artifacts in the summation result, shown on the right.

leaves mostly only the contribution of the near-offset traces. So, any residual multiple energy in the near-offset traces, will deteriorate the stacked section, while residual multiple energy at far-offsets is more likely to not add up constructively.
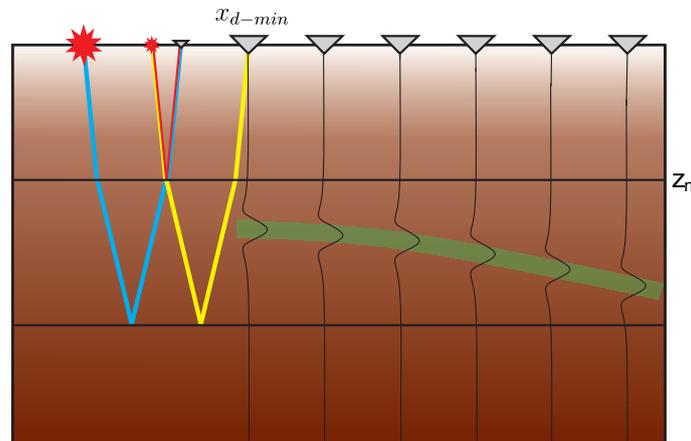


**Figure 5-25:** Physical ray-paths involved in the prediction of a near-offset multiple-trace. The resulting multiple-predictions are NMO-corrected using a primary-velocity-field. This shows that only the near-offset contributions add up constructively when stacked.

### 5-9-2   The importance of negative offset traces

For complex subsurfaces it is potentially necessary to use a grid-point aperture extending beyond the initial source-location. Figure 5-26 illustrates such a case, in the presence of a steeply dipping reflector. This means that the input-shot-gather also needs to contain negative offset-traces. Marine acquisition does not generate negative offsets, and therefore they have to be created either by extrapolation, or by applying reciprocity.

### 5-9-3   Interpolation of missing traces

As it is often not possible, or commercially viable to sample the wavefield on a regularized and dense spatial grid, we have to rely on the interpolation of the missing traces. This is a vital step for the successful application of 3D-IME implementations. The goal is to create traces with events that have the correct arrival-time, amplitude and phase-content. Often, the traces that are the closest to the sought after location, will give the best interpolation result.

Many different approaches exist, such as, partial NMO-correction, back and forth-transformation into the parabolic radon-domain [Kabir and Verschuur, 1995], re-mapping of recorded traces based on reciprocity, and interpolation by applying DMO-corrections [Baumstein et al., 2005]. However, not always do these approaches correctly account for the angle-dependency or the local dip of the recorded events. For example, in the case of a complex subsurface, the near-offset events are not exactly parabolic, and transformation into the parabolic radon-domain will not give the expected result.
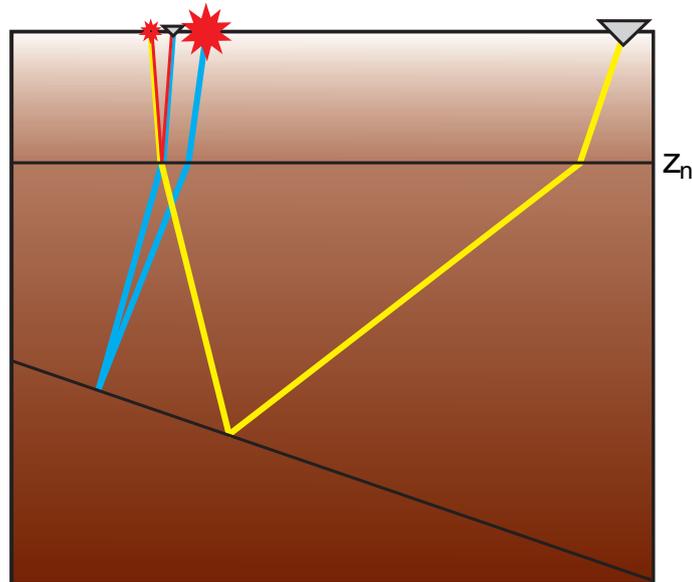
**Figure 5-26:** Physical primary ray-paths involved in the prediction of the multiple ray-path, extending to negative offsets in the presence of a complex subsurface.

