



TIAGO TAVARES LEITE BARROS

**IMPLEMENTATION ASPECTS OF EIGENDECOMPOSITION-BASED
HIGH-RESOLUTION VELOCITY SPECTRA**

***ASPECTOS DE IMPLEMENTAÇÃO DE ESPECTROS DE VELOCIDADE DE
ALTA RESOLUÇÃO BASEADOS EM DECOMPOSIÇÃO ESPECTRAL***

CAMPINAS

2012



**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO**

TIAGO TAVARES LEITE BARROS

**IMPLEMENTATION ASPECTS OF EIGENDECOMPOSITION-BASED
HIGH-RESOLUTION VELOCITY SPECTRA**

***ASPECTOS DE IMPLEMENTAÇÃO DE ESPECTROS DE VELOCIDADE DE
ALTA RESOLUÇÃO BASEADOS EM DECOMPOSIÇÃO ESPECTRAL***

Masters Dissertation presented to the Electrical Engineering Postgraduation Programm of the School of Electrical and Computer Engineering of the University of Campinas to obtain the Masters Degree in Electrical Engineering. Concentration area: Telecommunications and Telematics.

Dissertação de Mestrado apresentada ao programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Telecomunicações e Telemática.

Tutor: Associate Professor Renato da Rocha Lopes

Orientador: Prof. Dr. Renato da Rocha Lopes

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO TIAGO TAVARES LEITE BARROS, E ORIENTADA PELO PROF. DR. RENATO DA ROCHA LOPES

Assinatura do Orientador

**CAMPINAS
2012**

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

B278a Barros, Tiago Tavares Leite, 1972-
Aspectos de implementação de espectros de velocidade
de alta resolução baseados em decomposição espectral /
Tiago Tavares Leite Barros. – Campinas, SP: [s.n.], 2012.

Orientador: Renato da Rocha Lopes.
Dissertação de Mestrado - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Processamento de sinais. 2. Correlação
(Estatística). 3. Autovalores. 4. Autovetores. 5.
Processamento geofísico. I. Lopes, Renato da Rocha. II.
Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica e de Computação. III. Título.

Título em Inglês: Implementation aspects of eigendecomposition-based
high-resolution velocity spectra
Palavras-chave em Inglês: Signal processing, Correlation (statistics), Eigenvalues,
Eigenvectors, Geophysical processing
Área de concentração: Telecomunicações e Telemática
Titulação: Mestre em Engenharia Elétrica
Banca Examinadora: Leiv Jacob Gelius, Rafael Krummenauer
Data da defesa: 28-09-2012
Programa de Pós Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Tiago Tavares Leite Barros

Data da Defesa: 28 de setembro de 2012

Título da Tese: "Aspectos de Implementação de Espectros de Velocidade de Alta Resolução Baseados em Decomposição Espectral"

Prof. Dr. Renato da Rocha Lopes (Presidente): Renato Lopes

Prof. Dr. Leiv Jacob Gelius: Leiv J. Gelius

Dr. Rafael Kruppenauer: Rafael Kruppenauer

Agradecimentos

Este trabalho não seria possível sem a colaboração de muitas pessoas, às quais devo mais do que apenas agradecimentos. Por isso, começo agradecendo:

À minha mãe, Lúcia, e aos meus irmãos, pelos constantes apoio e incentivo. Sem vocês eu não teria concluído este objetivo.

À minha esposa, Ana Luiza, pela paciência em ser casada com um pesquisador e, principalmente, pelo amor e companheirismo. Espero que continuemos a ser muito felizes juntos.

Ao meu orientador, Prof. Renato da Rocha Lopes, por me guiar nesse caminho sempre com muita disposição e compreensão. Trabalhar com alguém como ele foi um privilégio.

À comissão julgadora deste trabalho, Prof. Leiv Gelius e Dr. Rafael Krummenauer, pelas contribuições e acompanhamento ao longo do mesmo. Não consigo imaginar pessoas mais qualificadas para julgá-lo.

Aos Profs. do DSPCom João Romano, Romis Attux e Leonardo Tomazeli Duarte, por todo apoio e também pelas “trocas de idéias”.

Aos Profs. Martin Tygel e Hervé Perroud, pelas contribuições na área de geofísica.

Aos colegas que diretamente ajudaram neste trabalho: Kenji Nose, por me apresentar às *toolboxes* de processamento sísmico. Endrias Asgedom, pelas contribuições no entendimento do método MUSIC. Everton Nadalin, Jorge Faccipieri Jr. e Kazuo Takahata, pelos palpites e por me acompanharem no congresso EAGE 2012, na Dinamarca. Rafael Ferrari, pela ajuda na análise das simulações.

Aos demais colegas e amigos do DSPCom.

Aos colegas de trabalho e amigos da Idea! Sistemas Eletrônicos, pela flexibilidade em trabalhar com um estudante de mestrado.

Ao consórcio WIT (*Wave Inversion Technology*).

A todos a quem eu possa ter cometido a injustiça de não mencioná-los.

“Digo: o real não está na saída nem na chegada: ele se dispõe para a gente é no meio da travessia.”

João Guimarães Rosa

Abstract

In this dissertation we discuss high-resolution coherence functions for the estimation of the stacking parameters in seismic signal processing. We focus on the MULTiple Signal Classification (MUSIC) algorithm, which uses the eigendecomposition of the seismic data to measure the coherence. MUSIC can outperform the traditional semblance in cases of close or interfering reflections. Our main contribution is to propose several simplifications to the implementation of MUSIC. First, we show how to compute MUSIC coherence measure in terms of the signal subspace of seismic data, which has lower dimension than the one currently used, the noise subspace. After that, we show how to obtain the signal subspace, iteratively, with the power method. We called this technique of Power Method MUSIC (PM-MUSIC). We also propose a new way to obtain the MUSIC spectrum, based on the eigendecomposition of the temporal correlation matrix of the seismic data. This is in contrast to the algorithms in the literature, which are based on the spatial correlation. Complexity reductions are obtained and discussed with the use of Power Method for both spatial and temporal variant of MUSIC. Finally, we propose a new normalization function for MUSIC, which we called semblance weighting. This function takes into account semblance coefficient and deals with high dynamic range in MUSIC velocity spectrum. We compared spatial and temporal correlation matrices, implemented with PM-MUSIC. Numerical examples with synthetic and real seismic data indicated that PM-MUSIC outperforms semblance and that the temporal variant of PM-MUSIC can present the same high-resolution as its spatial counterpart. Moreover, temporal PM-MUSIC is particularly useful when dealing with correlated signals.

Keywords: Signal processing, Correlation (statistics), Eigenvalues, Eigenvectors, Geophysical processing.

Resumo

Nesta dissertação, nós discutimos o cálculo de funções de coerência de alta resolução para a estimação dos parâmetros de empilhamento em processamento de sinais sísmicos. Nosso foco é o algoritmo de estimação por Classificação de Sinais Múltiplos (MUSIC, do inglês MULTiple SIGNAL Classification). Este pode superar a tradicional função de coerência *semblance* em casos em que há reflexões próximas ou interferentes. Nossa principal contribuição é a proposta de diversas simplificações para sua implementação. Primeiro, mostramos como obter os valores da função MUSIC a partir do subespaço de sinais do dado sísmico, que possui dimensão menor do que o subespaço de ruído, usualmente empregado. Depois disso, mostramos como obter o subespaço de sinais a partir do método da potência. Chamamos esta técnica de MUSIC com Método da Potência (PM-MUSIC). Também propusemos uma nova maneira de obtenção do espectro de MUSIC, baseada na decomposição em autovalores e autovetores da matriz de correlação temporal do dado sísmico. Este método contrasta com os algoritmos presentes na literatura, que se baseiam na correlação espacial. A partir do uso do Método da Potência, obtivemos reduções de complexidade tanto para a variante espacial quanto para a temporal do algoritmo MUSIC. Finalmente, também propusemos uma nova função de normalização para o cálculo de MUSIC, a qual chamamos de ponderação por *semblance*. Esta função leva em conta o espectro de velocidades obtido com a função de coerência *semblance* e lida com a alta variação dinâmica produzida pelo espectro de velocidades calculado com MUSIC. Nós comparamos a implementação de PM-MUSIC, a partir das correlações temporal e espacial. Exemplos numéricos com dados sísmicos sintéticos e de levantamentos reais demonstraram que o algoritmo PM-MUSIC supera o *semblance* e que sua variante temporal possui alta resolução, assim como sua variante espacial. Além disso, PM-MUSIC obtido a partir da correlação temporal mostrou-se extremamente robusto ao lidar com sinais correlacionados.

Palavras-chave: Processamento de sinais, Correlação (Estatística), Autovalores, Autovetores, Processamento geofísico.

Summary

List of Figures	x
List of Tables	xiii
1 Introduction	1
2 Eigendecomposition-based high-resolution velocity spectra	5
2.1 Overview	5
2.2 The model for eigendecomposition-based velocity spectra	6
2.3 Correlation matrix	8
2.3.1 Eigendecomposition of the correlation matrix	9
2.4 MULTiple Signal Classification	11
2.4.1 Correlated wavefronts	12
2.4.2 Wide-band signals	14
3 The windowing method for velocity spectra computation	16
3.1 Overview	16
3.2 Windowed data	16
3.3 The windowed spatial correlation matrix	18
3.4 The windowed spatial correlation matrix and MUSIC	20
3.4.1 Signal subspace dimension	21
3.5 The windowed spatial correlation matrix and semblance	22
4 Proposed method	24
4.1 Overview	24
4.2 Signal and noise subspaces	24
4.3 The windowed temporal correlation matrix	25
4.4 Power Method	28

4.4.1	MUSIC with Power Method for the windowed spatial correlation matrix . . .	28
4.4.2	MUSIC with Power Method for the windowed temporal correlation matrix .	29
4.5	Velocity spectra normalization	30
4.5.1	Semblance Balancing	31
4.5.2	Semblance Weighting	31
5	Numerical examples	33
5.1	Overview	33
5.2	Synthetic CMP data with two reflections	33
5.2.1	Comparison of the methods	34
5.2.2	Power Method	35
5.3	Real data set: Jequitinhonha Basin	38
5.3.1	Comparison of the methods	38
5.3.2	Power Method	40
5.4	Stacking of synthetic data	40
6	Conclusions	54
	References	56

List of Figures

1.1	(a) Common shot and (b) common-midpoint geometries.	2
1.2	NMO geometry for a single horizontal reflector.	2
1.3	Seismogram sorted in CMP gathers, with one reflection.	3
2.1	Spatial smoothing with K subarrays, each with $N_r - K + 1$ receivers	13
3.1	Non windowed and windowed data with the right (a) and wrong (b) velocity. The window forms the transpose seismic data matrix, $\mathbf{D}^H(\theta_k)$	17
3.2	For the window that fits a wavefront, the seismic data matrix is a repetition of the source wavelet in every trace. It can be written as a multiplication between the vector $\mathbf{1}$ and the source wavelet.	18
3.3	View of the spatial correlation matrix in terms of seismic data matrix.	19
4.1	View of temporal correlation matrix in terms of seismic data matrix.	26
4.2	Diagram of the estimated seismic wavelet. The summation is done along the receivers dimension, for each column of the figure.	27
5.1	Synthetic CMP section.	34
5.2	Semblance velocity spectrum in 3D (a) and in 2D close view (b).	35
5.3	S-MUSIC velocity spectrum in 3D (a) and in 2D close view (b).	35
5.4	T-MUSIC velocity spectrum in 3D (a) and in 2D close view (b).	36
5.5	S-MUSIC velocity spectrum computed with power method in 3D (a) and in 2D close view (b).	36
5.6	T-MUSIC velocity spectrum computed with power method in 3D (a) and in 2D close view (b).	36
5.7	Histogram of number of iterations for convergence of PM-MUSIC applied on spatial (a) and temporal (b) correlation matrices.	37

5.8	Images with the number of iterations of PM-MUSIC for velocity spectra obtainment with spatial (a) and temporal (b) correlation matrices.	37
5.9	CMP 1801 from marine data set.	38
5.10	Semblance velocity spectrum from CMP 1801 (a) and close view of the region inside the black box (b).	39
5.11	S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with SW. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra. .	41
5.12	S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with SB. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra. .	42
5.13	S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with the power method and normalized with SW. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra.	43
5.14	S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with the power method and normalized with SB. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra.	44
5.15	Histogram of number of iterations for convergence of PM-MUSIC applied on spatial (a) and temporal (b) correlation matrices.	45
5.16	Images with the number of iterations of PM-MUSIC for velocity spectra obtainment with spatial (a) and temporal (b) correlation matrices.	45
5.17	Velocity model (a) and close view of the region inside the red box, for stacking comparison (b).	47
5.18	CMP gather at the midpoint position of 1600 m (a) and velocity spectra computed with semblance (b), PM-S-MUSIC (c) and PM-T-MUSIC (f) without normalization and normalized with semblance balancing (d) and (g) and semblance weighting (e) and (h).	48
5.19	CMP gather at the midpoint position of 1800 m (a) and velocity spectra computed with semblance (b), PM-S-MUSIC (c) and PM-T-MUSIC (f) without normalization and normalized with semblance balancing (d) and (g) and semblance weighting (e) and (h).	49
5.20	CMP gather at the midpoint position of 1950 m (a) and velocity spectra computed with semblance (b), PM-S-MUSIC (c) and PM-T-MUSIC (f) without normalization and normalized with semblance balancing (d) and (g) and semblance weighting (e) and (h).	50
5.21	Stacking result close view for semblance.	51

5.22	Stacking result close view for PM-MUSIC, without normalization, computed for spatial (a) and temporal (b) correlation matrices.	51
5.23	Stacking result close view for PM-MUSIC, normalized with SW and computed for spatial (a) and temporal (b) correlation matrices.	52
5.24	Stacking result close view for PM-MUSIC, normalized with SB and computed for spatial (a) and temporal (b) correlation matrices.	52
5.25	Stacking velocities estimated with semblance (a), PM-S-MUSIC (b), SB PM-S-MUSIC (c), SW PM-S-MUSIC (d), PM-T-MUSIC (e), SB PM-T-MUSIC (f) and SB PM-S-MUSIC (g). The criterium was to choose, for each τ_0 , the velocity corresponding to the highest coherence value.	53

List of Tables

4.1	Number of Operations for semblance and PM-MUSIC Algorithms	30
-----	--	----

Chapter 1

Introduction

Seismic data acquisition is performed in shot-receiver coordinate (s, g) [Yilmaz, 2001]. Figure 1.1(a) illustrates the recording geometry and ray paths associated with a horizontal reflector. In this case the recorded data will form a common-shot gather with the same shot recorded at different receivers (traces). For processing seismic data we can sort the recorded traces in different ways. In figure 1.1(b) a common-midpoint (CMP) geometry is illustrated, also with the ray paths associated with a horizontal reflector, where all the reflections come from the same depth point. The CMP gather is generated by putting the recorded seismic traces with the same midpoint, y , between the shot and receiver locations associated with that trace in a gather [Yilmaz, 2001]. The coordinates of the CMP gather are the midpoint and half offset (y, h) , with the half offset being half the distance between source and receiver. The (y, h) coordinates are defined in terms of (s, g) by: $y = (g + s)/2$ and $h = (g - s)/2$.

The traveltimes of the traces in the CMP configuration in figure 1.2 is called normal moveout (NMO). We let h_i be the half-offset between the source and the receiver i , if we assume a homogeneous medium with a single horizontal reflector we can estimate the stacking velocity, v_k , with the NMO traveltimes, defined in Dix [1955], which is given by

$$t^2(h_i) = \tau_0^2 + \frac{4h_i^2}{v_k^2}. \quad (1.1)$$

In equation (1.1) τ_0 is the two-way zero-offset (ZO) traveltimes, which represents the time the seismic wave takes to propagate from the shot to the reflector and reflect back to the receiver, if both the source and receiver were at the midpoint position of the CMP being analyzed, known as the ZO position. The stacking velocity is the velocity that presents a better curve fitting of a reflection in the equation presented in Dix [1955]. This equation is exact for a single horizontal reflector under a homogeneous medium, as in figure 1.2, and can be obtained by simple geometry [Yilmaz, 2001]. For a generic

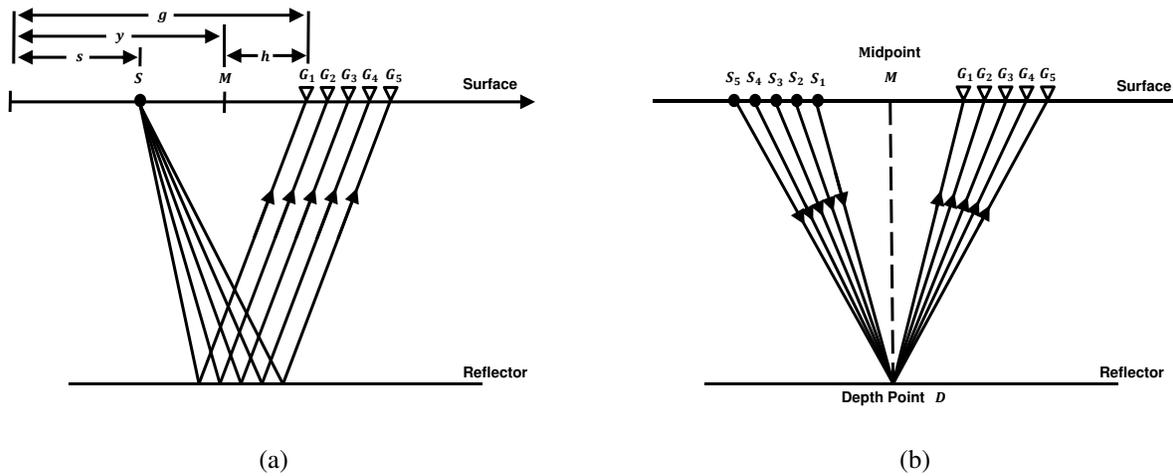


Figure 1.1: (a) Common shot and (b) common-midpoint geometries.

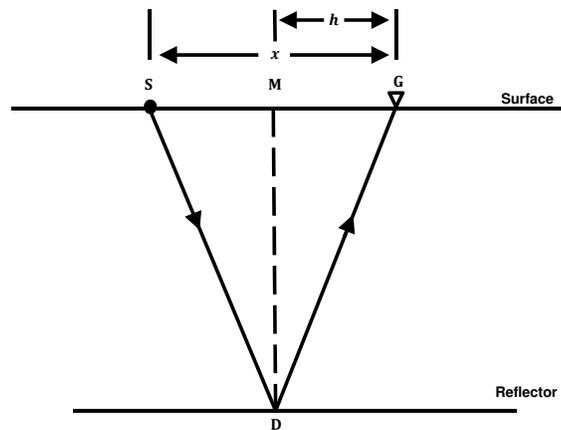


Figure 1.2: NMO geometry for a single horizontal reflector.

reflector and also for more than one reflector equation (1.1) is still a good approximation [Yilmaz, 2001]. In figure 1.3 we show an example of traces in a CMP gather, in which there is a single reflection.

We can sum the traces from the same CMP gather along the traveltime curve from equation (1.1), which will result in a trace with increased signal-to-noise ratio (SNR). This operation is called stacking and if performed for several CMP gathers, with different midpoint coordinates, will generate a stacked section. The stacked trace can be viewed as a simulated ZO trace, where we say that each trace was produced by the virtual experiment of shooting and receiving a seismic shot at the same point. For stacking the traces from a CMP gather, we must estimate the stacking parameters from equation (1.1), which is done by means of a velocity analysis applied directly to the CMP data. The

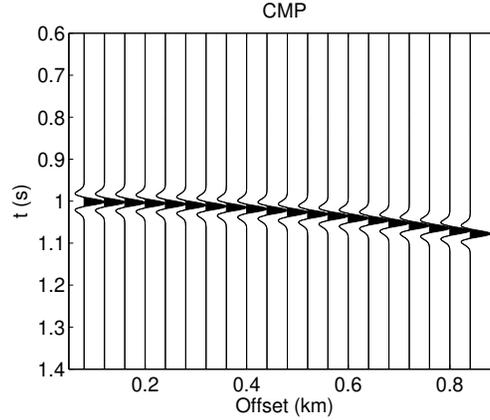


Figure 1.3: Seismogram sorted in CMP gathers, with one reflection.

velocity analysis procedure consists in, for each τ_0 , to design several traveltimes curves using equation (1.1) with possible stacking velocities, v_k , and to choose the v_k that results in a maximum value of a coherence function, which measures the similarity between the samples of the traces along that curve [Yilmaz, 2001].

The standard coherence function is a second-order energy measure, called semblance [Neidell and Taner, 1971]. Semblance is computed for windows of N_t samples taken from traces at N_r receivers. Each window follows the moveout defined by the parameters being estimated, and consists of a few samples before and after the window center. Given the n -th time sample of data $d_i(n)$, at the i -th receiver, the semblance is defined as

$$S_c = \frac{\sum_{l=k-(N_t-1)/2}^{l=k+(N_t-1)/2} \left| \sum_{i=1}^{N_r} d_i(n) \right|^2}{N_r \sum_{l=k-(N_t-1)/2}^{l=k+(N_t-1)/2} \sum_{i=1}^{N_r} |d_i(n)|^2}. \quad (1.2)$$

Still within the framework of velocity analysis, Biondi and Kostov [1989] and Kirlin [1992] showed that eigenstructure methods for velocity (coherence) analysis can lead to parameter estimations (in this case velocity spectra) with higher resolution than semblance. One of most commonly used high-resolution methods is MULTiple SIGNAL Classification (MUSIC), introduced by Schmidt [1986], which is based in some properties of the eigendecomposition of the seismic data. Recently, MUSIC has been used in Asgedon et al. [2011] for estimating the common-reflection-surface (CRS) attributes.

The implementation of MUSIC-based velocity spectra is the main focus of this work. We propose a number of improvements, namely:

1. To reduce its computational complexity, we compute the MUSIC coherence measure based on a single eigenvector, the one associated to the largest eigenvalue, which will be referred to as the *largest eigenvector*. As a consequence, the full eigenvalue decomposition is not required. Moreover, the largest eigenvector can be efficiently computed by means of the power method [Golub and Van Loan, 1996].
2. To obtain further computational savings, we propose a coherence function based on the eigendecomposition of the temporal correlation matrix of seismic data. This matrix presents lower dimension when compared to the spatial correlation matrix, currently used in the literature [Biondi and Kostov, 1989, Kirilin, 1992]. In consequence, it is simpler to compute its eigenvectors.
3. As a byproduct, the use of this lower-dimensional matrix seems to improve the performance of the method when dealing with correlated wavefronts, as indicated by numerical experiments.
4. In order to reduce the arbitrary amplitude values from the eigendecomposition-based velocity spectra [Asgedon et al., 2011, Abbad and Ursin, 2012] we propose a normalization function that we called *semblance weighting*.

The remainder of the dissertation is organized in this way:

- **Chapter 2:** We present the basics of high-resolution eigendecomposition methods applied to velocity analysis in seismic parameter estimation. In this chapter we outline how the Direction of Arrival (DOA) techniques are mapped to seismic CMP gathers.
- **Chapter 3:** In this chapter, we describe the windowing operation, applied to standard high-resolution seismic methods.
- **Chapter 4:** Here, we present our methodology. Based on seismic windowing, we develop an implementation of high-resolution coherence functions.
- **Chapter 5:** In this chapter, we present the numerical examples we did to test the proposed method and compare it to methods in the literature.
- **Chapter 6:** We present the general conclusions and prospects involving our work in this chapter.

During the period of the master, the following article was published:

- T. Barros, R. Lopes, M. Tygel, J.M.T. Romano, "Implementation Aspects of Eigenstructure-based Velocity Spectra", Proceedings of the 74-th EAGE Conference & Exhibition, 2012, Copenhagen.

Chapter 2

Eigendecomposition-based high-resolution velocity spectra

2.1 Overview

Eigendecomposition-based algorithms have been widely used for high-resolution parameter estimation [Van Trees, 2002]. In seismic, one of its first use was in Key et al. [1987]. This method was later extended in Key and Smithson [1990], where the eigenvalues from the seismic data covariance matrix calculated for data windowed around hyperbolic normal moveout (NMO) trajectories were used for the estimation of stacking parameters. A higher resolution was obtained in Biondi and Kostov [1989], by the use of subspace partition of the data covariance matrix, in accordance to the methods of direction of arrival (DOA) estimation, traditionally applied to sensor arrays [Krim and Viberg, 1996]. The same higher resolution subspace-based methods were analyzed in Kirilin [1992], where the windowing properties are discussed and the semblance coherence function is written in terms of the covariance matrices. Later, in Asgedon et al. [2011], those methods were used for estimating parameters from a different moveout equation, the common-reflection-surface (CRS) moveout.

In this chapter we show a little of the history of the application of high-resolution DOA estimation methods in seismic stacking parameters estimation. In section 2.2 we illustrate the model assumptions for the application of these methods for seismic velocity spectra computation. In section 2.3 we introduce the correlation matrix and some of its properties. Finally, in section 2.4 we introduce the MULTiple SIGNAL Classification (MUSIC) method applied to seismic velocities estimation, which makes use of the eigendecomposition of the correlation matrix.

2.2 The model for eigendecomposition-based velocity spectra

A classical estimation problem is to determine parameters of a signal, from the observation of this signal in the presence of additive white noise [Haykin, 1989]. In array signal processing, a finite number of sensors (e.g., antennas, hydrophones, geophones), named the array, is arranged, typically linearly and uniformly spaced, to record a propagating wave. The processing of the recorded signal can provide estimates of the desired parameters.

A common-midpoint (CMP) gather is a type of array, from which we can determine the velocities and zero-offset (ZO) traveltimes from the different reflections.

The seismic data from a CMP gather that contains N_r seismic traces can be modeled as a combination of N_s reflections of the wavefront source, caused by rock interfaces at the subsurface. At the time instant t and at the i -th receiver, the recorded data can be described as

$$x_i(t) = \sum_{k=1}^{N_s} s_k(t - \tau_k(i)) + n_i(t), \quad (2.1)$$

where $s_k(t)$ is the k -th observed ZO reflection from the source $s(t)$ arriving at τ_0 , $n_i(t)$ represents an additive noise supposed to be white with zero mean and variance σ_n^2 and $\tau_k(i)$ is the time difference, referred to as moveouts (delays), between the wavefronts arrivals at i -th receiver and the ZO traveltime τ_0 , for the k -th reflection.

Equation (2.1) describes a signal continuous in time impinging on the receivers. In practice, the receivers take uniform samples from the signal, with a fixed sampling period T . The sampled version of the signal $x_i(t)$, at the time instant $t = nT$, would result in a recorded discrete-time signal $x_i(n) = x_i(nT)$, for $n = 1, 2, \dots, N_T$, where N_T is the total number of samples.

The moveouts in equation (2.1), can be approximated by the hyperbolic moveout equation of Dix [1955]

$$\tau_k(i) = \sqrt{\tau_0^2 + \frac{4h_i^2}{v_k^2}} - \tau_0, \quad (2.2)$$

where h_i is the half-offset between the i -th pair of source and receiver and v_k is the stacking velocity of the k -th reflection.

In order to apply most of the classical DOA techniques, including in this group the eigenstructure-based methods, to estimate velocities v_k , we assume that the model described in (2.1) is governed by the following assumptions [Haykin, 1989, Biondi and Kostov, 1989]:

1. The reflections from the source are uncorrelated.
2. The additive white noise $n_i(t)$ is uncorrelated with each reflection component from the received

signal.

3. The propagating waves are *narrow-band*, which means adequately characterization by a single frequency ω .
4. The number of receivers, N_r , is larger than the number of reflections, N_s .
5. The number of reflections, N_s , from the source, is known.

For seismic wave propagation, many of the above assumptions are not true, but there are some ways of dealing with this fact.

Regarding the assumption number 1, seismic reflections are indeed correlated. As will be discussed further, correlated reflections can generate a singular spatial covariance matrix, what can cause the method to loose its ability to distinguish different events. A spatial smoothing of the covariance matrix [Shan et al., 1985] can be applied used to overcome this issue of correlated signals, as will be discussed later.

Seismic signals are also wide-band, which goes against assumption number 3. Biondi and Kostov [1989] perform a frequency domain coherence computation, in which case time delays can be approximated by phase shifts.

The moveout parameters may vary rapidly with time, *i.e.*, seismic signals are highly non-stationary, which implies that when processing seismic data only a few samples can be used for the estimation of the covariance matrix. This contradicts assumption number 4 and may result in a bad conditioning of the covariance matrix. Forward-backward (FB) averaging [Williams et al., 1988] technique together with the spatial smoothing Shan et al. [1985] of the covariance matrix can be used in the attempt to overcome this situation. In Kirilin [1991], an analysis about the effects of non-idealness of temporal wide-band and temporal stationarity assumptions has been made. It is shown that non-stationary signals may present high correlation values, even when they are not correlated, decreasing the rank of the correlation matrix whenever the signals are coherent.

Usually the number of reflections is not known, but it can be estimated [Wax and Kailath, 1985, Wang and Kaveh, 1985] to satisfy the assumption number 5.

Returning to the data model from equation (2.1), using the assumption that the source reflections $s_k(t)$ are narrow-band, with center frequency ω , the moveouts can be approximated as phase shifts [Biondi and Kostov, 1989], so the data can be expressed as

$$x_i(t) = \sum_{k=1}^{N_s} s_k(t) e^{j\omega\tau_k(i)} + n_i(t). \quad (2.3)$$

In order to simplify data manipulations, equation (2.3) can be written in matrix notation, considering a subset¹ of $N_{T'}$ time samples, from the total amount of N_T samples:

$$\mathbf{X} = \mathbf{A}(\Theta)\mathbf{S} + \mathbf{N}, \quad (2.4)$$

where \mathbf{X} and \mathbf{N} are $N_r \times N_{T'}$ data and noise matrices, \mathbf{S} is a $N_s \times N_{T'}$ source matrix, and $\mathbf{A}(\Theta)$ is a $N_r \times N_s$ matrix formed by the N_s steering vectors:

$$\mathbf{A}(\Theta) = \begin{bmatrix} \mathbf{a}_1(\theta_1) & \cdots & \mathbf{a}_k(\theta_k) & \cdots & \mathbf{a}_{N_s}(\theta_{N_s}) \end{bmatrix}. \quad (2.5)$$

Given that $\theta_k \triangleq 1/v_k$ and that $\tau_k(i)$ and θ_k are related by equation (2.2), we can form the $N_r \times 1$ columns of $\mathbf{A}(\Theta)$ by the wavefront arrival delays at each receiver

$$\mathbf{a}_k(\theta_k) = \frac{1}{\sqrt{N_r}} \begin{bmatrix} e^{j\omega\tau_k(1)} \\ e^{j\omega\tau_k(2)} \\ \vdots \\ e^{j\omega\tau_k(N_r)} \end{bmatrix}, \quad (2.6)$$

where the normalization factor is to ensure steering vectors with unitary norm.

2.3 Correlation matrix

The eigenstructure-based methods for velocity spectra calculation are based on the spatial covariance matrix of the seismic data. If we look at the seismic digital data $x_i(n)$ at the i -th receiver and n -th time sample we can define the vectors

$$\mathbf{x}(n) = [x_1(n) \ x_2(n) \ \cdots \ x_{N_r}(n)]^H. \quad (2.7)$$

Assuming that the reflections are zero-mean random process, that the noise is uncorrelated with the source and that the signals are spatially and temporally stationary, the covariance matrix, in this case

¹A small number of samples for the computation of the covariance matrix can soften the temporal non-stationarity of seismic signals.

also called correlation matrix, is defined as

$$\mathbf{R} = \mathbb{E}\{\mathbf{x}(n)\mathbf{x}^H(n)\} \quad (2.8a)$$

$$= \begin{bmatrix} \mathbb{E}\{x_1(n)x_1^*(n)\} & \mathbb{E}\{x_1(n)x_2^*(n)\} & \dots & \mathbb{E}\{x_1(n)x_{N_r}^*(n)\} \\ \mathbb{E}\{x_2(n)x_1^*(n)\} & \mathbb{E}\{x_2(n)x_2^*(n)\} & \dots & \mathbb{E}\{x_2(n)x_{N_r}^*(n)\} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}\{x_{N_r}(n)x_1^*(n)\} & \mathbb{E}\{x_{N_r}(n)x_2^*(n)\} & \dots & \mathbb{E}\{x_{N_r}(n)x_{N_r}^*(n)\} \end{bmatrix} \quad (2.8b)$$

where $\mathbb{E}\{\}$ is the expectation operator and the superscripts H and $*$ denote matrix Hermitian and complex conjugate. Note that \mathbf{R} is an $N_r \times N_r$ matrix. For stationary signals, the correlation function $R_x(p, m, n, k) = \mathbb{E}\{x_p(n)x_m^H(k)\}$, for $n = k$, can be written in terms of the difference $l = m - p$, *i.e.*, $R_x(p, m, n, k) = R_x(l, 0) \triangleq R_x(l)$. Since $R_x(l) = R_x(-l)$, the correlation matrix can be written, then, as

$$\mathbf{R} = \begin{bmatrix} R_x(0) & R_x(1) & \dots & R_x(N_r - 1) \\ R_x(1) & R_x(0) & \dots & R_x(N_r - 2) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(N_r - 1) & R_x(N_r - 2) & \dots & R_x(0) \end{bmatrix}. \quad (2.9)$$

With a few manipulations of equations (2.4) and (2.8a), and recalling that the signal and noise are uncorrelated, the correlation matrix can also be written as

$$\mathbf{R} = \mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}^H(\Theta) + \sigma_n^2\mathbf{I}, \quad (2.10)$$

where \mathbf{R}_s is the source spatial correlation diagonal matrix, σ_n^2 is the noise variance and \mathbf{I} is the identity matrix. Note that the last expansion should not be confused with the orthogonal similarity transformation. In general, the diagonal elements of \mathbf{R}_s are not the eigenvalues of $\mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}^H(\Theta)$. On the other hand, the structure of equation (2.10) can be used to give an insight into the eigendecomposition of \mathbf{R} , as will be discussed next.

2.3.1 Eigendecomposition of the correlation matrix

Let $\lambda_1 \geq \dots \geq \lambda_{N_r}$ and $\nu_1 \geq \dots \geq \nu_{N_r}$ be the eigenvalues of \mathbf{R} and $\mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}^H(\Theta)$, respectively. From equation (2.10), these eigenvalues are related by

$$\lambda_i = \nu_i + \sigma_n^2, \quad i = 1, \dots, N_r. \quad (2.11)$$

The columns of $\mathbf{A}(\Theta)$ are linearly independent and if we assume N_s wavefronts arriving at the array, with different velocities and ZO times, it will imply that $\mathbf{A}(\Theta)$ will have rank N_s . As a consequence, the $(N_r - N_s)$ smallest eigenvalues of $\mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}^H(\Theta)$ are equal to zero, so we may write,

$$\nu_{N_s+1} = \cdots = \nu_{N_r} = 0. \quad (2.12)$$

By comparing equations (2.11) and (2.12), we can see that

$$\lambda_{N_s+1} = \cdots = \lambda_{N_r} = \sigma_n^2. \quad (2.13)$$

Equation (2.13) states that the smallest eigenvalues of the correlation matrix \mathbf{R} are equal to σ_n^2 with multiplicity $(N_r - N_s)$.

Now, let \mathbf{v}_i denote the eigenvector of the correlation matrix, \mathbf{R} , associated with the eigenvalue λ_i . By definition,

$$\mathbf{R}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \quad i = 1, \cdots, N_r. \quad (2.14)$$

For the eigenvectors associated with the $(N_r - N_s)$ smallest eigenvalues² of \mathbf{R} , it follows that $\lambda_i = \sigma_n^2$. Then, equation (2.14) can be written as

$$(\mathbf{R} - \sigma_n^2)\mathbf{v}_i = 0, \quad i = N_s + 1, \cdots, N_r. \quad (2.15)$$

By combining equations (2.10) and (2.15), we can write

$$\mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}^H(\Theta)\mathbf{v}_i = 0, \quad i = N_s + 1, \cdots, N_r. \quad (2.16)$$

The columns of $\mathbf{A}(\Theta)$ are linearly independent and considering the N_s wavefronts uncorrelated \mathbf{R}_s is a full-rank diagonal matrix, which implies that equation (2.16) can be satisfied if,

$$\mathbf{A}^H(\Theta)\mathbf{v}_i = 0, \quad i = N_s + 1, \cdots, N_r. \quad (2.17)$$

Using (2.5), we can write equation (2.17) as

$$\mathbf{a}_k^H(\theta_k)\mathbf{v}_i = 0, \quad k = 1, \cdots, N_s, \quad i = N_s + 1, \cdots, N_r. \quad (2.18)$$

From equation (2.18), we can state that the set of N_s steering vectors, $\{\mathbf{a}_1(\theta_1), \cdots, \mathbf{a}_{N_s}(\theta_{N_s})\}$,

²In order to simplify our explanations, we will commit an abuse of notation and refer to the eigenvector associate to the smallest eigenvalue as the smallest eigenvector, and vice-versa.

from $\mathbf{A}(\Theta)$, is orthogonal to the set of $(N_r - N_s)$ smallest eigenvectors, $\{\mathbf{v}_{N_s+1}, \dots, \mathbf{v}_{N_r}\}$ from the correlation matrix \mathbf{R} . It is also important to note that, since the eigenvectors of \mathbf{R} form a complete basis, the set of eigenvectors $\{\mathbf{v}_1, \dots, \mathbf{v}_{N_s}\}$ is orthogonal to $\{\mathbf{v}_{N_s+1}, \dots, \mathbf{v}_{N_r}\}$.

The matrix $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_{N_r}]$ can be formed with the N_r eigenvectors of \mathbf{R} . The set of all those eigenvectors is called the *vector space* \mathbb{R}^{N_r} . The properties of the eigenvectors suggest the partition of the space \mathbb{R}^{N_r} into two subspaces, which are the orthogonal complement of each other:

1. A *signal subspace* that is spanned by the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_{N_s}$ associated with the N_s largest eigenvalues from \mathbf{R} .
2. A *noise subspace* that is spanned by the eigenvectors $\mathbf{v}_{N_s+1}, \dots, \mathbf{v}_{N_r}$ associated with the smallest eigenvalues from \mathbf{R} , with multiplicity $(N_r - N_s)$.

In this case, the eigendecomposition of \mathbf{R} can be written as [Kirlin, 1992]

$$\mathbf{R} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H \quad (2.19a)$$

$$= \mathbf{V}_s\mathbf{\Lambda}_s\mathbf{V}_s^H + \mathbf{V}_n\mathbf{\Lambda}_n\mathbf{V}_n^H, \quad (2.19b)$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_{N_r})$ is a diagonal matrix that contains the eigenvalues of \mathbf{R} . We assume that $\lambda_1 \geq \dots \geq \lambda_{N_r}$. $\mathbf{\Lambda}_s$ contains the N_s largest eigenvalues and $\mathbf{\Lambda}_n$ contains the smallest eigenvalues of \mathbf{R} , all equal to σ_n^2 . The matrix \mathbf{V}_n spans the noise subspace, which is orthogonal to the image of $\mathbf{A}(\Theta)$, and the matrix \mathbf{V}_s spans the signal subspace. Finally, \mathbf{V} is a unitary matrix that can be decomposed as $\mathbf{V} = [\mathbf{V}_s \mathbf{V}_n]$.

Since the noise subspace is orthogonal to the steering vectors that constitute the signal, we may search for all the possible steering vectors and measure how orthogonal they are to \mathbf{V}_n . Then, we may estimate the parameters Θ as those that yield the steering vectors with the smallest projection onto the noise subspace.

2.4 Multiple Signal Classification

In order to estimate the parameters θ_k and τ_0 , the Multiple Signal Classification (MUSIC) [Schmidt, 1986] algorithm can be used. It consists of a search that tests the projection of several steering vectors $\mathbf{a}_k(\theta_k)$ onto the noise subspace of matrix \mathbf{R} . All the tested values form the MUSIC spectrum, which can be defined, for each tested value, as

$$P_{MU}(\theta_k) = \frac{\mathbf{a}_k^H(\theta_k)\mathbf{a}_k(\theta_k)}{\mathbf{a}_k^H(\theta_k)\mathbf{P}_n\mathbf{a}_k(\theta_k)}, \quad (2.20)$$

where $\mathbf{a}_k(\theta_k)$ is a candidate steering vector and \mathbf{P}_n is the projection matrix onto the noise subspace, given by $\mathbf{P}_n = \mathbf{V}_n \mathbf{V}_n^H$. When the candidate steering vector and noise subspaces are orthogonal, the denominator in (2.20) tends to zero. Thus, large values of $P_{MU}(\theta_k)$ will generally correspond to the actual steering vectors, in other words, to the index k of the true parameters.