A different approach to the sampling problem would be to delay the interpolation-process till the cross-line summation step. As proposed, the summation is then replaced with a sparse inversion scheme [Dedem and Verschuur, 2005]. This is done after mapping the events in the cross-line direction of the multiple contribution gathers into a model space using a high-resolution hyperbolic Radon transform.

However, such dedicated approaches usually work in the case the sampling is dense in one direction and coarse in the other, like the typical narrow azimuth streamer configurations. Nowadays, there is a trend towards multi-azimuth, wide azimuth and full azimuth acquisition, often with multiple source boats. This creates acquisition patterns without a specific dense and a coarse direction. This means that for each predicted multiple trace, a dedicated set of input traces need to be selected or created from the available data. Therefore, current interpolation techniques for SRME and IME are based on on-the-fly interpolation of required data traces from the existing ones (see [Dragoset et al., 2010] for 3D SRME and [Hung et al., 2013] for 3D IME).

# 6

# Conclusions

Multiple elimination methods have progressed considerably over the last 25 years. Back then, the newly conceptualized 2D-SRME-approach was at first believed to not be able to produce good results on field data, and even if so, too computationally expensive to be of importance for the industry. Today, most competitive processing-companies have 3D-SRME standardly offered, and are even starting to develop simplified implementations of 3D-IME. I firmly believe that hardware advancements will make 3D IME a standard commodity in the next decade. Therefore, I come to the conclusion that computational cost should be less of a concern for currently researched methods and not be regarded as a restriction. In the same sense, I also think that the other discussed limitations have differently strong impacts on the prediction accuracy, or can already be handeled to a different degree.

As the detailed and correct theory for the three dimensional implementation of IME exists and can already be applied for smaller data-sets, the errors of neglecting the subsurface-complexity become less relevant. Also the other theoretical error-source, incorrectly implemented transmission operators, is not much of an issue, as it affects all multiples related to one boundary in a similar way. Regarding the practical difficulties; with enough awareness, care and proper effort, most problems regarding muting, amplitude balancing, the unknown multiple generating boundary, and the complications leading to the iterative and top-down implementation, can be eliminated.

On the other hand, noise and the issue of appropriate spatial sampling, are a bigger concern, as we cannot completely remove the effect they have on the prediction result, without a disproportionate amount of work for acquisition. This would be quite beneficial though, as IME is a data-driven approach, and relies on well-sampled, clean recordings, acting as prediction operators. This can be achieved in two ways; during acquisition and post-acquisition, during processing. Commonly, seismic acquisition is a lot more expensive than processing. Therefore, it is preferred to only apply as much effort in acquisition as is necessary to allow for correct processing. So, any methods improving upon noise-attenuation, to increase the S/N-ratio, and trace-interpolation, would benefit IME. An interesting aspect

between the relation of commercial interests and theoretical advancement, is that probably any progress achieved on the side of noise-attenuation and trace interpolation is met with cost-cutting efforts on the acquisition-side.

The other very important factor, which would help to increase the accuracy of the adaptive subtracting result, is the correct estimation of the source-properties. Not only in a global manner, by retrieving one wavelet for the whole survey, but in all its angle-dependency. This would first of all alleviate the previously discussed problems caused by wavelet-interference, and secondly allow the adaptive subtraction to focus on other amplitude spectrum errors, such as the ones caused by the additional attenuation. Interestingly, SRME is regarded as one way to retrieve the deconvolution filters during the adaptive subtraction process, and might serve as an estimate for the following IME-application, as it is subject to less potential errors.

These three aspects of noise, spatial sampling, and source-properties, are also very important error-sources for Marchenko imaging, and in fact any method based on wavefield convolution. Therefore, any advancements made in these individual areas of research would improve upon these methods. For example, the development of accurate interpolation algorithms for complex subsurfaces, honoring the angle-dependency of the phase spectrum and amplitude, would allow to forgo near-offset sampling, while still producing correct arrival times and amplitudes.