For the application of MUSIC, recorded seismic data are assumed to be complex-valued and narrow-band, with central frequency ω . Actually, the seismic signal is real-valued and wide-band, but it can be transformed into an analytic and narrow-band signal, by the use of a Hilbert transform and a band-pass filter [Biondi and Kostov, 1989].

The subspace decomposition of the spatial correlation matrix has been widely used in the literature for the calculation of high resolution velocity spectra coherence functions [Biondi and Kostov, 1989, Key and Smithson, 1990, Kirlin, 1992, Sacchi, 1998, Asgedon et al., 2011]. In practice, the spatial correlation matrix must be estimated from the data. The usual approach to estimate it is by the computation of the sample correlation matrix³ $\hat{\mathbf{R}}$

$$\hat{\mathbf{R}} = \frac{1}{N_t} \mathbf{X} \mathbf{X}^H, \quad (2.21)$$

where it is assumed, implicitly, stationarity along the snapshots that compose the data set \mathbf{X} .

In order to use a spectrum search algorithm for parameter estimation it is mandatory to know the number of parameters (in this case, reflections) to be estimated. Wax and Kailath [1985] have proposed two criteria for estimating the number of wavefronts, based on the application of the information theoretic criteria for model selection, introduced by Akaike [1973] and by Schwartz [1978] and Rissanen [1978].

Both criteria tend to overestimate the number of signals when few time samples are used in the estimate of the spatial correlation matrix. For the seismic case, however, the maximum number of wavefronts can usually be set to a low number, because the need to detect more than two or three interfering wavefronts is not common [Biondi and Kostov, 1989].

2.4.1 Correlated wavefronts

Seismic signals are strongly correlated, as they are reflections of the same signal source at different layers of the earth's subsurface. If two correlated wavefronts are being analyzed at the same window, the source correlation matrix \mathbf{R}_s tends to be ill-conditioned and in the extreme case of coherent signals it will be rank deficient, causing in both cases a mix between signal and noise subspaces. The eigenstructure methods will then fail to resolve the correlated events in the coherence spectra.

³We can subtract the mean from the data before calculating the expression in (2.21), which will result in the estimation of the sample covariance matrix [Kirlin, 1992].

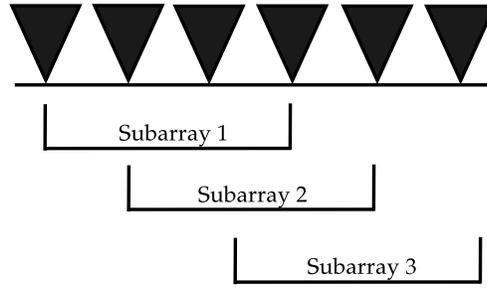


Figure 2.1: Spatial smoothing with K subarrays, each with $N_r - K + 1$ receivers

To deal adequately with correlated signals, a method called spatial smoothing in the correlation matrix is used [Shan et al., 1985]. Spatial smoothing consists in dividing the original array of N_r receivers into K overlapping subarrays of $(N_r - K + 1)$ receivers, as illustrated in Figure 2.1. A correlation matrix $\hat{\mathbf{R}}^k$ can be computed for each subarray. The spatial smoothed correlation matrix, $\hat{\mathbf{R}}^K$, will be the average from the K correlation matrices from each subarray

$$\hat{\mathbf{R}}^K = \frac{1}{K} \sum_{k=1}^K \hat{\mathbf{R}}^k. \quad (2.22)$$

Spatial smoothing has two main drawbacks. First, it reduces the effective number of sensors in the array from N_r to $(N_r - K + 1)$, reducing the resolution of the eigenstructure methods. It also increases the computational complexity to determine the estimated correlation matrix.

When dealing with correlated signals, we can also apply to the spatial smoothed correlation matrix another operation called forward-backward (FB) averaging. FB averaging consists on averaging forward and backward correlation matrices, in order to increase the array aperture [Williams et al., 1988]. We can assume that the forward correlation matrix, \mathbf{R}^f , is equal to the matrix $\hat{\mathbf{R}}$, or its spatial smoothed version. In that sense, the backward correlation matrix is defined as

$$\mathbf{R}^b \triangleq \mathbf{J}(\mathbf{R}^f)^H \mathbf{J}, \quad (2.23)$$

where \mathbf{J} is the exchange matrix

$$\mathbf{J} \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}. \quad (2.24)$$

Therefore, FB averaging of the correlation matrix, $\tilde{\mathbf{R}}$, can be computed as

$$\tilde{\mathbf{R}} = \frac{\mathbf{R}^f + \mathbf{R}^b}{2}. \quad (2.25)$$

2.4.2 Wide-band signals

The narrow-band methods for DOA estimation can not be directly applied to seismic data, since seismic data is wide-band. Biondi and Kostov [1989] propose to decompose the data in frequency components, either by Fourier transform or by filtering with a bank of filters, in order to estimate its parameters with an algorithm like MUSIC.

The number of wavefronts and stacking parameters remain the same in the frequency domain, for all frequencies. This gives rise to a search for different ways to estimate those parameters, once we do not need to use all the frequencies components to estimate them. The frequency domain methods for parameters estimation can either use the average of the final estimates of the coherence spectrum [Wax et al., 1984], for different frequencies, or use the averages of the spatial correlation matrices of the different components [Wang and Kaveh, 1985]. The later method can use techniques that linearly transform the correlation matrices of different frequencies, in order to make them approximately *coherent* to each other. Those techniques are known by coherent signal subspace approach and employ focusing matrices for wide-band array processing [Hung and Kaveh, 1988, Doron and Weiss, 1992]. The main advantage of focusing matrices is the increase of the statistical robustness of the estimates, but the number of correlation matrices combined must be carefully chosen, since they are only approximately *coherent* to each other.

The frequency-domain method proposed by Biondi and Kostov [1989], applied to MUSIC spectrum, could be summarized by:

1. Correct the data by a time moveout $\tau_k(i)$, where the parameters v_k and τ_0 have been previously chosen;
2. Decompose the time-corrected data into few N_ω wide frequency bands, by using a bank of band-pass filters;
3. Obtain the correlation matrices in frequency domain, by focusing and averaging several correlation matrices for each of the N_ω frequency bands;
4. Determine the number of wavefronts, N_s , impinging in the array of receivers, by minimizing, for each of the N_ω frequency bands, some of the criteria presented in Wax and Kailath [1985].

5. Obtain the MUSIC coherence spectrum by averaging the parameter spectra of all frequency bands, as in equation (2.26):

$$P_{MU}(\theta_k) = \frac{1}{N_\omega} \sum_{l=1}^{N_\omega} P_{MU}(\theta_k, \omega_l) = \frac{1}{N_\omega} \sum_{l=1}^{N_\omega} \frac{\mathbf{a}_k^H(\theta_k, \omega_l) \mathbf{a}_k(\theta_k, \omega_l)}{\mathbf{a}_k^H(\theta_k, \omega_l) \mathbf{P}_n(\omega_l) \mathbf{a}_k(\theta_k, \omega_l)}. \quad (2.26)$$

Now, $\mathbf{a}_k(\theta_k, \omega_l)$ and $\mathbf{P}_n(\omega_l)$ are the tested steering vector and the noise subspace projection at the l -th frequency ω_l .

Chapter 3

The windowing method for velocity spectra computation

3.1 Overview

In this chapter we focus in the seismic data windowing method applied to velocity spectra computation. In the eigendecomposition-based methods presented in chapter 2 a correlation matrix is computed with a large number of time samples and several steering vectors are tested, in order to estimate the seismic velocities and two-way *ZO* traveltime. The main difference in the windowed eigendecomposition-based methods is that, for each tested velocity and two-way *ZO* traveltime, a different correlation matrix, commonly referred to as the steered correlation matrix [Kirlin, 1991], is computed and a fixed vector is tested to determine the right set of parameters. One important advantage of the windowing-based methods is that, as the correlation matrix is computed with a small number of samples, we do not need to estimate the number of wavefronts, because we assume that there is only one event with the right pair of parameters being tested per window.

In section 3.2 we show how the windowing of the seismic data is performed, operation in which our proposed method is also based on. Then, in section 3.3 we introduce the windowed spatial correlation matrix and in sections 3.4 and 3.5 we explain how this matrix is related with MUSIC and show an interpretation of writing semblance in terms of the windowed seismic data, respectively.

3.2 Windowed data

As discussed in chapter 1, coherence is computed on a window of data centered at some time $\tau_k(i)$, where k corresponds to a given value of the parameters being tested and i corresponds to the receiver. The value of $\tau_k(i)$ depends on some parameters to be estimated. We assume a hyperbolic moveout, so

that $\tau_k(i)$ depends on the normal moveout (NMO) velocity v_k and the two-way zero-offset traveltime τ_0 . For each τ_0 , the data \mathbf{X} from equation (2.4) is windowed with windows of N_t samples¹ centered at the hyperbolic traveltime given by τ_0 and v_k . Each window corresponds to a matrix $\mathbf{D}(\theta_k)$. The dimension of $\mathbf{D}(\theta_k)$ is $N_r \times N_t$, where N_t is the number of samples in the window, N_r is the number of receivers (traces) considered and $\theta_k = 1/v_k$. In figure 3.1, we illustrate how the windowing operation is applied in seismic data. Note, however, that the rows of $\mathbf{D}(\theta_k)$ appear as vertical lines on the right of figure 3.1.

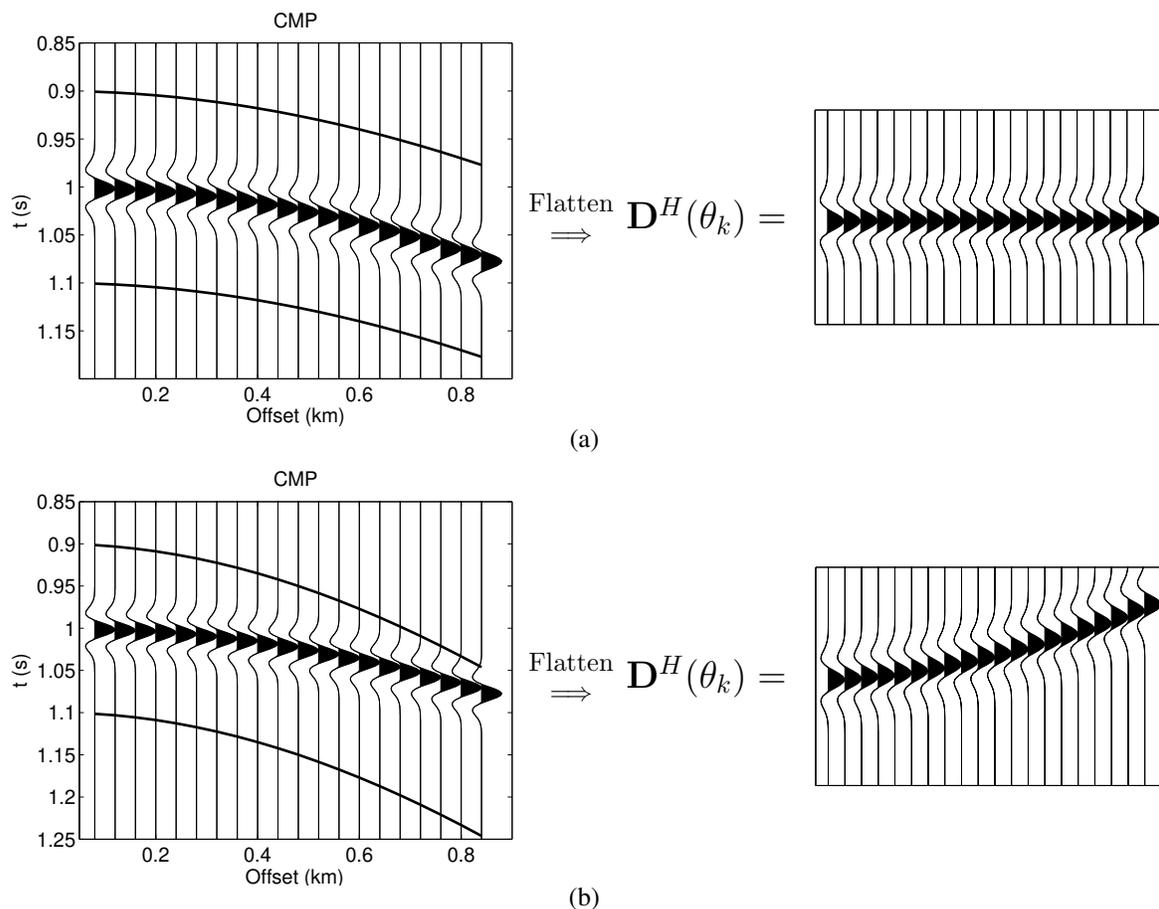


Figure 3.1: Non windowed and windowed data with the right (a) and wrong (b) velocity. The window forms the transpose seismic data matrix, $\mathbf{D}^H(\theta_k)$.

The hyperbolic windowing can also be used for eigenstructure-based coherence calculation. In section 2.2, we have explained the theory behind eigenstructure and MUSIC-based velocity spectra computation.

As will be discussed now, for each τ_0 , different values of θ_k result in different windows, and

¹Usually the number of samples used for the windowing of seismic data is very small compared to the number of receivers.

thus in different data matrices. Thus, to each θ_k corresponds a different spatial correlation matrix. Indeed, when a window with correct values of τ_0 and v_k is applied, the windowed data matrix will be represented as in figure 3.2. In other words, the data contains several repetitions of the reflection, all arriving at the same time instant at all the receivers, plus noise terms. This is the first difference between the windowed method and the one in the previous chapter. In this case, the data can be written as

$$\mathbf{D}(\theta_k) = \mathbf{1}\mathbf{s}^H + \mathbf{N}, \quad (3.1)$$

where \mathbf{s} is a $N_t \times 1$ vector that contains the samples from the reflected wavelet, $\mathbf{1}$ is a $N_r \times 1$ vector of ones, \mathbf{N} is an $N_r \times N_t$ noise matrix independent of \mathbf{s} , which may also contain interfering reflections, and the superscript H refers to the transpose conjugate operation.

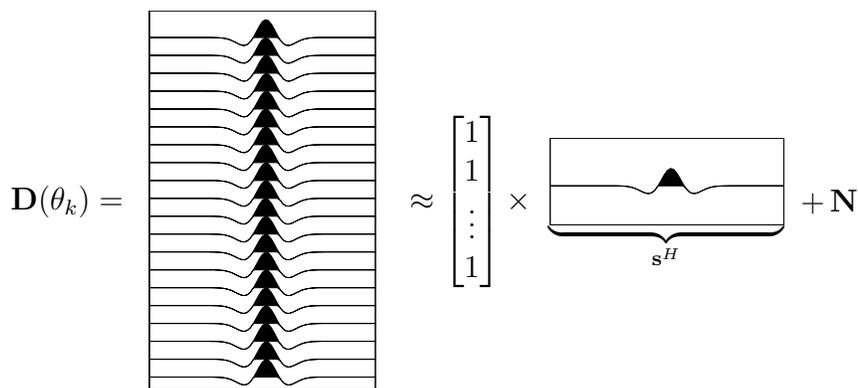


Figure 3.2: For the window that fits a wavefront, the seismic data matrix is a repetition of the source wavelet in every trace. It can be written as a multiplication between the vector $\mathbf{1}$ and the source wavelet.

3.3 The windowed spatial correlation matrix

Assuming that the seismic digital data $d_i(n)$, at the i -th receiver and n -th time sample, has already been windowed around the hyperbola centered in τ_0 , with velocity v_k , with $N_t/2$ samples above and below the window center, as in figure 3.1, we can write the estimated windowed correlation matrix in terms of the vectors

$$\mathbf{d}(n) = [d_1(n) \ d_2(n) \ \cdots \ d_{N_r}(n)]^H. \quad (3.2)$$

The windowed spatial correlation matrix would be, then,

$$\mathbf{R}(\theta_k) = \mathbf{E}\{\mathbf{d}(n)\mathbf{d}^H(n)\}. \quad (3.3)$$

The estimated spatial sample correlation matrix is given by

$$\hat{\mathbf{R}}(\theta_k) = \frac{1}{N_t} \mathbf{D}(\theta_k) \mathbf{D}^H(\theta_k), \quad (3.4a)$$

$$\approx \frac{\|\mathbf{s}\|^2}{N_t} \mathbf{1} \mathbf{1}^T + \sigma_n^2 \mathbf{I}, \quad (3.4b)$$

where σ_n^2 is the noise variance, \mathbf{I} is the identity matrix of appropriate dimension and the superscript T indicates the matrix transpose operation. Note that we disregard the cross terms resulting from $\mathbf{D}(\theta_k) \mathbf{D}^H(\theta_k)$, because we assume that the noise is zero-mean and uncorrelated with the signal. In figure 3.3 we illustrate the correlation matrix in terms of seismic windowed data matrix. The sample correlation matrix is related with the correlation matrix as $\mathbf{R}(\theta_k) = E\{\hat{\mathbf{R}}(\theta_k)\}$.

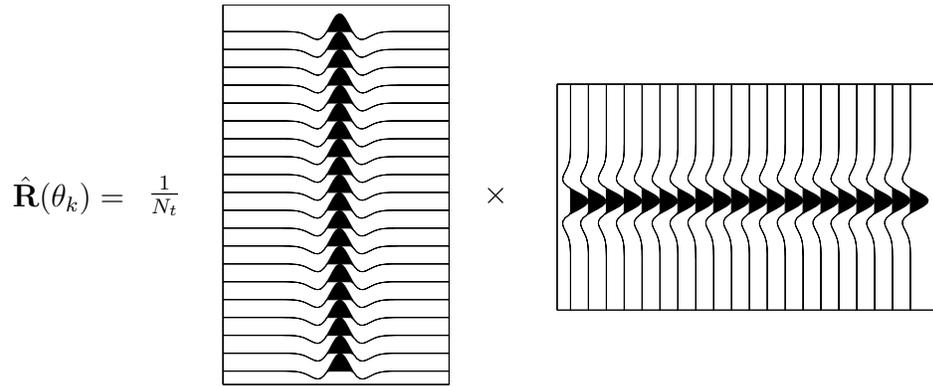


Figure 3.3: View of the spatial correlation matrix in terms of seismic data matrix.

We can now develop an intuition about the eigendecomposition of the windowed data. Assume that λ_1 is the largest eigenvalue of $\hat{\mathbf{R}}(\theta_k)$, associated with the eigenvector \mathbf{v}_1 . We can write, then,

$$\hat{\mathbf{R}}(\theta_k) \mathbf{v}_1 = \lambda_1 \mathbf{v}_1. \quad (3.5)$$

If we combine equations (3.4b) and (3.5), and if we right-multiply $\hat{\mathbf{R}}(\theta_k)$ by $(1/\sqrt{N_r}) \mathbf{1}$, with dimension $N_r \times 1$, we will have

$$\frac{1}{N_t} \|\mathbf{s}\|^2 \frac{1}{\sqrt{N_r}} \mathbf{1} \mathbf{1}^T \mathbf{1} + \sigma_n^2 \mathbf{I} \frac{1}{\sqrt{N_r}} \mathbf{1} \approx \lambda_1 \frac{1}{\sqrt{N_r}} \mathbf{1}, \quad (3.6a)$$

$$\frac{N_r}{N_t} \|\mathbf{s}\|^2 \frac{1}{\sqrt{N_r}} \mathbf{1} + \sigma_n^2 \frac{1}{\sqrt{N_r}} \mathbf{1} \approx \lambda_1 \frac{1}{\sqrt{N_r}} \mathbf{1}, \quad (3.6b)$$

where $\mathbf{1}^T \mathbf{1} = N_r$ and $\|\mathbf{s}\|^2$ is the energy of the wavelet at one receiver (trace). From equation (3.6b),

clearly, the largest eigenvalue of $\hat{\mathbf{R}}(\theta_k)$ is

$$\lambda_1 \approx \frac{N_r}{N_t} \|\mathbf{s}\|^2 + \sigma_n^2, \quad (3.7)$$

associated with the eigenvector $\mathbf{v}_1 \approx \frac{1}{\sqrt{N_r}} \mathbf{1}$. Any vector orthogonal to $\mathbf{1}$ is also an eigenvector, with eigenvalue σ_n^2 .

3.4 The windowed spatial correlation matrix and MUSIC

The matrix $\mathbf{R}(\theta_k)$ in equation (3.3) computes the correlation between traces at different receivers, and so is referred to as the spatial correlation matrix. The MUSIC-based methods in the literature are based on $\mathbf{R}(\theta_k)$, so they will be called spatial, or S-MUSIC [Wang et al., 2001]. Essentially, they can be seen as an attempt to answer the question: “Is $\mathbf{1}$ proportional to the largest eigenvector or $\hat{\mathbf{R}}(\theta_k)$?” If this answer is positive, then we may assume that $\hat{\mathbf{R}}(\theta_k)$ was formed from a window that contains a reflection. The S-MUSIC spectrum is a coherency measure that presents large values when $\mathbf{1}$ is close to the largest eigenvector of $\hat{\mathbf{R}}(\theta_k)$:

$$P_S(\theta_k) = \frac{\mathbf{1}^T \mathbf{1}}{\mathbf{1}^T \mathbf{V}_n(\theta_k) \mathbf{V}_n^H(\theta_k) \mathbf{1}}, \quad (3.8a)$$

$$= \frac{N_r}{|\mathbf{1}^T \mathbf{V}_n(\theta_k)|^2}. \quad (3.8b)$$

As discussed in section 2.2, $\mathbf{V}_n(\theta_k) \mathbf{V}_n^H(\theta_k)$ in equation (3.8a) is the projection matrix on the so-called noise subspace, consisting of the subspace spanned by the eigenvectors associated to the smallest eigenvalues of $\mathbf{R}(\theta_k)$. The S subscript on $P_S(\theta_k)$ refers to the S-MUSIC computation for the windowed spatial correlation matrix.

If the windowed data contains an event, the average of the smallest $(N_r - 1)$ eigenvalues, $\lambda_2, \dots, \lambda_{N_r}$, of the estimated correlation matrix $\hat{\mathbf{R}}(\theta_k)$ will be equal to σ_n^2 :

$$\sigma_n^2 = \frac{1}{(N_r - 1)} \sum_{i=2}^{N_r} \lambda_i. \quad (3.9)$$

For this case, the largest eigenvector of $\hat{\mathbf{R}}(\theta_k)$, \mathbf{v}_1 , that spans the signal subspace, will also be orthogonal to the remaining ones, $\mathbf{v}_2, \dots, \mathbf{v}_{N_r}$, that span the noise subspace.

We now show a reason for choosing $\mathbf{1}$ as an eigenvector of $\hat{\mathbf{R}}(\theta_k)$, for windowed data. Assume that we have flattened a wavefront in the window, by the selection of the right pair of parameters (v_k, τ_0) . All the wavefronts will arrive in the traces at the same time instant, as illustrated in figure 3.1, with

all the relative delays, $\tau_k(i)$, being equal to zero. The steering vector defined in equation (2.6), for the windowed data, becomes, then,

$$\mathbf{a}_k(\theta_k) = \frac{1}{\sqrt{N_r}} \begin{bmatrix} e^{j\omega\tau_k(1)} \\ e^{j\omega\tau_k(2)} \\ \vdots \\ e^{j\omega\tau_k(N_r)} \end{bmatrix} = \frac{1}{\sqrt{N_r}} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \quad (3.10)$$

Some advantages of the windowing operation have been discussed in Kirlin [1992]. Perhaps one of the main advantages, from the point of view of this work, is that, when searching for events with zero delay, the methods apply to both wide-band and narrow-band signals. This can be confirmed if we analyze the data model from equation (2.1), since in this case $s_k(t - \tau_k(i)) = s_k(t) \forall \omega$ if $\tau_k(i) = 0$.

3.4.1 Signal subspace dimension

The signal subspace dimension is an important parameter for eigendecomposition-based algorithms [Biondi and Kostov, 1989, Kirlin, 1992]. As discussed in Kirlin [1992], if the number of wavefronts in the windowed data is greater than one, even with one wavefront perfectly flat, the vector $\mathbf{1}$ will not be proportional to an eigenvector of $\hat{\mathbf{R}}(\theta_k)$. We can, then, assume that the wavefronts have similar but slightly different parameter values, which will make only one wavefront to be flattened by the window based on its parameters. In this case, the other reflections will still appear in the windows, but as slightly incoherent interference. In consequence, we may assume that the dimension of the signal subspace of $\hat{\mathbf{R}}(\theta_k)$, $\mathbf{V}_s(\theta_k)$, is one, so that it consists only of the largest eigenvector of the correlation matrix. This is the second difference between the windowed method and the one in the previous chapter.

Note that, if the wavefronts are similar, as might happen when there is a multiple interfering with a primary, even eigenstructure methods may not resolve the two different events. Eigenstructure-based velocity spectra methods have higher resolution than methods like semblance, which are based on energy measures, but we must know, or estimate, as in Biondi and Kostov [1989], the number of wavefronts. If we overestimate or underestimate the signal subspace dimension, MUSIC coherence method will fail to estimate the right parameters. However, the assumption of a single event in each window seems reasonable, as indicated the numerical results in chapter 5.

3.5 The windowed spatial correlation matrix and semblance

In this section we show an interpretation of semblance in terms of writing seismic data in matrix notation. We have seen that seismic data is windowed in order to get aligned events, but we still need to measure coherence (alignment) of traces. Intuitively, we know that if there is an event, \mathbf{s} (the signal that repeats in all the traces, as illustrated in figure 3.2) is “large” and if there is no event, \mathbf{s} is “small”. So, a good coherence function would present large or small values in the presence or absence of an event.

Consider again the case of a window with correct parameters, resulting in a data matrix, also illustrated in figure 3.2, which can be approximated as $\mathbf{D}(\theta_k) = \mathbf{1}\mathbf{s}^H + \mathbf{N}$, where \mathbf{s} is the $N_t \times 1$ vector that contains the samples from the reflected wavelet. As before, if the noise is zero-mean and uncorrelated with \mathbf{s} , the spatial correlation matrix can be estimated by $\hat{\mathbf{R}}(\theta_k) \approx \frac{\|\mathbf{s}\|^2}{N_t} \mathbf{1}\mathbf{1}^T + \sigma_n^2 \mathbf{I}$, as shown in equation (3.4b).

As illustrated in Kirilin [1992], the semblance coherence function from equation (1.2) can be written in terms of the estimated spatial correlation matrix as

$$S_c = \frac{\mathbf{1}^T \hat{\mathbf{R}}(\theta_k) \mathbf{1}}{N_r \text{Tr}[\hat{\mathbf{R}}(\theta_k)]}, \quad (3.11)$$

where $\text{Tr}[\hat{\mathbf{R}}(\theta_k)]$ is the trace of the spatial correlation matrix and can be written in terms of its eigenvalues as

$$\text{Tr}[\hat{\mathbf{R}}(\theta_k)] = \sum_{i=1}^{N_r} \lambda_i. \quad (3.12)$$

The total energy of $\hat{\mathbf{R}}(\theta_k)$, in the denominator of (3.11), can also be written, with a few manipulations, as

$$\text{Tr}[\hat{\mathbf{R}}(\theta_k)] = N_r \left(\frac{\|\mathbf{s}\|^2}{N_t} + \sigma_n^2 \right). \quad (3.13)$$

By replacing equation (3.4b) in the numerator of equation (3.11), we have that

$$\mathbf{1}^T \hat{\mathbf{R}}(\theta_k) \mathbf{1} = \frac{\|\mathbf{s}\|^2 N_r^2}{N_t} + \sigma_n^2 N_r. \quad (3.14)$$

If we now replace equations (3.14) and (3.13) in equation (3.11), we can write semblance coefficient as

$$S_c \approx \frac{\frac{\|\mathbf{s}\|^2}{N_t} + \frac{\sigma_n^2}{N_r}}{\frac{\|\mathbf{s}\|^2}{N_t} + \sigma_n^2}. \quad (3.15)$$

From the analysis of equation (3.15), we can see that the extreme cases for semblance coefficient would be:

1. The absence of event, where $\|\mathbf{s}\|^2 = 0$, resulting in $S_c \approx \frac{1}{N_r}$.
2. The presence of event, where $\|\mathbf{s}\|^2 \gg \sigma_n^2$, which gives $S_c \approx 1$.

In that sense, semblance is a “good” coherence function, since it presents large values in the presence of a coherent signal and small ones when there is only noise. It is interesting to note that the lower value of semblance can be decreased by increasing of the number of traces.

Chapter 4

Proposed method

4.1 Overview

The implementation of MUSIC-based velocity spectra is the main focus of this work, for which we propose a number of improvements, presented in this chapter. In section 4.2 we apply the assumption of a single event in the seismic windowed data and compute S-MUSIC from equation (3.8b) with the signal subspace, as in Biondi and Kostov [1989] and Kirilin and Done [1999]. We then define a new way of computing MUSIC for seismic data, in terms of the temporal correlation matrix, namely T-MUSIC, in section 4.3. In section 4.4 we show how to compute iteratively S- and T-MUSIC with the power method and, finally, in section 4.5 we outline the importance of searching for MUSIC normalization functions and propose a new one, which we called semblance weighting MUSIC.

4.2 Signal and noise subspaces

We begin the simplification of MUSIC by noting that $\mathbf{V}(\theta_k)$ is unitary, so that

$$\mathbf{1}^T \mathbf{V}(\theta_k) \mathbf{V}^H(\theta_k) \mathbf{1} = \mathbf{1}^T \mathbf{1} = N_r. \quad (4.1)$$

On the other hand, $\mathbf{V}(\theta_k) = [\mathbf{V}_s(\theta_k) \ \mathbf{V}_n(\theta_k)]$, so that

$$\mathbf{1}^T \mathbf{V}(\theta_k) \mathbf{V}^H(\theta_k) \mathbf{1} = \mathbf{1}^T \mathbf{V}_s(\theta_k) \mathbf{V}_s^H(\theta_k) \mathbf{1} + \mathbf{1}^T \mathbf{V}_n(\theta_k) \mathbf{V}_n^H(\theta_k) \mathbf{1}. \quad (4.2)$$

Now, recall that $\mathbf{P}_n(\theta_k) = \mathbf{V}_n(\theta_k) \mathbf{V}_n^H(\theta_k)$ is the projection matrix onto the noise subspace. If we

let $\mathbf{P}_s(\theta_k) = \mathbf{V}_s(\theta_k)\mathbf{V}_s^H(\theta_k)$ be the projection matrix onto the signal subspace, then (4.2) yields

$$\mathbf{1}^T \mathbf{P}_n(\theta_k) \mathbf{1} + \mathbf{1}^T \mathbf{P}_s(\theta_k) \mathbf{1} = N_r. \quad (4.3)$$

We can, therefore, modify the MUSIC equation (3.8b), in order to use the signal subspace projection instead of the noise one [Biondi and Kostov, 1989, Kirilin and Done, 1999]. Recalling that the signal subspace is spanned by the largest eigenvector of $\hat{\mathbf{R}}(\theta_k)$, $\mathbf{V}_s(\theta_k) = \mathbf{v}_1$, we have

$$P_S(\theta_k) = \frac{N_r}{N_r - \mathbf{1}^T \mathbf{v}_1 \mathbf{v}_1^H \mathbf{1}}, \quad (4.4a)$$

$$= \frac{N_r}{N_r - |\mathbf{1}^T \mathbf{v}_1|^2}. \quad (4.4b)$$

The benefit of using the signal subspace projection is that the dimension of the signal subspace is smaller than the noise one (specially for the windowed data) and, due to that, the eigendecomposition that obtains the signal subspace has lower complexity than the one used to obtain the noise subspace. Computing, for the windowed correlation matrices, $\mathbf{1}^T \mathbf{P}_s(\theta_k) \mathbf{1}$ is also simpler than computing $\mathbf{1}^T \mathbf{P}_n(\theta_k) \mathbf{1}$, whenever that the signal subspace is smaller than the noise one.

4.3 The windowed temporal correlation matrix

Assuming that the digital seismic data $d_p(n)$ at the p -th receiver and n -th time sample has been windowed with the parameter θ_k , we can base the eigenstructure methods for parameter estimation in the temporal correlation matrix. If now, for the p -th receiver, we define the vector

$$\mathbf{d}_p = [d_p(1) \ d_p(2) \ \cdots \ d_p(N_t)], \quad (4.5)$$

with dimension $1 \times N_t$, we can define the temporal correlation matrix as

$$\mathbf{r}(\theta_k) = \mathbf{E}\{\mathbf{d}_p^H \mathbf{d}_p\} \quad (4.6a)$$

$$= \begin{bmatrix} \mathbf{E}\{d_p^*(1)d_p(1)\} & \mathbf{E}\{d_p^*(1)d_p(2)\} & \cdots & \mathbf{E}\{d_p^*(1)d_p(N_t)\} \\ \mathbf{E}\{d_p^*(2)d_p(1)\} & \mathbf{E}\{d_p^*(2)d_p(2)\} & \cdots & \mathbf{E}\{d_p^*(2)d_p(N_t)\} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{E}\{d_p^*(N_t)d_p(1)\} & \mathbf{E}\{d_p^*(N_t)d_p(2)\} & \cdots & \mathbf{E}\{d_p^*(N_t)d_p(N_t)\} \end{bmatrix}. \quad (4.6b)$$

The dimension of $\mathbf{r}(\theta_k)$ is $N_t \times N_t$, and it contains the correlation between different time samples of the windowed data. Recall that N_t is the number of samples in the window, which is usually smaller than the number of receivers, N_r . Thus, the dimension of $\mathbf{r}(\theta_k)$ is usually smaller than that of $\mathbf{R}(\theta_k)$.

Assuming spatially and temporally stationary signals within the window, the temporal correlation function defined as $r_d(p, m, n, k) = E\{d_p^*(n)d_m(k)\}$, for $p = m$, can also be written in terms of the difference $l = n - k$, resulting in $r_d(p, m, n, k) = r_d(0, l) \triangleq r_d(l)$. For writing $\mathbf{r}(\theta_k)$ in terms of $r_d(l)$, we also use the fact that $r_d(l) = r_d(-l)$, which results in

$$\mathbf{r} = \begin{bmatrix} r_d(0) & r_d(1) & \dots & r_d(N_t - 1) \\ r_d(1) & r_d(0) & \dots & r_d(N_t - 2) \\ \vdots & \vdots & \ddots & \vdots \\ r_d(N_t - 1) & r_d(N_t - 2) & \dots & r_d(0) \end{bmatrix}. \quad (4.7)$$

The temporal correlation matrix can also be estimated in terms of the sample correlation matrix as

$$\hat{\mathbf{r}}(\theta_k) = \frac{1}{N_r} \mathbf{D}^H(\theta_k) \mathbf{D}(\theta_k), \quad (4.8)$$

which can be viewed in terms of the seismic windowed data matrix as in figure 4.1. Using the same reasoning as in section 3.2, the estimated temporal correlation matrix can be approximated as

$$\hat{\mathbf{r}}(\theta_k) \approx \mathbf{s} \mathbf{s}^H + \sigma^2 \mathbf{I}. \quad (4.9)$$

Now, σ^2 is the noise variance for the temporal correlation matrix, not necessarily equal to σ_n^2 from the spatial correlation matrix. Note that we also disregard the cross terms from the multiplication, because of the assumption that the noise is zero-mean and uncorrelated with \mathbf{s} and that the time snapshots are independent.

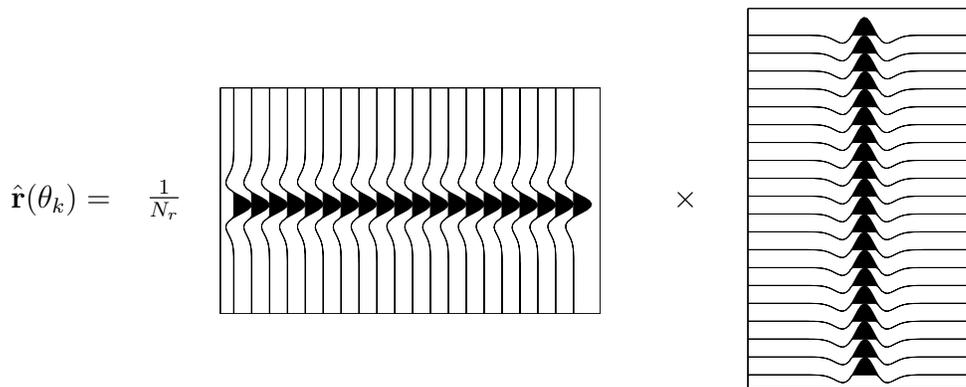


Figure 4.1: View of temporal correlation matrix in terms of seismic data matrix.

Now, the spatial correlation matrix is estimated as $\hat{\mathbf{R}}(\theta_k) = \frac{1}{N_t} \mathbf{D}(\theta_k) \mathbf{D}^H(\theta_k)$. We now show that both $\hat{\mathbf{r}}(\theta_k)$ and $\hat{\mathbf{R}}(\theta_k)$ have the same number of non-zero eigenvalues. In order to check how the eigenvectors and eigenvalues of $\hat{\mathbf{r}}(\theta_k)$ and $\hat{\mathbf{R}}(\theta_k)$ are related we can write, for $i = 1, \dots, \min(N_r, N_t)$,

$$\hat{\mathbf{R}}(\theta_k) \mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad (4.10a)$$

$$\frac{1}{N_t} \mathbf{D}(\theta_k) \mathbf{D}^H(\theta_k) \mathbf{v}_i = \lambda_i \mathbf{v}_i. \quad (4.10b)$$

Now, if we right-multiply both sides of equation (4.10b) by $(1/N_r) \mathbf{D}^H(\theta_k)$, we will have

$$\frac{1}{N_r} \mathbf{D}^H(\theta_k) \mathbf{D}(\theta_k) \mathbf{D}^H(\theta_k) \mathbf{v}_i = \frac{N_t}{N_r} \lambda_i \mathbf{D}^H(\theta_k) \mathbf{v}_i, \quad (4.11a)$$

$$\hat{\mathbf{r}}(\theta_k) \mathbf{D}^H(\theta_k) \mathbf{v}_i = \frac{N_t}{N_r} \lambda_i \mathbf{D}^H(\theta_k) \mathbf{v}_i. \quad (4.11b)$$

Clearly, the eigenvalues, κ_i , and eigenvectors, \mathbf{u}_i , of $\hat{\mathbf{r}}(\theta_k)$ are related to the eigenvalues, λ_i , and eigenvectors, \mathbf{v}_i , of $\hat{\mathbf{R}}(\theta_k)$ by $\mathbf{u}_i = \mathbf{D}^H(\theta_k) \mathbf{v}_i$ and $\kappa_i = (N_t/N_r) \lambda_i$. We expect that $\mathbf{1}$ is the largest eigenvector from $\hat{\mathbf{R}}(\theta_k)$ and, therefore, that $\mathbf{D}^H(\theta_k) \mathbf{1}$ is the largest eigenvector from $\hat{\mathbf{r}}(\theta_k)$.

In the absence of noise, we have, from equation (3.1), that $\mathbf{s} = \frac{1}{N_r} \mathbf{D}^H(\theta_k) \mathbf{1}$. In the presence of noise, we may have a good estimative of the wavelet \mathbf{s} if we compute $\hat{\mathbf{s}} = \frac{1}{N_r} \mathbf{D}^H(\theta_k) \mathbf{1}$. The vector $\hat{\mathbf{s}}$ can also be viewed as the mean value of the traces in a window, as illustrated in figure 4.2.

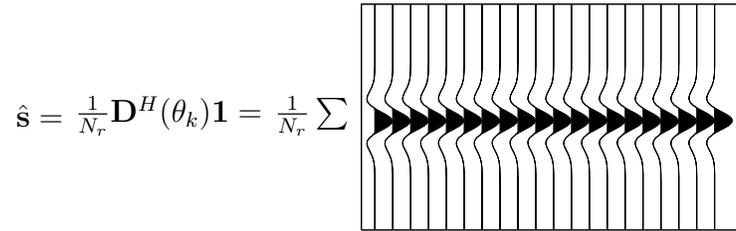


Figure 4.2: Diagram of the estimated seismic wavelet. The summation is done along the receivers dimension, for each column of the figure.