But of course, avoiding any errors is only the first step. The second step is to account for inevitable errors by applying adaptive subtraction. I was astonished by how much research already went into something so seemingly simple as subtracting one amplitude-value from another. Despite all the efforts though, it seems as if the general approach is still standard least-squares subtraction using the L2-norm. More sophisticated approaches, potentially differentiating between the needed adjustments and managing to reformulate the wavelet-reshaping into a step-wise correction process, would be helpful as it is easier to estimate the contribution from one error source, than the interference of many errors. However, with improved 3D predictions, adaptive subtraction may also achieve better results and be simpler to implement, such as only requiring one large window.

The future of seismics may also be the to apply IME as a pre-processing step for full-waveform inversion. This would allow to take all multiples into account in the forward modeling, and also reduce the required sampling-effort.

# Bibliography

[Abma et al., 2005] Abma, R., Kabir, N., Matson, K. H., Michell, S., Shaw, S. A., and McLain, B. (2005). Comparisons of adaptive subtraction methods for multiple attenuation. *The Leading Edge*, 24(3):277–280.

[Alá'i et al., 2003] Alá'i, R., Verschuur, D. J., Reagan, R., et al. (2003). Land data prestack surface multiple elimination strategy, case study in north africa. In *2003 SEG Annual Meeting*. Society of Exploration Geophysicists.

[Alá'i and Verschuur, 2006] Alá'i, R. and Verschuur, E. (2006). Case study of surface-related and internal multiple elimination on land data. In *2006 SEG Annual Meeting*.

[Oppenheim, 1983] A.V. Oppenheim, A.S.Willsky, I. Y. (1983). *Signals and Systems*. Prentice Hall.

[Baumstein et al., 2005] Baumstein, A., Hadidi, M. T., Hinkley, D. L., and Ross, W. S. (2005). A practical procedure for application of 3D SRME to conventional marine data. *The Leading Edge*, 24(3):254–258.

[Berkhout, 1982] Berkhout, A. (1982). *Seismic Migration: Imaging of Acoustic Energy by Wave Field Extrapolation, Part A: Theoretical Aspects*. Elsevier.

[Berkhout and Verschuur, 2005a] Berkhout, A. and Verschuur, D. (2005a). Removal of internal multiples with the common-focus-point approach (CFP): Part 1 - explanation of the theory. *GEOPHYSICS*, 70(3):V45–V60.

[Berkhout and Verschuur, 2005b] Berkhout, A. and Verschuur, D. (2005b). Removal of internal multiples with the common-focus-point (CFP) approach: Part 1explanation of the theory. *Geophysics*, 70(3):V45–V60.

[Berkhout and Verschuur, 1994] Berkhout, A. and Verschuur, D. J. (1994). Multiple technology. part 2: migration of multiple reflections. Technical report, Society of Exploration Geophysicists, Tulsa, OK (United States).

[Berryhill and Kim, 1986] Berryhill, J. and Kim, Y. (1986). Deep-water peg legs and multiples: Emulation and suppression. *Geophysics*, 51(12):2177–2184.

[Broggini et al., 2012] Broggini, F., Snieder, R., and Wapenaar, K. (2012). Focusing the wavefield inside an unknown 1D medium: Beyond seismic interferometry. *Geophysics*, 77(5):A25–A28.

[Chapman, 2004] Chapman, C. (2004). *Fundamentals of Seismic Wave Propagation*. Cambridge.

[Coates and Weglein, 1996] Coates, R. and Weglein, A. (1996). Internal multiple attenuation using inverse scattering: Results from prestack 1 & 2D acoustic and elastic synthetics. In *1996 SEG Annual Meeting*. Society of Exploration Geophysicists.

[Wapenaar, 1989] C.P.A. Wapenaar, A. (1989). *Elastic Wave Field Extrapolation: Redatuming of Single- and Multi-Component Seismic Data*. Elsevier Science.

[Wapenaar and Berkhout, 1990] C.P.A. Wapenaar, P. Herrmann, D. V. and A.J.Berkhout (1990). Decomposition of multicomponent seismic data into primary p- and s-wave responses. *Geophysical Prospecting*, 38:633 – 661.