Now, for windowed seismic data, $\hat{\mathbf{r}}(\theta_k)$ will have the largest eigenvalue $\kappa_1 = (N_t/N_r) \lambda_1 \approx \|\mathbf{s}\|^2 + \sigma^2$ associated with eigenvector $\mathbf{u}_1 \approx \mathbf{s} / \|\mathbf{s}\|$. The average of the $(N_t - 1)$ remaining eigenvalues will be equal to σ^2 and they will be associated with eigenvectors orthogonal to \mathbf{s} . Therefore, instead of testing whether the all-ones vector, $\mathbf{1}$, is the largest eigenvector of $\hat{\mathbf{R}}(\theta_k)$, as is done in the usual S-MUSIC, we may test whether $\hat{\mathbf{s}} = \frac{1}{N_r} \mathbf{D}^H(\theta_k) \mathbf{1}$ is the largest eigenvector of $\hat{\mathbf{r}}(\theta_k)$. The advantage is that, in the second case, we have to compute the eigenvectors of a smaller matrix. We have, then, that for each window formed with the parameters τ_0 and v_k , we will test, using an eigenstructure-based coherence measure, if $\hat{\mathbf{s}}$ is aligned to the largest eigenvector from matrix $\hat{\mathbf{r}}(\theta_k)$. The MUSIC spectrum, based on the temporal correlation matrix, known as T-MUSIC [Wang et al., 2001], can be

obtained as

$$P_T(\theta_k) = \frac{\hat{\mathbf{s}}^H \hat{\mathbf{s}}}{\hat{\mathbf{s}}^H \mathbf{U}_n(\theta_k) \mathbf{U}_n^H(\theta_k) \hat{\mathbf{s}}}, \quad (4.12a)$$

$$= \frac{\|\hat{\mathbf{s}}\|^2}{|\hat{\mathbf{s}}^H \mathbf{U}_n(\theta_k)|^2}, \quad (4.12b)$$

where, now, we have that $\mathbf{U}_n(\theta_k)$ is the matrix formed with the eigenvectors that span the noise subspace of $\hat{\mathbf{r}}(\theta_k)$. We can also use the eigenvector that spans the signal subspace and write the MUSIC coherence function as

$$P_T(\theta_k) = \frac{\hat{\mathbf{s}}^H \hat{\mathbf{s}}}{\hat{\mathbf{s}}^H \hat{\mathbf{s}} - \hat{\mathbf{s}}^H \mathbf{U}_s(\theta_k) \mathbf{U}_s^H(\theta_k) \hat{\mathbf{s}}}, \quad (4.13a)$$

$$= \frac{\|\hat{\mathbf{s}}\|^2}{\|\hat{\mathbf{s}}\|^2 - |\hat{\mathbf{s}}^H \mathbf{u}_1|^2}. \quad (4.13b)$$

The benefit of writing T-MUSIC in terms of the signal subspace is that, as for S-MUSIC, we only need to compute one eigenvector to obtain the coherence function. In the next section we will introduce a method to estimate the largest eigenvector from both the spatial and the temporal correlation matrices.

4.4 Power Method

One of the most well-known methods to calculate the largest eigenvalues and eigenvectors of a matrix is the power method [Golub and Van Loan, 1996]. It starts with an initial vector, and iteratively updates this vector, until it converges to the largest eigenvector of the matrix. The main benefit of using a coherence function, as in equations (4.4b) and (4.13b), based on the signal subspace of seismic windowed data, is that, as we assume that the signal subspace has dimension one, we only need to estimate the largest eigenvector of $\hat{\mathbf{R}}(\theta_k)$ and $\hat{\mathbf{r}}(\theta_k)$. In that way, the power method (PM) suits perfectly the problem. Next, we will discuss how to initialize the PM and to use its result for S-MUSIC and T-MUSIC computation.

4.4.1 MUSIC with Power Method for the windowed spatial correlation matrix

As mentioned before, if we use a coherence measure based on the signal subspace of the spatial correlation matrix, as in equation (4.4b), we only need to estimate the largest eigenvector $\mathbf{v} = \mathbf{v}_1$ of the correlation matrix. As we suspect that this eigenvector should be proportional to $\mathbf{1}$, we initialize

the power method with the vector $\mathbf{v}^{(0)} = \mathbf{1}$. For the n -th iteration, the estimated eigenvector will be:

$$\mathbf{v}^{(n)} = \frac{\hat{\mathbf{R}}(\theta_k)\mathbf{v}^{(n-1)}}{\|\hat{\mathbf{R}}(\theta_k)\mathbf{v}^{(n-1)}\|} \quad (4.14)$$

The stopping criterion for the power method is based on the difference between consecutive estimates of the eigenvector, $\mathbf{v}^{(n)}$. We say that $\hat{\mathbf{v}} = \mathbf{v}^{(n)}$ is the desired eigenvector if

$$\|\mathbf{v}^{(n)} - \mathbf{v}^{(n-1)}\| < \xi, \quad (4.15)$$

where ξ is a threshold value, which controls the desired precision of the algorithm. For the right parameters, the proposed initialization yields that usually only one iteration is enough for convergence, as indicated in simulations.

The S-MUSIC spectrum, combined with the power method (PM-S-MUSIC) can be written as:

$$P_S(\theta_k) = \frac{N_r}{N_r - |\mathbf{1}^T \hat{\mathbf{v}}|^2} \quad (4.16)$$

PM-S-MUSIC has lower complexity than the conventional S-MUSIC from equation (3.8b), since S-MUSIC requires a full eigendecomposition, with the computation of $N_r - 1$ eigenvectors, which results in a complexity of order $O(N_r^3)$. Its complexity, however, is larger than that of semblance. This is illustrated in Table 4.1, where we show the approximate number of calculations needed for the semblance and PM-S-MUSIC algorithms. In this Table, n represents the number of iterations of the power method. In general, the convergence of PM-S-MUSIC is very fast (as will be illustrated by simulations, in chapter 5), which implies that $N_r \gg n$ and $N_t \gg n$. By considering only the number of multiplications from Table 4.1, we can see that, asymptotically, PM-S-MUSIC has order $O(N_r^2)$, while semblance has order $O(N_r N_t)$. Because $N_t < N_r$, semblance is faster

4.4.2 MUSIC with Power Method for the windowed temporal correlation matrix

For the temporal correlation matrix, we know that, when there is an event in the window, the largest eigenvector is approximately aligned with $\frac{1}{N_r} \mathbf{D}^H(\theta_k) \mathbf{1}$, so a good initialization for PM-T-MUSIC is $\mathbf{u}^{(0)} = \hat{\mathbf{s}} = \frac{1}{N_r} \mathbf{D}^H(\theta_k) \mathbf{1}$. The n -th iteration of PM-T-MUSIC will give the estimated eigenvector:

$$\mathbf{u}^{(n)} = \frac{\hat{\mathbf{r}}(\theta_k)\mathbf{u}^{(n-1)}}{\|\hat{\mathbf{r}}(\theta_k)\mathbf{u}^{(n-1)}\|}. \quad (4.17)$$

Table 4.1: Number of Operations for semblance and PM-MUSIC Algorithms

Algorithm	Additions	Multiplications	Complexity
Semblance	$2[(N_r - 1) + N_t]$	$N_r N_t$	$O(N_r N_t)$
S-MUSIC			$O(N_r^3)$
T-MUSIC			$O(N_t^3)$
PM-S-MUSIC	$n(N_r^2 + 2N_r) + N_r$	$n(N_r^2 + 2N_r) + N_r$	$O(N_r^2)$
PM-T-MUSIC	$N_r N_t + n(N_t^2 + 2N_t) + N_t$	$N_r N_t + n(N_t^2 + 2N_t) + 2N_t^2$	$O(N_r N_t)$

We assume the same convergence criterion as PM-S-MUSIC, *i.e.*, $\hat{\mathbf{u}} = \mathbf{u}^n$ if

$$\|\mathbf{u}^{(n)} - \mathbf{u}^{(n-1)}\| < \xi. \quad (4.18)$$

After convergence, the PM-T-MUSIC is computed as

$$P_T(\theta_k) = \frac{\|\hat{\mathbf{s}}\|^2}{\|\hat{\mathbf{s}}\|^2 - |\hat{\mathbf{s}}^H \hat{\mathbf{u}}|^2}. \quad (4.19)$$

For the temporal correlation matrix, T-MUSIC requires the computation of $N_t - 1$ eigenvectors, resulting in a complexity of order $O(N_t^3)$. Besides the computational savings obtained by using only one eigenvector, we obtain further savings because the dimension of $\hat{\mathbf{r}}(\theta_k)$ is usually smaller than the one of $\hat{\mathbf{R}}(\theta_k)$, in other words, usually $N_t < N_r$. Since we assume this condition, the number of multiplications of PM-T-MUSIC, presented in Table 4.1 will be dominated by the term $N_r N_t$, resulting in a complexity of $O(N_r N_t)$, also illustrated in Table 4.1. Thus, both semblance and PM-T-MUSIC have complexities of same order. T-MUSIC also has the advantage of requiring less processing than S-MUSIC, since it does not require the computation of spatial smoothing and FB averaging.

4.5 Velocity spectra normalization

One of the main complains about MUSIC-based velocity spectra is that it yields arbitrary amplitude values, since it is a measure of the orthogonality between two vectors from orthogonal subspaces. Such behavior makes the simple replacement of semblance with MUSIC as a coherency measure, for example, in standard velocity analysis, not adequate, even though MUSIC presents a higher resolution.

On the other hand, semblance is a very robust coherence function and produces normalized values between 0 and 1. It would be very interesting to develop a type of normalization which combines semblance and MUSIC values, in an attempt to get a velocity spectrum without those arbitrary amplitude values. For instance, we may get events with coherency measures differing by several orders

of magnitude, which makes visualizing the spectrum very difficult.

To that end, we present, in this section, two normalization functions, which make use of semblance coefficient and can be used either for S- and T-MUSIC. The first one, proposed by Asgedon et al. [2011] and called semblance balancing, uses the energy of both semblance and MUSIC. The second one, named semblance weighting, is a proposal of this work and uses the concept of coherence multiplication, as presented in Abbad and Ursin [2012].

4.5.1 Semblance Balancing

In Asgedon et al. [2011], a normalization function that makes use of semblance, named semblance balancing, is proposed. It assumes the knowledge of the velocity spectra, computed with both MUSIC and semblance, and balances the MUSIC coefficient according to the semblance and the MUSIC energies, in a given time window. We can obtain semblance-balanced MUSIC by assuming, for a given CMP location, the possession of semblance and MUSIC coherence values for N_τ zero-offset time samples, and N_v trial velocities. If we denote $m_{i,j}$ and $s_{i,j}$ as the coherence values obtained from MUSIC and semblance, respectively, for the i -th ZO time sample and j -th trial velocity, semblance-balanced MUSIC, $\hat{m}_{i,j}$, will be given by

$$\hat{m}_{i,j} = \sqrt{\frac{A_{s,i}}{A_{m,i}}} m_{i,j}, \quad (4.20)$$

where, for a window with an odd number of L ZO time samples,

$$A_{s,i} = \sum_{l=i-(L-1)/2}^{i+(L-1)/2} \sum_{k=1}^{N_v} s_{lk}^2, \quad A_{m,i} = \sum_{l=i-(L-1)/2}^{i+(L-1)/2} \sum_{k=1}^{N_v} m_{lk}^2. \quad (4.21)$$

Application of the above conditioning makes sure that the amplitude anomalies inherent to the original MUSIC velocity spectrum are balanced according to the energy level of semblance.

4.5.2 Semblance Weighting

In Abbad and Ursin [2012], a high-resolution bootstrapped differential semblance coherence function is presented, which uses the multiplication of several coherence functions. This idea can be used to combine MUSIC and semblance measures, so as to map the MUSIC values to a well-defined range, while keeping its resolution. This can be accomplished with the use of a weighting of the MUSIC coherence function ($m_{i,j}$) by semblance ($s_{i,j}$). First, we normalize the MUSIC measure, ensuring that, for each τ_0 , its maximum value is one. To that end, we define a new coherence function, $m_{i,j}^{(1)}$,

by dividing $m_{i,j}$ by its maximum value, for the i -th ZO time:

$$m_{i,j}^{(1)} = \frac{m_{i,j}}{\max\{m_{i,j}\}_{j=1}^{N_v}}. \quad (4.22)$$

The normalized function, $m_{i,j}^{(1)}$, has a disadvantage of amplifying the noise, since there will be a coherence value equal to 1 for each τ_0 , even when there is no event. Now, if we use semblance as a weighting factor, we can decrease the amplified noise and still maintain some of the robustness of semblance. The semblance weighting normalization of MUSIC can be written, then, as

$$m_{i,j}^{(2)} = s_{i,j} \times m_{i,j}^{(1)}. \quad (4.23)$$

As will be shown in chapter 5, semblance weighting presents amplitude values between 0 and 1, without the amplitude anomalies found in MUSIC.

Chapter 5

Numerical examples

5.1 Overview

In this chapter we present the numerical examples that we performed for comparison between MUSIC-based velocity spectra calculated from spatial and temporal correlation matrices and also to evaluate the velocity spectra computed with the power method and normalized with both semblance weighting and semblance balancing functions. In section 5.2 we present simulations with a simple synthetic CMP gather, containing two reflections. In section 5.3 we present a numerical example with real marine data set from the Jequitinhonha Basin, in Brazil. Finally, we show, for synthetic data, a stacking example for velocities obtained with different coherence functions, in section 5.4.

5.2 Synthetic CMP data with two reflections

In this section we compare the use of spatial and temporal correlation matrices to obtain high-resolution velocity spectra for a synthetic data. We use the semblance coherence function as a benchmark. We also show high-resolution velocity spectra obtained with the power method, for both spatial and temporal correlation matrices.

In the simulations, we used a simple synthetic model with two reflections, generated by equation (1.1). The first one has a zero offset travelttime of 1 s and a velocity of 4000 m/s; the second one has a zero offset travelttime of 1.06 s and a velocity of 4500 m/s. Both reflections are modeled by a zero-phase Ricker wavelet, with a dominant frequency of 25 Hz and are fully correlated. The CMP section contains 64 receivers. The offset of the first one is 80 m and the distance between them is also 80 m. The sample period is 2 ms and white Gaussian noise was added to the data. The CMP section can be viewed in figure 5.1.

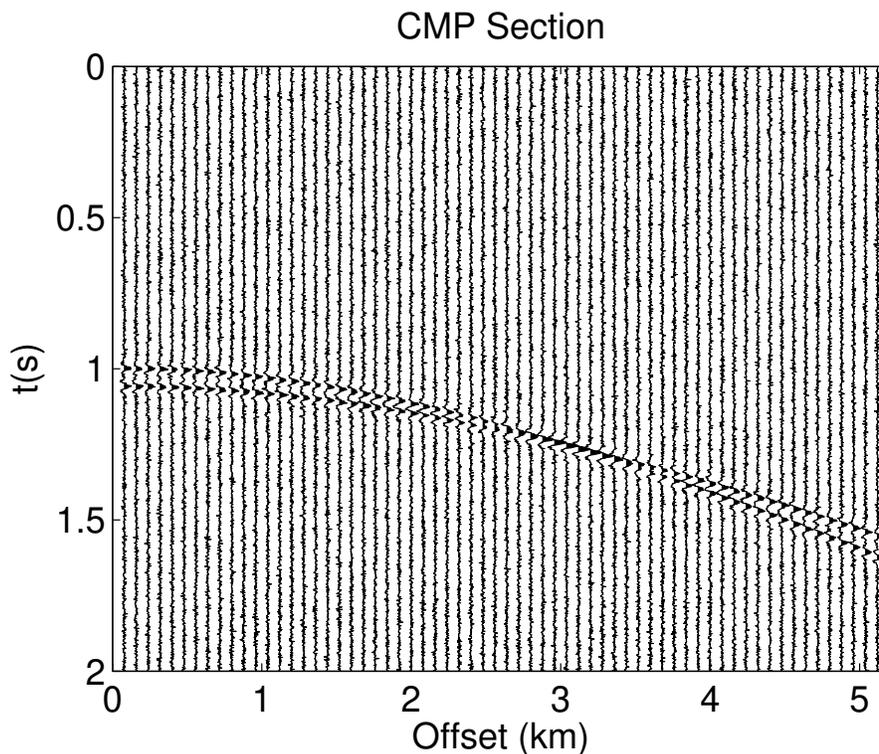


Figure 5.1: Synthetic CMP section.

5.2.1 Comparison of the methods

Figures 5.2, 5.3 and 5.4 show velocity spectra calculated with semblance from equation (1.2) and MUSIC from equations (4.4b) and (4.13b), using spatial and temporal correlation matrices, respectively. The figures show both the full three-dimensional spectrum and a zoom into the two-dimensional spectrum part close to the actual values of the parameters. The window size used was $N_t = 19$ samples and velocities were tested from 3000 m/s to 6000 m/s, in increments of 10 m/s. We have assumed that the signal subspace has rank one, *i.e.*, it is formed by a single signal. For the spatial correlation matrix, we performed spatial smoothing, described in 2.4.1, using 47 subarrays, each consisting of 18 receivers. We have also used forward-backward averaging, described in 2.4.1, in the spatially smoothed version of the spatial correlation matrix.

The white squares in the velocity spectra indicate the true location of the parameters and the MUSIC velocity spectra in figures 5.3 and 5.4 are normalized with respect to the largest coherence value of the corresponding spectrum. The results in the figures clearly show that both MUSIC algorithms outperform semblance in terms of resolution, resulting in more precise velocity estimates. It is possible to observe a large spreading in semblance function, a small spreading in S-MUSIC and almost no spreading in T-MUSIC. These results indicate that MUSIC with the temporal correlation

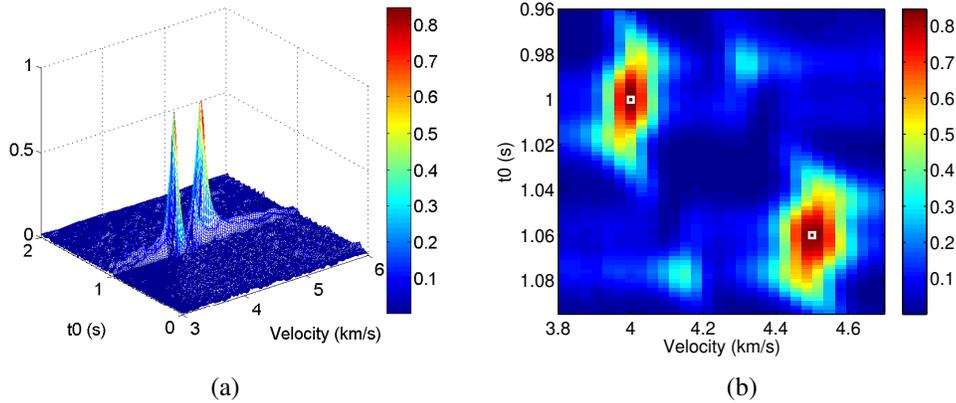


Figure 5.2: Semblance velocity spectrum in 3D (a) and in 2D close view (b).

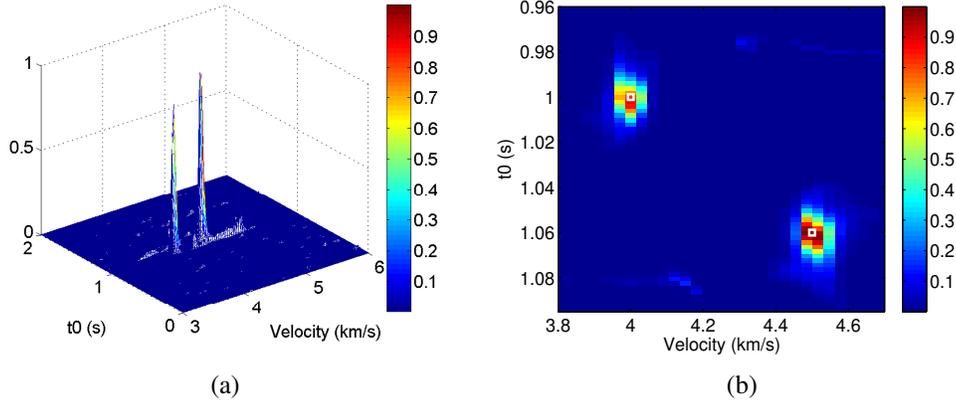


Figure 5.3: S-MUSIC velocity spectrum in 3D (a) and in 2D close view (b).

may present even better resolution than the spatial correlation one, in time and velocity, despite its lower complexity.

5.2.2 Power Method

Figures 5.5 and 5.6 show, for the same CMP section of the previous example, velocity spectra calculated with PM-MUSIC for spatial and temporal correlation matrices, normalized with respect to the largest coherence value of the corresponding spectrum. Clearly, the use of the power method has no impact on the results, when compared to figures 5.3 and 5.4. However, S-MUSIC measurements seem to be noisier with the power method, especially in the areas with no events.

In figure 5.7, we show the histogram of the number of iterations needed for PM-MUSIC convergence, for both spatial and temporal correlation matrices example. We used $\xi = 0.3$ in the simulations. As seen in figure 5.7(a), for the spatial correlation matrix, in 84.71% of the cases the power

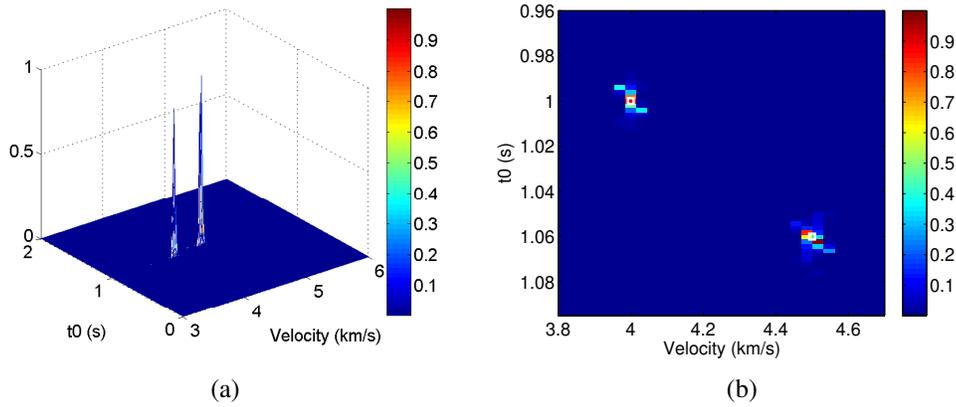


Figure 5.4: T-MUSIC velocity spectrum in 3D (a) and in 2D close view (b).

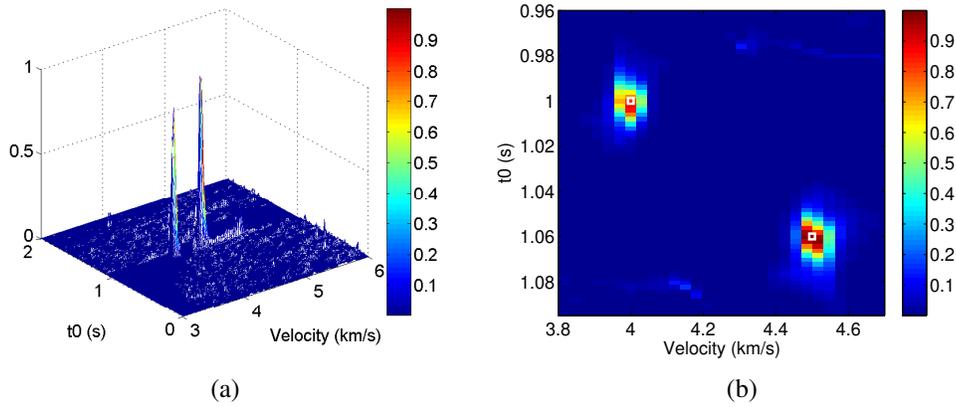


Figure 5.5: S-MUSIC velocity spectrum computed with power method in 3D (a) and in 2D close view (b).

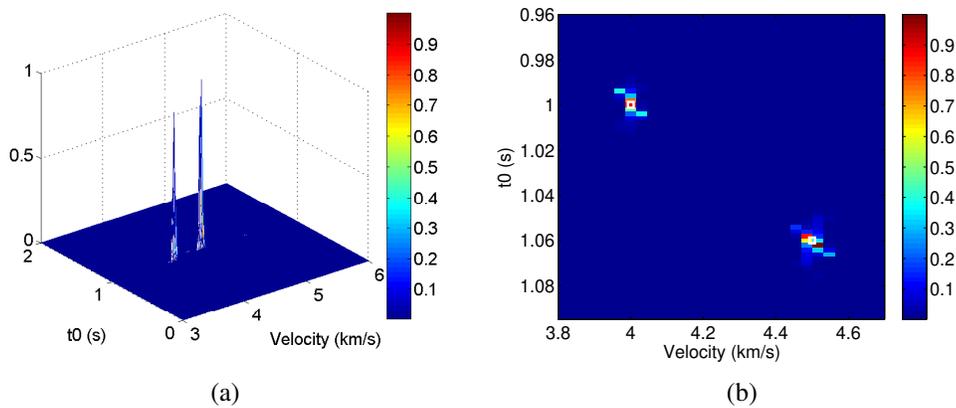


Figure 5.6: T-MUSIC velocity spectrum computed with power method in 3D (a) and in 2D close view (b).

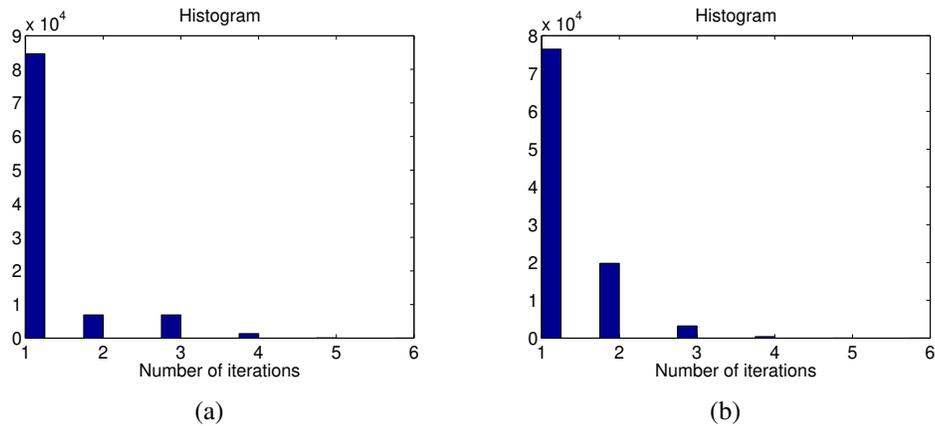


Figure 5.7: Histogram of number of iterations for convergence of PM-MUSIC applied on spatial (a) and temporal (b) correlation matrices.

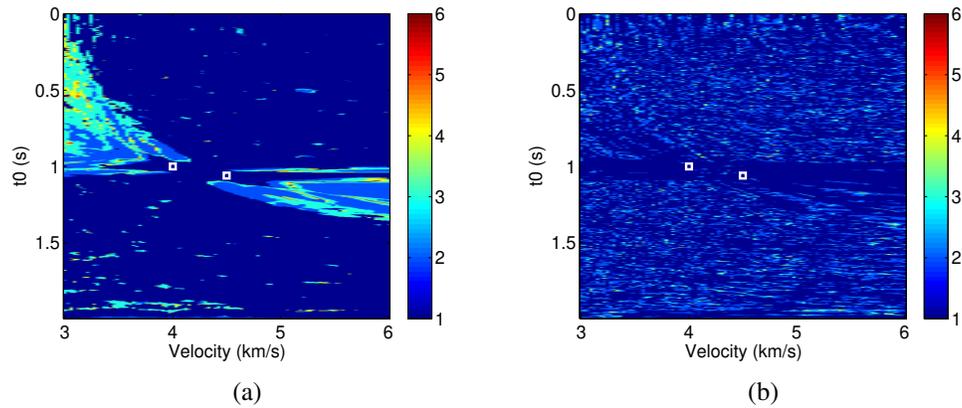


Figure 5.8: Images with the number of iterations of PM-MUSIC for velocity spectra obtainment with spatial (a) and temporal (b) correlation matrices.

method converged in a single iteration. For the temporal correlation matrix, figure 5.7(b) shows that PM-MUSIC converged in one iteration in 76.53% of the cases. Both figures illustrate that the power method converges quickly, hardly ever requiring more than three iterations.

In figure 5.8, we show images with the number of iterations performed by PM-MUSIC, for each point of the velocity spectra. It is possible to see that when we are close to the true stacking parameters the number of iterations is small.

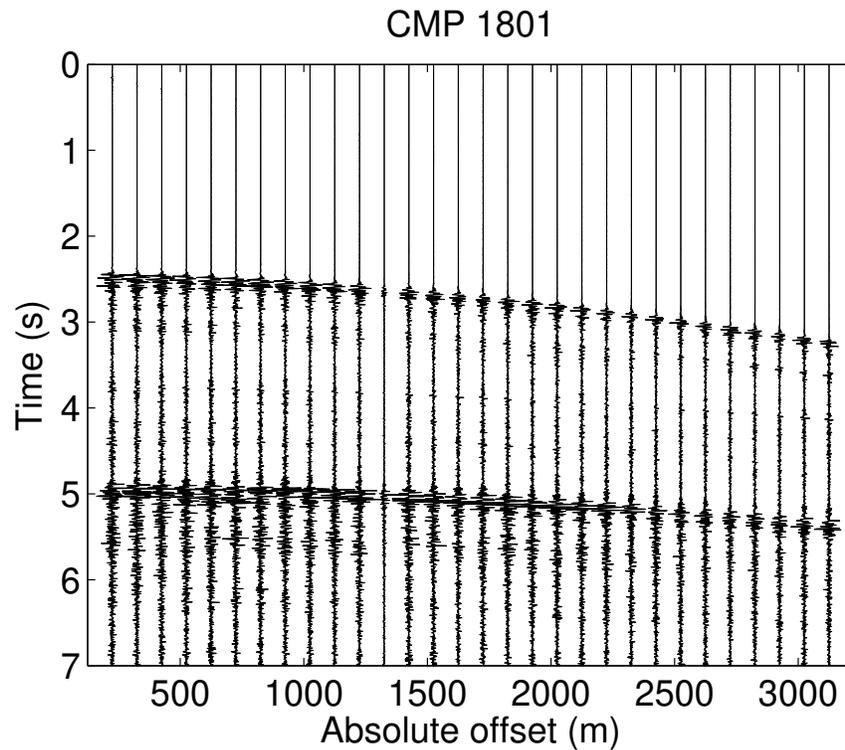


Figure 5.9: CMP 1801 from marine data set.

5.3 Real data set: Jequitinhonha Basin

For the numerical examples presented in this section, we used the real marine data set, acquired in the Jequitinhonha Basin and provided by PETROBRAS¹. It consists of 981 shots with a distance of 25 m between consecutive shots. For each shot there are 120 receivers displaced with intervals of 25 m. The data has a sample period of 4 ms and the total time of recording was 7 s.

In figure 5.9 we show the pre-processed² CMP gather 1801, which was used in the numerical examples presented in this section. This CMP gather contains 30 receivers. Figure 5.10 illustrates the velocity spectrum computed with semblance coefficient, using a window of $N_t = 15$ samples. Semblance function was computed for each time sample, with velocities going from 1000 m/s to 3000 m/s, increased of 20 m/s.

5.3.1 Comparison of the methods

Figure 5.11 illustrates the velocity spectra obtained with S-MUSIC and T-MUSIC, both computed with a window of $N_t = 15$ samples, using semblance weighting normalization function. In figure 5.12

¹Line 214-2660 - azimuth 50° N, acquired towards the NE.

²This data was pre-processed with band-pass filters, geometrical amplitude corrections, deconvolution, etc.

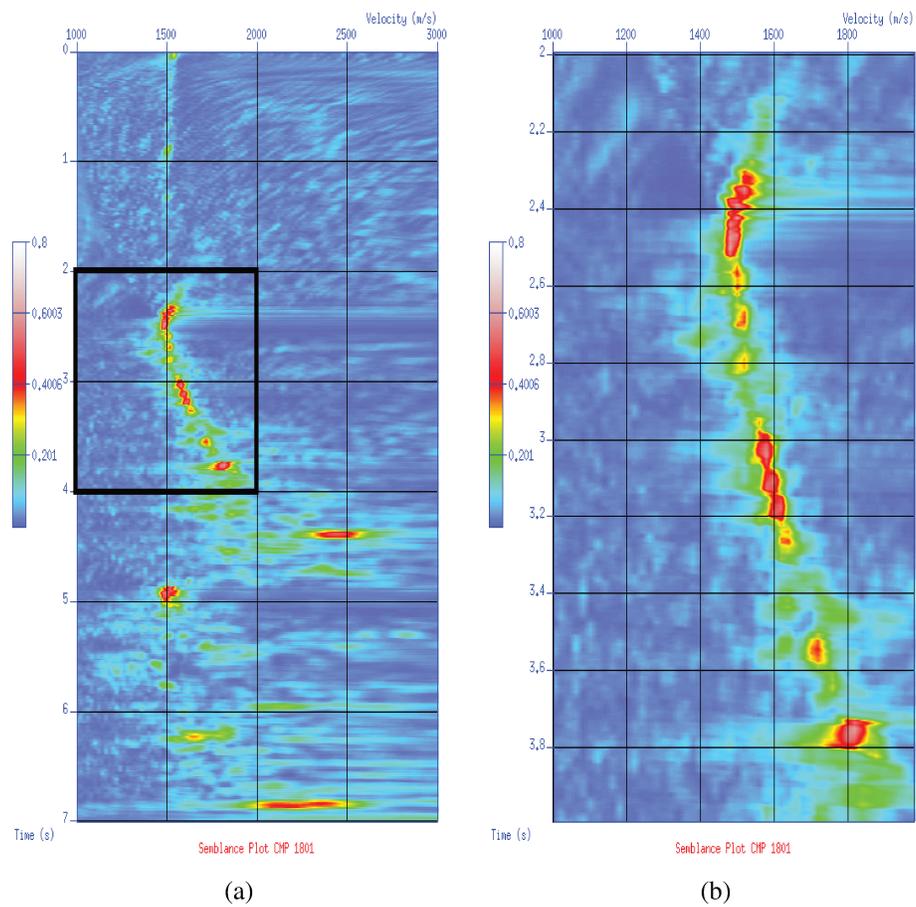


Figure 5.10: Semblance velocity spectrum from CMP 1801 (a) and close view of the region inside the black box (b).

we also show S-MUSIC and T-MUSIC velocity spectra, also computed with $N_t = 15$ samples, but normalized with semblance balancing, applied for a window of $L = 3$ samples. For S-MUSIC, we performed FB averaging and spatial smoothing of the correlation matrix using 15 subarrays of 16 receivers.

We can see that, for this data, semblance weighting normalization has a better preservation of the coherence. We can also see that both S- and T-MUSIC present higher resolution than semblance. However, the spectra computed with T-MUSIC seem a little noisier than the ones computed with S-MUSIC. The cause of that might be the equally averaging of the traces in the estimative of the tested eigenvector, as illustrated in figure 4.2. A possible improvement of T-MUSIC could be a better estimative of this eigenvector, taking into account the different amplitudes and shapes of the wavelets along the traces.

5.3.2 Power Method

In figures 5.13 and 5.14 we show S- and T-MUSIC normalized with semblance weighting and semblance balancing functions and computed with the power method, for the same window of $N_t = 15$ samples. When comparing these figures with figures 5.11 and 5.12 we can see that power method maintains the high-resolution property of MUSIC, with the benefit of presenting lower computation complexity, as illustrated in Table 4.1.

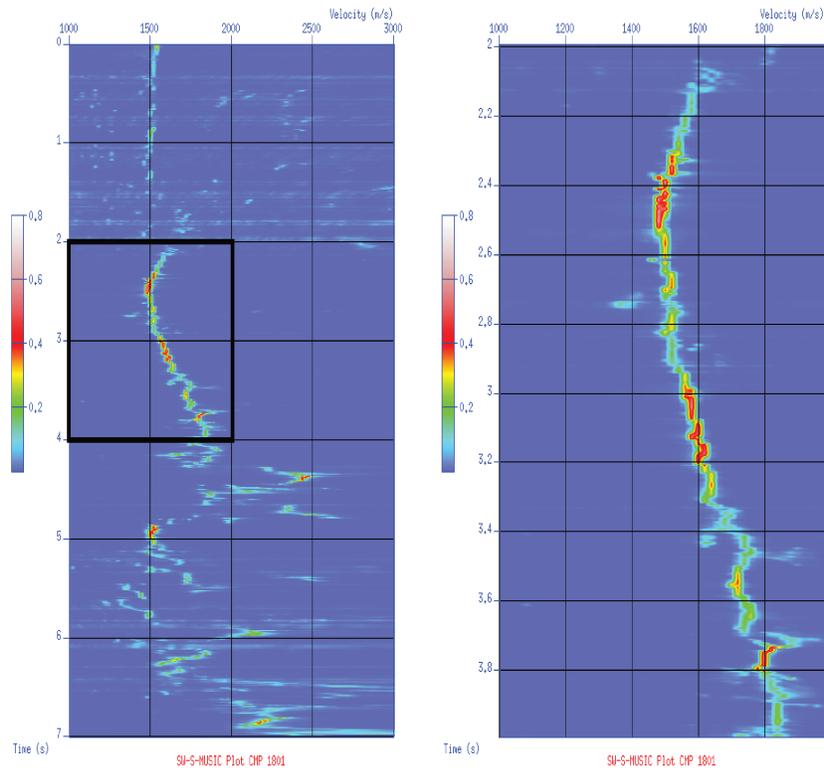
Figure 5.15 presents the histogram of the number of iterations needed for PM-MUSIC convergence, for both spatial and temporal correlation matrices example. We also used $\xi = 0.3$. In this example, for the spatial correlation matrix, in 47% of the cases the power method converged in a single iteration. For the temporal correlation matrix, the PM-MUSIC converged in one iteration in 89% of the cases. In figure 5.16, we show images with the number of iterations performed by PM-MUSIC, for each point of the velocity spectra.

5.4 Stacking of synthetic data

In this section we present the stacking results of the synthetic data³ generated with the velocity model from figure 5.17(a). The velocity values in figure 5.17 vary from 1500 m/s to 3500 m/s. This example demonstrates how the high-resolution coherence functions affect the stacking of regions with pinch-outs, as the one illustrated in the red box of figure 5.17(a).

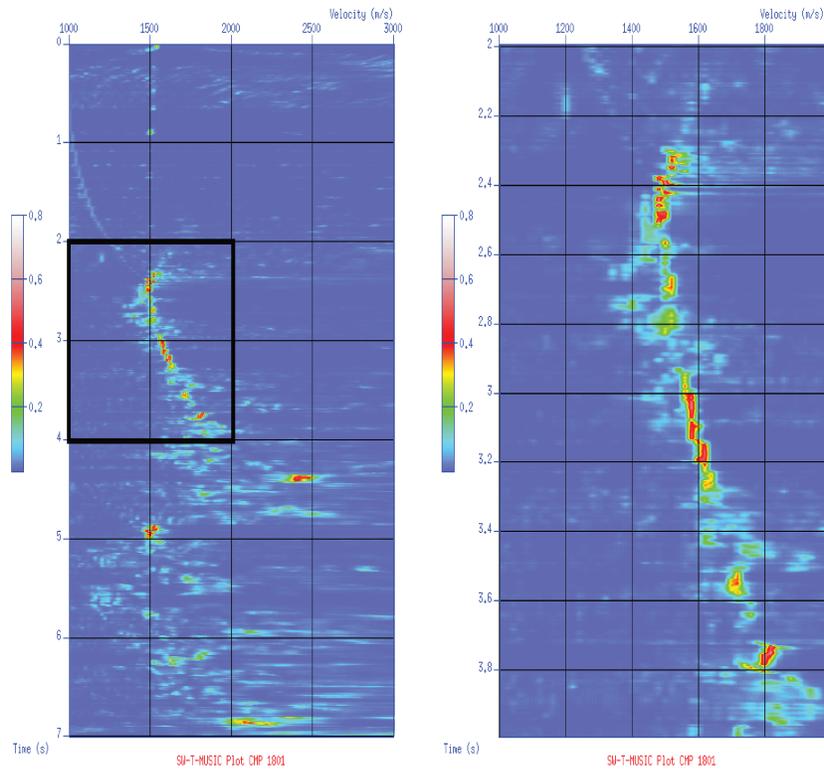
The data contains 179 shots with 180 receivers, displaced at intervals of 10 m. The distance between the source and the first receiver is 10 m. The first shot was performed at the offset of 10 m

³The data was generated by a ray-tracing program kindly provided by Prof Hervé Perroud, from the Université de Pau et des Pays de l'Adour.