[Hampson, 1986] D. Hampson (1986). Inverse velocity stacking for multiple elimination. *Journal of the Canadian Society of Exploration Geophysics*, 22(1):44–55.

[Dudgeon and Mersereau, 1984] Dan E. Dudgeon, R. M. Mersereau (1984). *Multidimensional digital signal processing*. Prentice-Hall.

[Dedem and Verschuur, 2005] Dedem, E. v. and Verschuur, D. (2005). 3D surface-related multiple prediction: A sparse inversion approach. *Geophysics*, 70(3):V31–V43.

[Dragoset, 1995] Dragoset, B. (1995). Geophysical applications of adaptive-noise cancellation sp2. 7.

[Dragoset et al., 2010] Dragoset, B., Verschuur, E., Moore, I., and Bisley, R. (2010). A perspective on 3D surface-related multiple elimination. *Geophysics*, 75(5):75A245–75A261.

[Dragoset and Jericevic, 1998] Dragoset, W. H. and Jericevic, Z. (1998). Some remarks on surface multiple attenuation. *Geophysics*, 63(2):772–789.

[Guitton et al., 2002] Guitton, A. et al. (2002). Shot-profile migration of multiple reflections. In *Annual International Meeting, SEG, Expanded Abstracts*, volume 72, pages 1296–1299.

[Guitton and Verschuur, 2004] Guitton, A. and Verschuur, D. J. (2004). Adaptive subtraction of multiples using the L1-norm. *Geophysical Prospecting*, 52(1):27–38.

[Hung et al., 2013] Hung, B., Wang, M., and Griffiths, M. (2013). True-azimuth 3D internal multiple attenuation without subsurface information. In *75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*.

[IEA, 2013] IEA (2013). World energy outlook 2013. Technical report, International Energy Association.

[Jakubowicz, 1998] Jakubowicz, H. (1998). *Wave equation prediction and removal of interbed multiples*, chapter 402, pages 1527–1530. SEG.

[Kabir and Verschuur, 1995] Kabir, M. N. and Verschuur, D. (1995). Restoration of missing offsets by parabolic radon transform. *Geophysical Prospecting*, 43(3):347–368.

[Kelamis and Verschuur, 2000] Kelamis, P. G. and Verschuur, D. J. (2000). Surfacerelated multiple elimination on land seismic datastrategies via case studies. *GEOPHYSICS*, 65(3):719–734.

[Koefoed, 1981] Koefoed, O. (1981). Aspects of vertical seismic resolution. *Geophysical Prospecting*, 29(1):21–30.

[Levinson, 1947] Levinson, N. (1947). The wiener rms (root mean square) error criterion in filter design and prediction.

[Li and Hu, 2009] Li, X. and Hu, T.-Y. (2009). Surface-related multiple removal with inverse scattering series method. *Chinese Journal of Geophysics*, 52(3):716–724.

[Lu et al., 2013] Lu, S., Whitmore, N., LeGleut, H., and Long, A. (2013). 3D high-resolution imaging using separated wavefields. In *75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*.

[Meles et al., 2014] Meles, G. A., Löer, K., Ravasi, M., Curtis, A., and da Costa Filho, C. A. (2014). Internal multiple prediction and removal using Marchenko autofocusing and seismic interferometry. *Geophysics*, 80(1):A7–A11.

[Retailleau et al., 2012] M.G.Retailleau, N. Benjamin, A. P. et al. (2012). Advanced 3D land internal multiple modeling and subtraction, a WAZ oman case study. In *2012 SEG Annual Meeting*. Society of Exploration Geophysicists.

[Taner, Koehler and Sheriff, 1979] M.T.Taner, F. Koehler and Sheriff, R. (1979). Complex seismic trace analysis. *GEOPHYSICS*, 44(6):1041 – 1063.

[O'Doherty and Anstey, 1971] O'Doherty, R. and Anstey, N. (1971). Reflections on amplitudes. *Geophysical Prospecting*, 19(3):430 – 458.