(a)

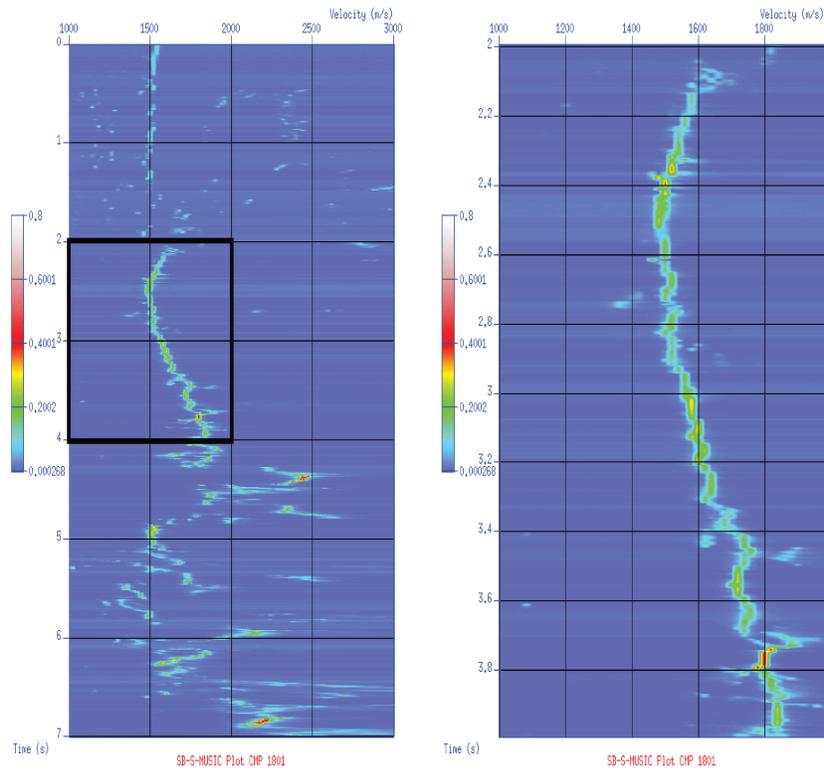
(b)



(c)

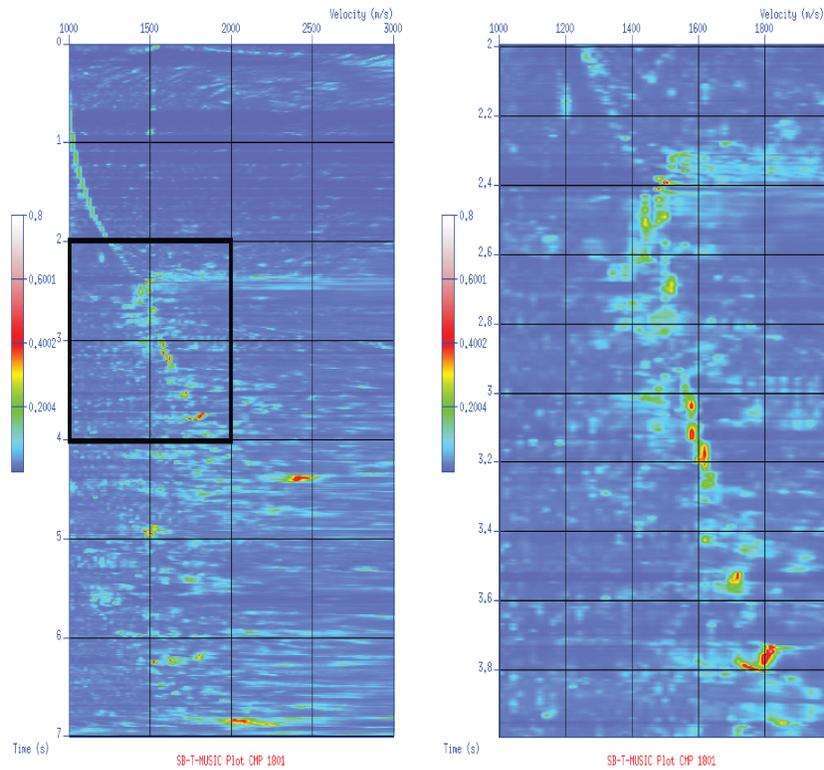
(d)

Figure 5.11: S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with SW. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra.



(a)

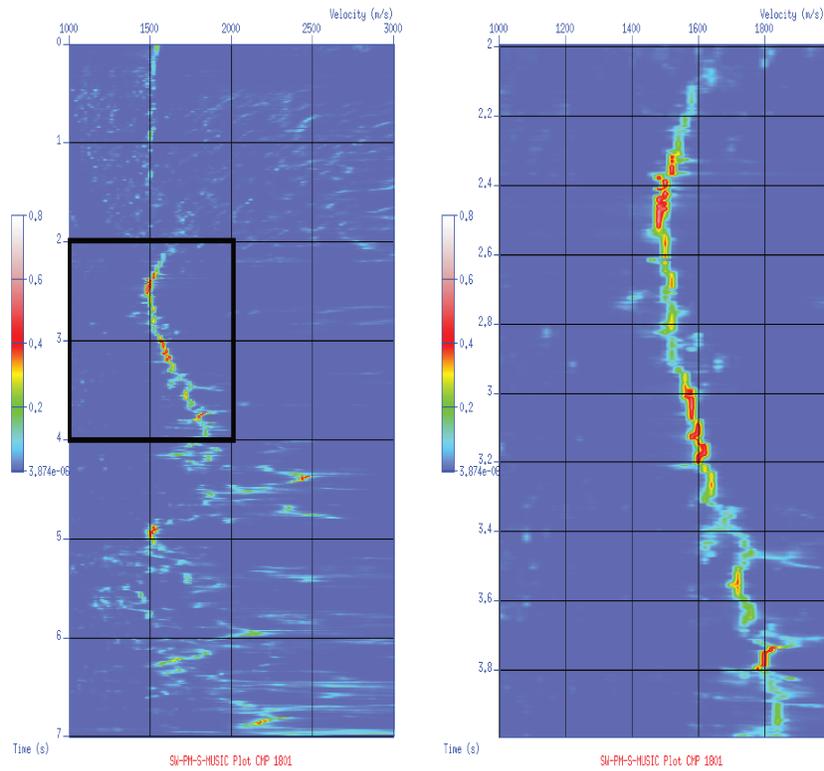
(b)



(c)

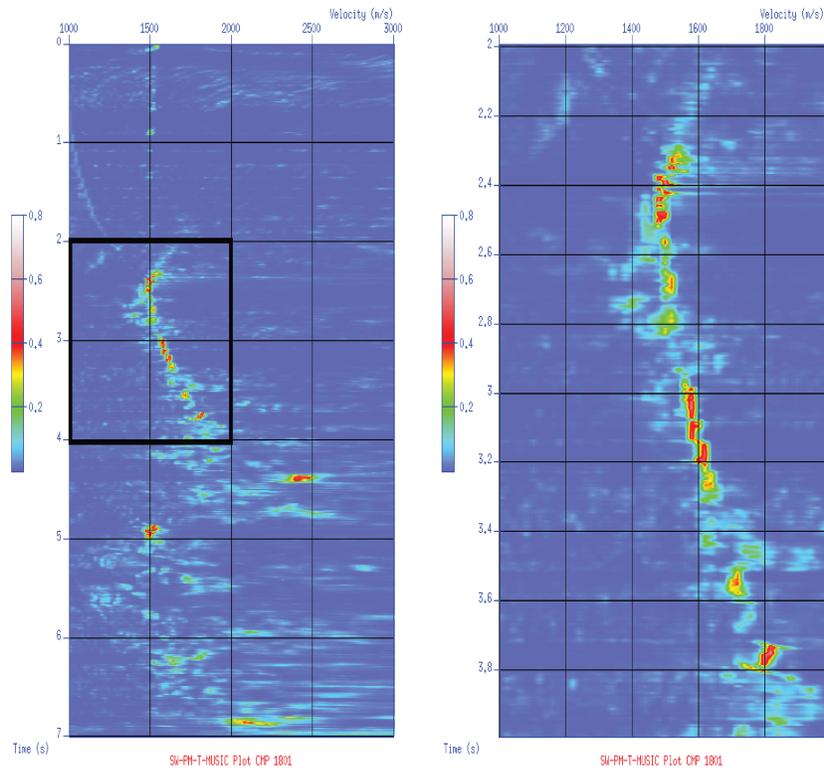
(d)

Figure 5.12: S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with SB. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra.



(a)

(b)



(c)

(d)

Figure 5.13: S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with the power method and normalized with SW. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra.

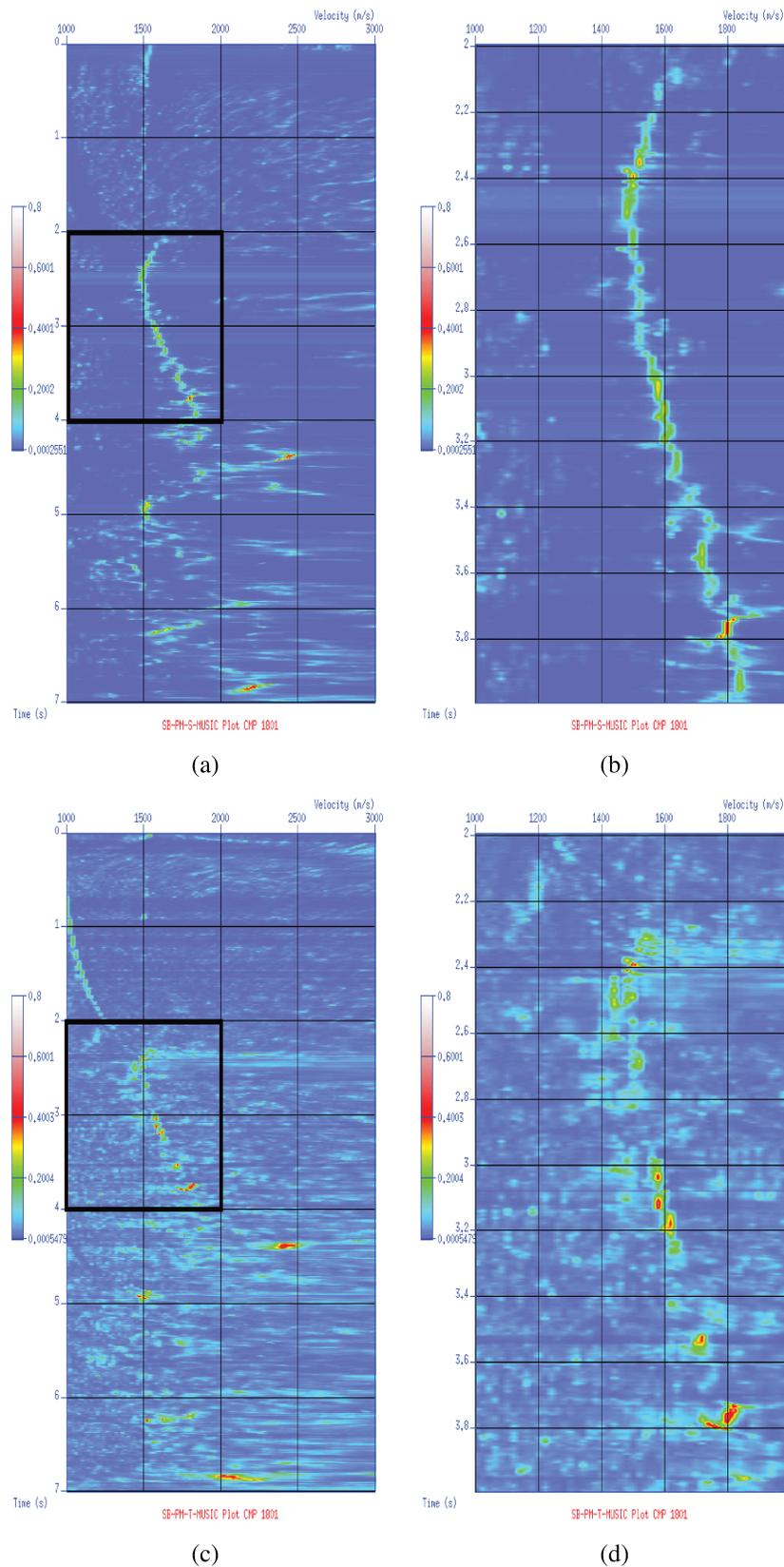


Figure 5.14: S-MUSIC (a) and T-MUSIC (c) velocity spectra computed with the power method and normalized with SB. Close view of the region inside the black box for S-MUSIC (b) and T-MUSIC (d) velocity spectra.

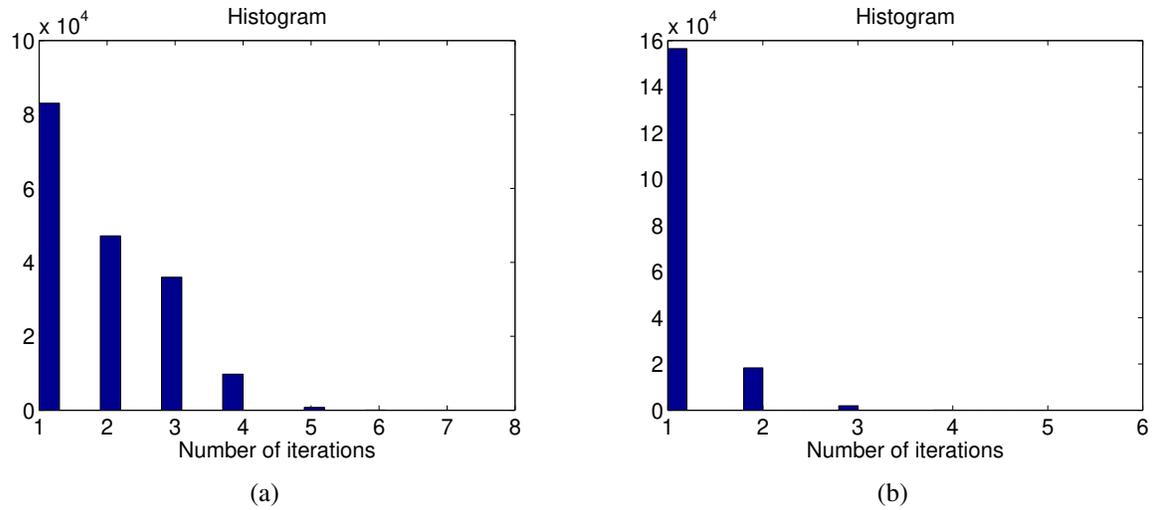


Figure 5.15: Histogram of number of iterations for convergence of PM-MUSIC applied on spatial (a) and temporal (b) correlation matrices.

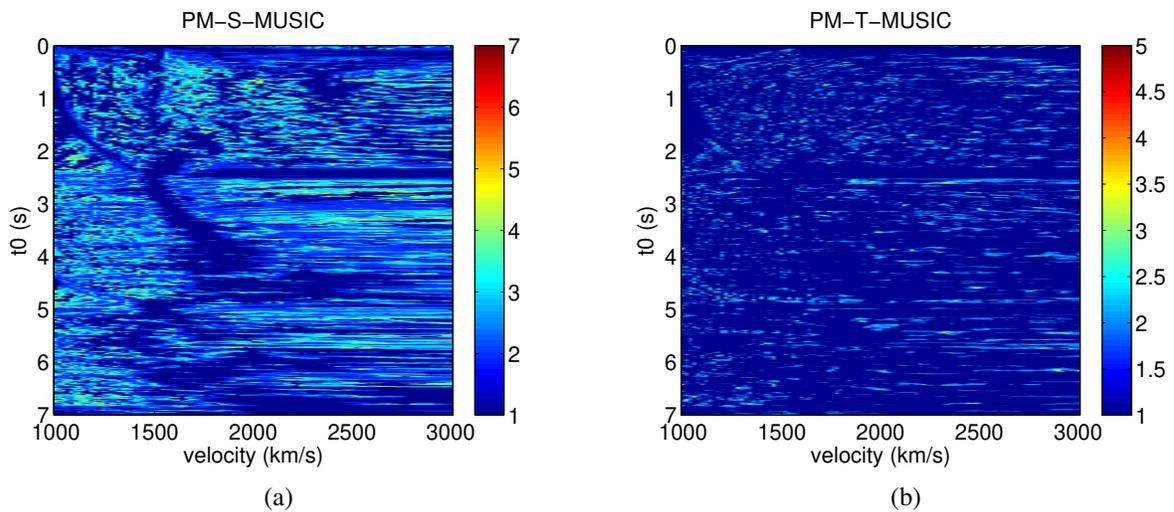


Figure 5.16: Images with the number of iterations of PM-MUSIC for velocity spectra obtainment with spatial (a) and temporal (b) correlation matrices.

and the last one at 1790 m. Each shot was increased of 10 m in relation to the last one. The data has a sample period of 2 ms and the total time of recording was 2 s. We stacked 101 CMP gathers, going from offset 1600 m to 2000 m. The fold of the CMP's is of 90 receivers. In the coherence computations, we tested 151 velocities, going from 1000 m/s to 4000 m/s, with a window of $N_t = 15$ samples. We performed FB averaging and spatial smoothing with 61 sub-arrays of 30 receivers for S-MUSIC. We added white Gaussian noise to the data. For the stacking operation, we used, for each time sample, the velocity which presented the highest coherency value.

Figures 5.18, 5.19 and 5.20 illustrate the CMP's located at the midpoints of 1600 m, 1800 m and 1950 m (the last one being the exact location of the pinch-out, where there are more than one reflection superposed arriving around 0.3 s) with different velocity spectra computed with semblance, PM-S-MUSIC and PM-T-MUSIC with and without normalization functions. For a better visualization, the non-normalized PM-S-MUSIC and PM-T-MUSIC spectra were clipped, respectively, for coherence values larger than 300 and 700 in figure 5.18, 250 and 300 in figure 5.19 and 300 and 500 in figure 5.20. It is possible to see, in 5.18(c) and 5.18(f), that both PM-S- and PM-T-MUSIC can resolve the two reflections from the interfaces with velocities 2000 m/s and 2500 m/s, arriving around 0.3 s and 0.4 s. In 5.19(c) and 5.19(f), only PM-T-MUSIC can resolve these two reflections and in 5.20(c) and 5.20(f) neither PM-S- and PM-T-MUSIC are capable to resolve it. This behavior influences the stacking results, as shown in figures 5.21, 5.22, 5.23 and 5.24.

In figure 5.21 we show a close view of the stacking performed with the velocities provided by semblance coherence function. In figures 5.22, 5.23 and 5.24 we show the results of the stacking performed with velocities obtained with PM-MUSIC for spatial and temporal correlation matrices without normalization and normalized with semblance weighting and semblance balancing. We observe from the results that the stacking performed with the velocities estimated from PM-T-MUSIC have a better definition of the layers close to the pinch-out than the ones performed with semblance and PM-S-MUSIC. Moreover, the stacking from semblance balancing normalization seem to present a better resolution near the pinch-out than the ones from semblance weighting. After the pinch-out, the stacking with the best alignment are the ones with the velocities estimated with semblance and with semblance weighting normalization. We show the velocities used for stacking the data, estimated with semblance and with the different types of PM-S- and PM-T-MUSIC, in figure 5.25. It is possible to see several arbitrary values in most of the spectra. These values appear due to the automatic algorithm we used for choosing the velocities: when there were no reflections at the τ_0 being analyzed, the algorithm simply picked out the velocity which resulted in the highest coherence value. This has no impact in the stacking result because as there is no event in that τ_0 it does not matter the velocity used for stacking the data. It is possible to see, from figures 5.25(e), 5.25(f) and 5.25(g), that PM-T-MUSIC presents better estimative of the velocities near the pinch-out.

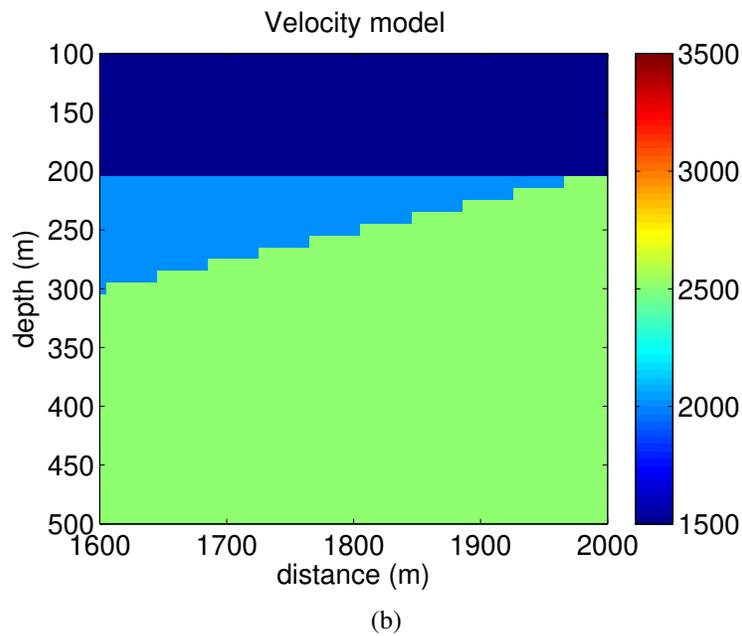
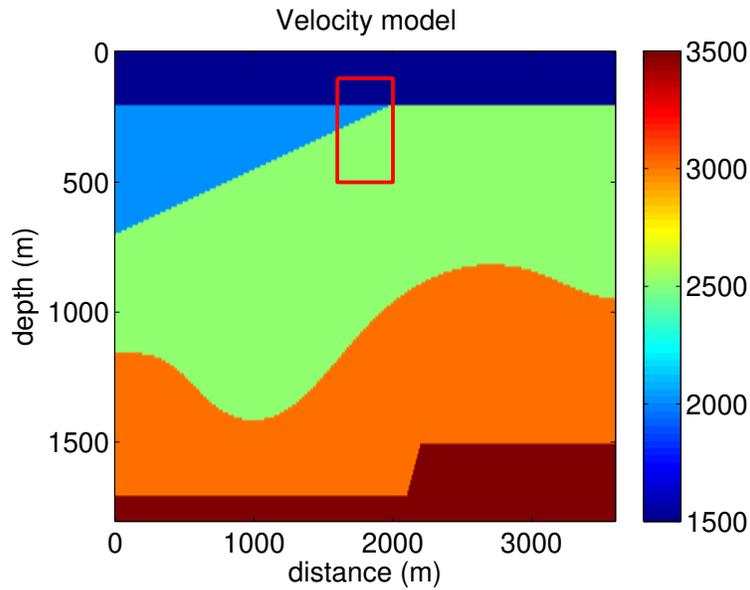


Figure 5.17: Velocity model (a) and close view of the region inside the red box, for stacking comparison (b).