[Pica et al., 2008] Pica, A., Delmas, L., et al. (2008). Wave equation based internal multiple modeling in 3D. In *2008 SEG Annual Meeting*. Society of Exploration Geophysicists.

[Pica et al., 2005] Pica, P. et al. (2005). 3D surface-related multiple modeling. *The Leading Edge*, 24(3):292 – 296.

[Ramírez et al., 2008] Ramírez, A. C., Weglein, A. B., et al. (2008). Inverse scattering internal multiple elimination: Leading-order and higher-order closed forms. In *2008 SEG Annual Meeting*. Society of Exploration Geophysicists.

[Robinson, 1954] Robinson, E. (1954). *Predictive decomposition of time series with application to seismic exploration*. PhD thesis, Massachusetts Institute of Technology.

[Ryu, 1982] Ryu, J. V. (1982). Decomposition (decom) approach applied to wave field analysis with seismic reflection records. *Geophysics*, 47(6):869–883.

[Spitz, 1999] Spitz, S. (1999). Pattern recognition, spatial predictability, and subtraction of multiple events. *The Leading Edge*, 18(1):55–58.

[Thorbecke, 1997] Thorbecke, J. (1997). *Common Focus Point Technology*. PhD thesis, TU Delft.

[van der Neut et al., 2015] van der Neut, J., Wapenaar, K., Thorbecke, J., Slob, E., and Vasconcelos, I. (2015). An illustration of adaptive Marchenko imaging. *The Leading Edge*, 34(7):818–822.

[van Groenestijn and Verschuur, 2009] van Groenestijn, G. and Verschuur, D. (2009). Estimating primaries by sparse inversion and application to near-offset data reconstruction. *Geophysics*, 74(3):A23–A28.

[Van Groenestijn, 2010] Van Groenestijn, G.-J. A. (2010). *Estimation of Primaries and Multiples by Sparse Inversion*. PhD thesis, TU Delft.

[Verschuur, 2006a] Verschuur, D. (2006a). *Seismic multiple removal techniques*. EAGE.

[Verschuur, 2006b] Verschuur, D. (2006b). Tracking of seismic events using a combined global and local optimization strategy. In *68th EAGE Conference & Exhibition*.

[Verschuur, 1991] Verschuur, D. J. (1991). *Surface-related multiple elimination, an inversion approach*. PhD thesis, TU Delft.

[Verschuur et al., 1992] Verschuur, D. J., Berkhout, A., and Wapenaar, C. (1992). Adaptive surface-related multiple elimination. *Geophysics*, 57(9):1166–1177.

[Verschuur and Berkhout, 1997a] Verschuur, D. and Berkhout, A. (1997a). Estimation of multiple scattering by iterative inversion, part I: Theoretical considerations. *Geophysics*, 62(5):1586–1595.

[Verschuur and Berkhout, 1997b] Verschuur, D. and Berkhout, A. (1997b). Estimation of multiple scattering by iterative inversion, part II: Practical aspects and examples. *Geophysics*, 62(5):1596–1611.

[Wapenaar et al., 1992] Wapenaar, C., Verschuur, D., and Herrmann, P. (1992). Amplitude preprocessing of single and multicomponent seismic data. *Geophysics*, 57(9):1178–1188.

[Wapenaar et al., 2014] Wapenaar, K., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E., and Snieder, R. (2014). Marchenko imaging. *Geophysics*, 79(3):WA39–WA57.

[Weglein et al., 1997] 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(6):1975–1989.

[Wiggins, 1988] Wiggins, J. W. (1988). Attenuation of complex water-bottom multiples by wave-equation-based prediction and subtraction. *Geophysics*, 53(12):1527–1539.

[Yang et al., 2012] Yang, K., Hung, B., et al. (2012). Shallow water demultiple with seafloor reflection modeling using multichannel prediction operator. In *2012 SEG Annual Meeting*. Society of Exploration Geophysicists.

[Ypma and Verschuur, 2013] Ypma, F. and Verschuur, D. (2013). Estimating primaries by sparse inversion, a generalized approach. *Geophysical Prospecting*, 61(s1):94–108.