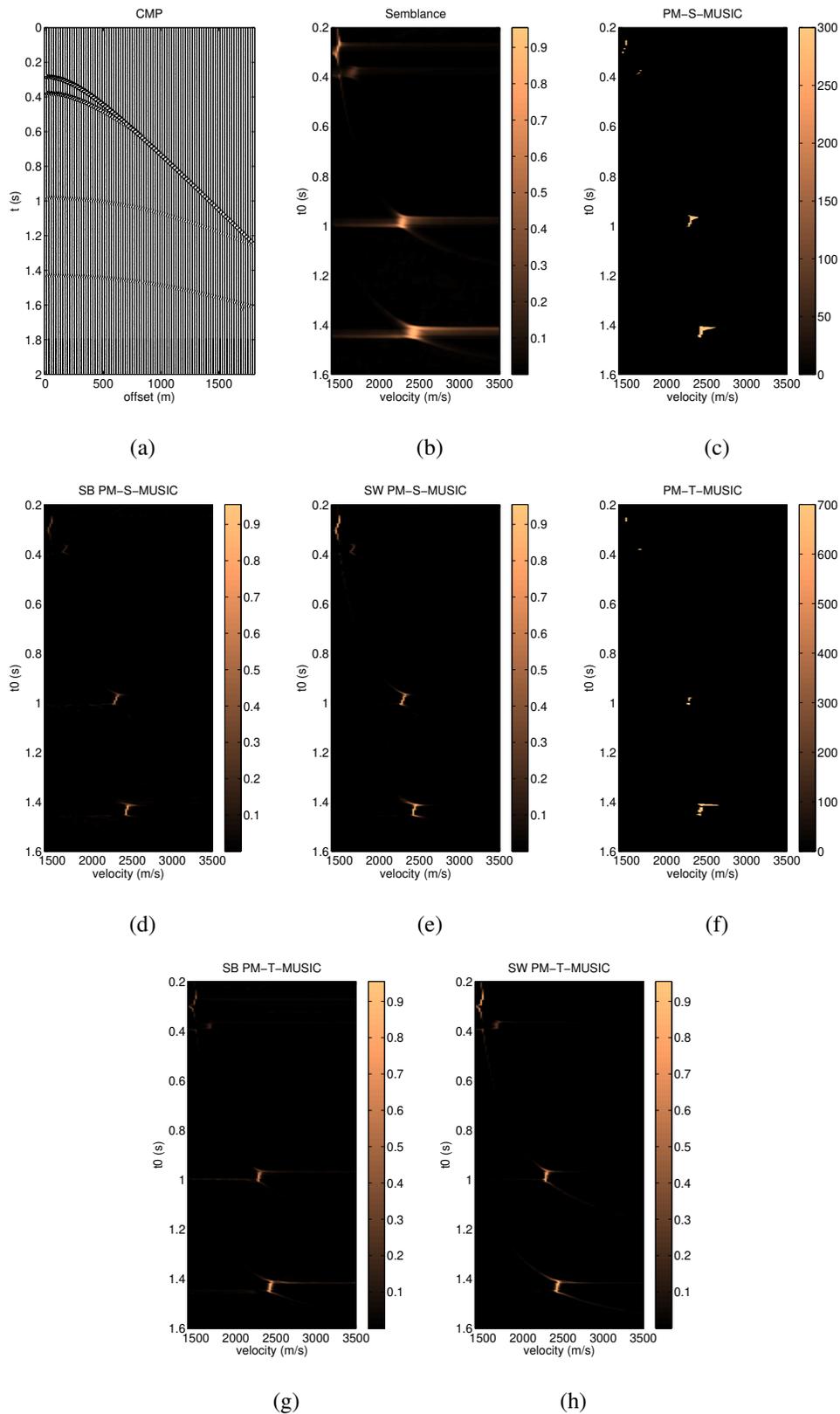


Figure 5.18: CMP gather at the midpoint position of 1600 m (a) and velocity spectra computed with semblance (b), PM-S-MUSIC (c) and PM-T-MUSIC (f) without normalization and normalized with semblance balancing (d) and (g) and semblance weighting (e) and (h).

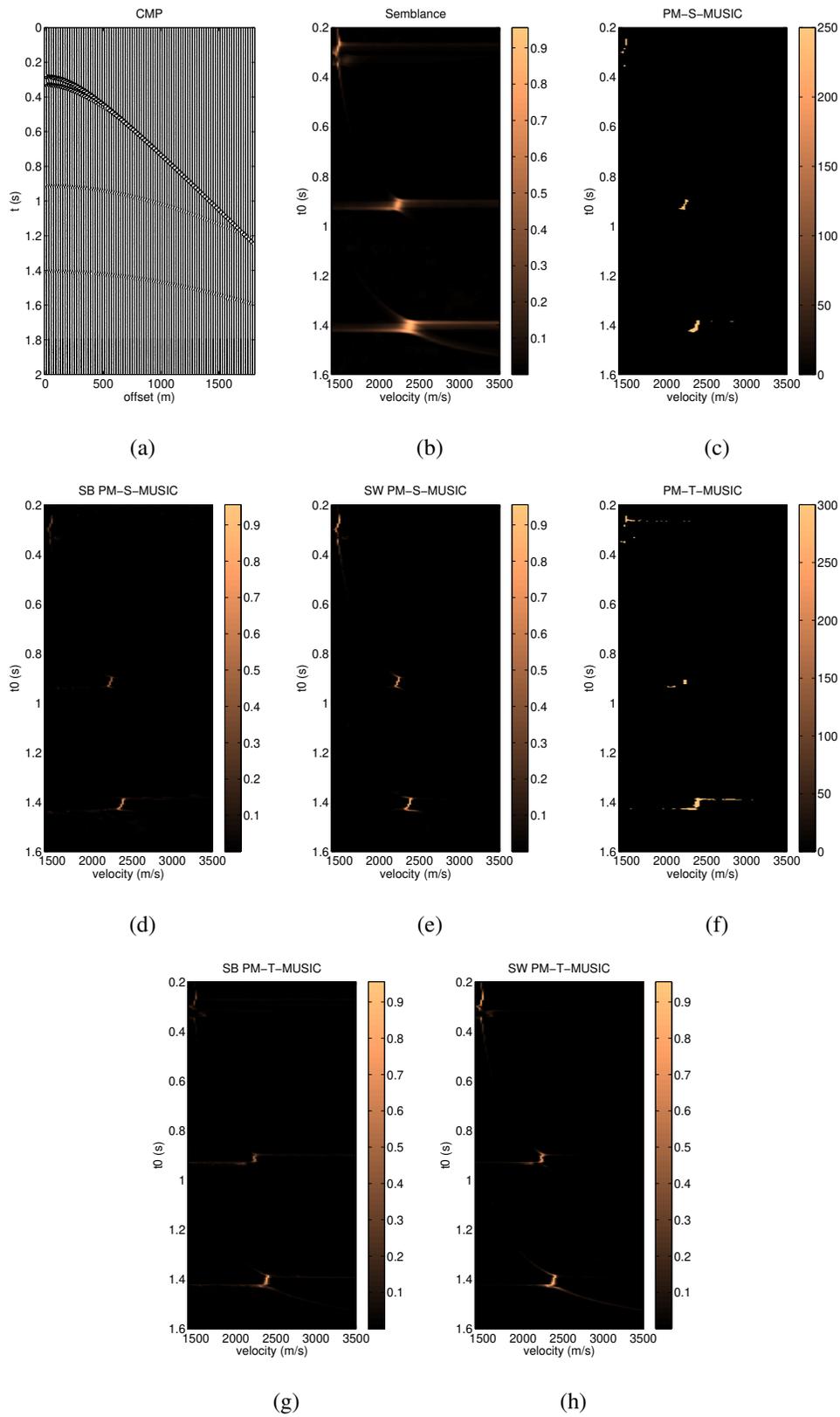


Figure 5.19: CMP gather at the midpoint position of 1800 m (a) and velocity spectra computed with semblance (b), PM-S-MUSIC (c) and PM-T-MUSIC (f) without normalization and normalized with semblance balancing (d) and (g) and semblance weighting (e) and (h).

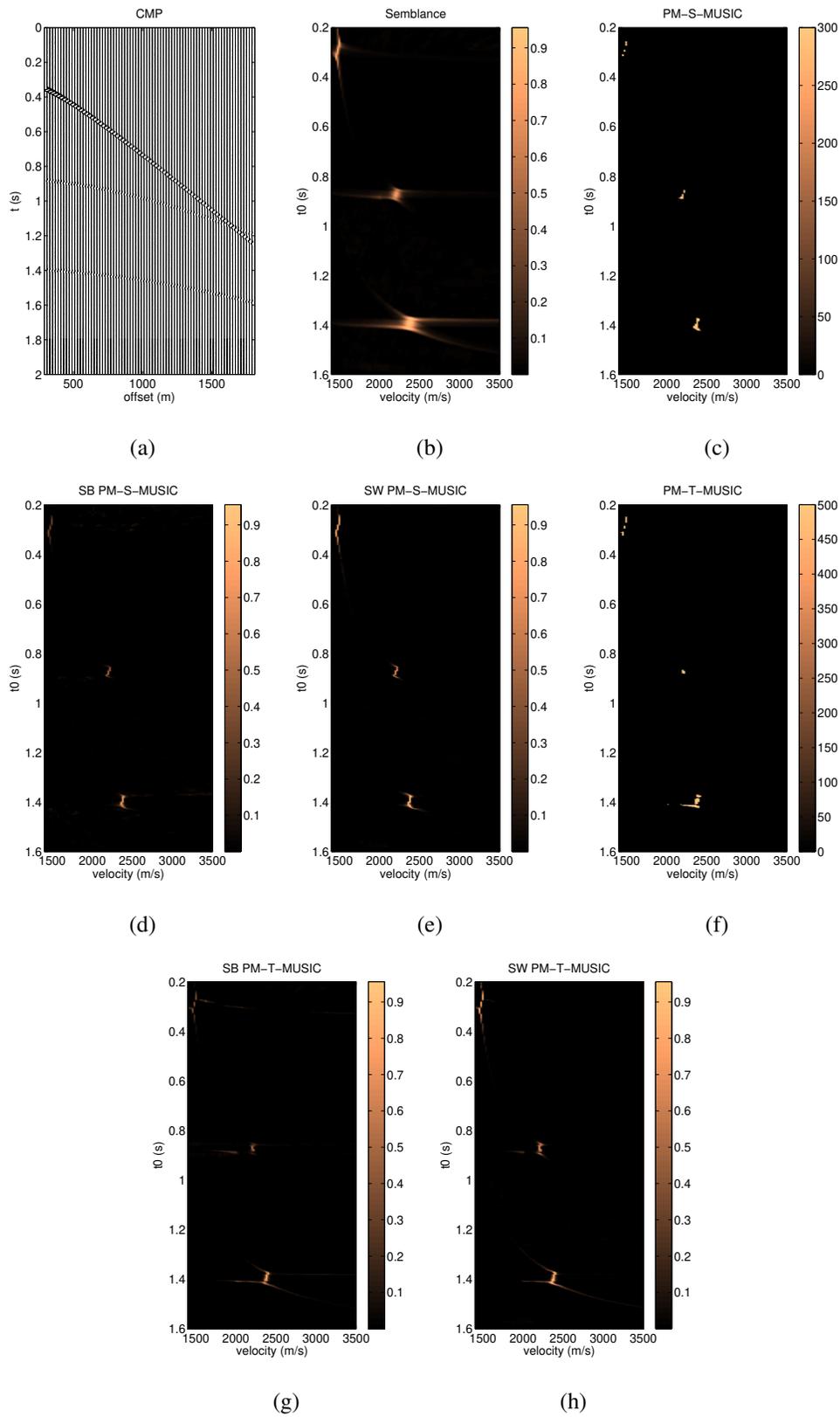


Figure 5.20: CMP gather at the midpoint position of 1950 m (a) and velocity spectra computed with semblance (b), PM-S-MUSIC (c) and PM-T-MUSIC (f) without normalization and normalized with semblance balancing (d) and (g) and semblance weighting (e) and (h).

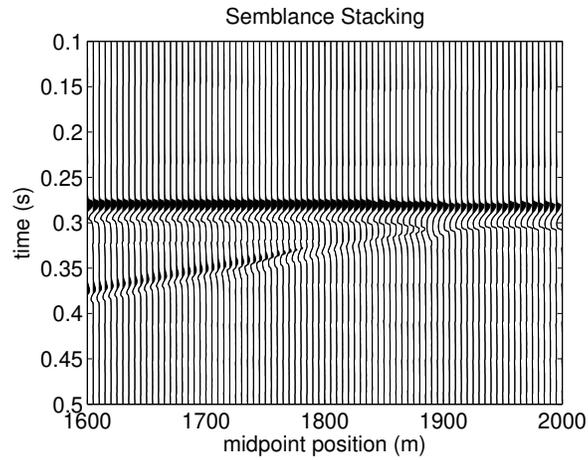


Figure 5.21: Stacking result close view for semblance.

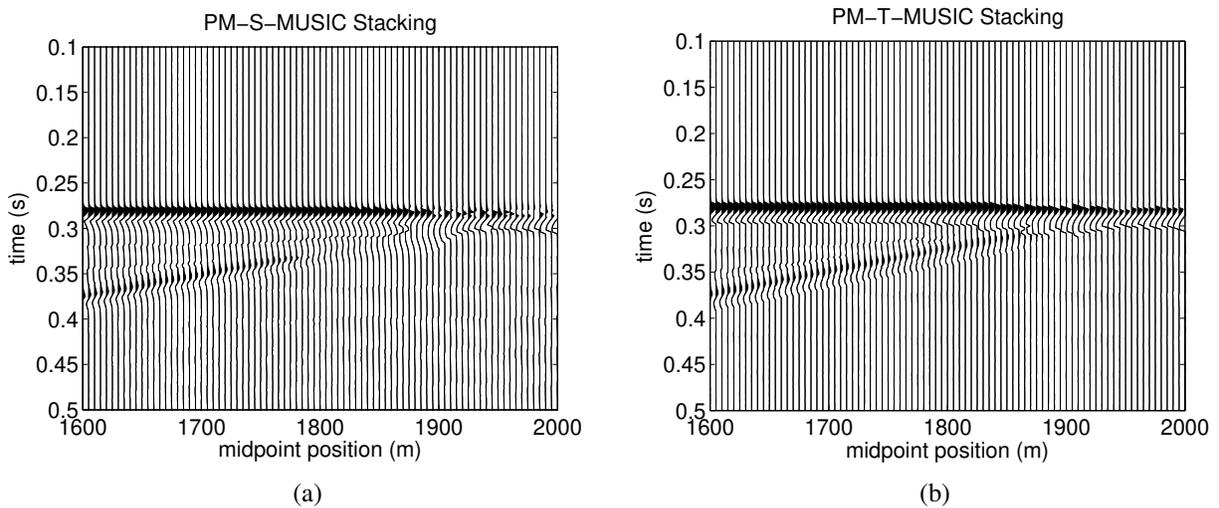


Figure 5.22: Stacking result close view for PM-MUSIC, without normalization, computed for spatial (a) and temporal (b) correlation matrices.

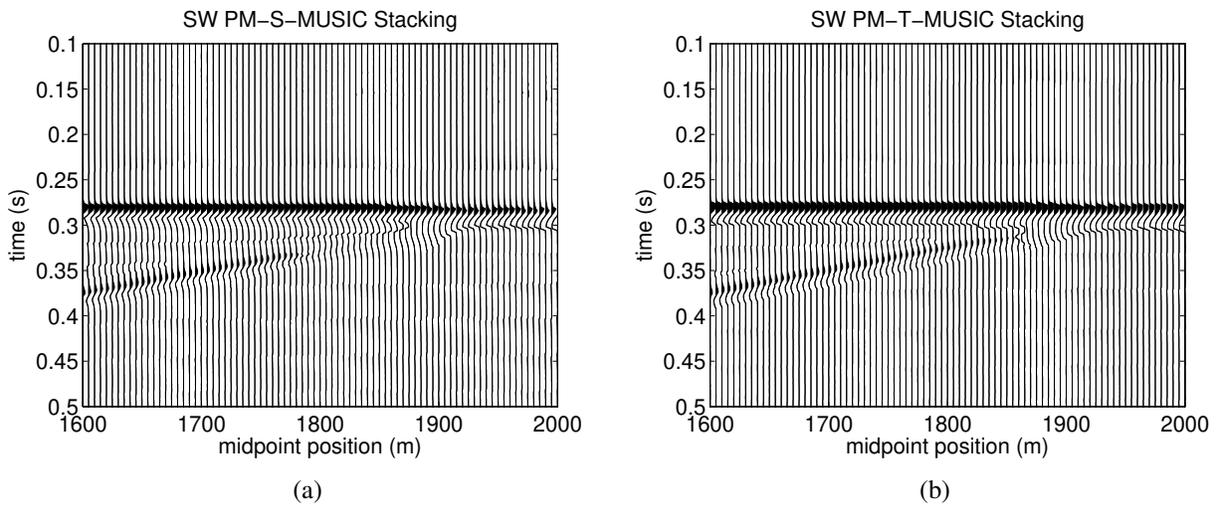


Figure 5.23: Stacking result close view for PM-MUSIC, normalized with SW and computed for spatial (a) and temporal (b) correlation matrices.

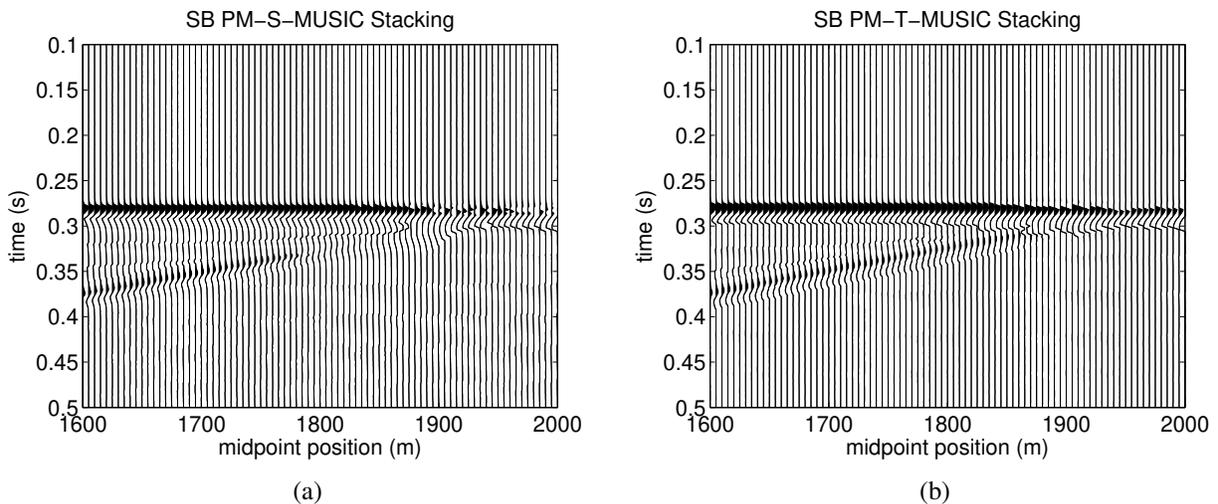


Figure 5.24: Stacking result close view for PM-MUSIC, normalized with SB and computed for spatial (a) and temporal (b) correlation matrices.

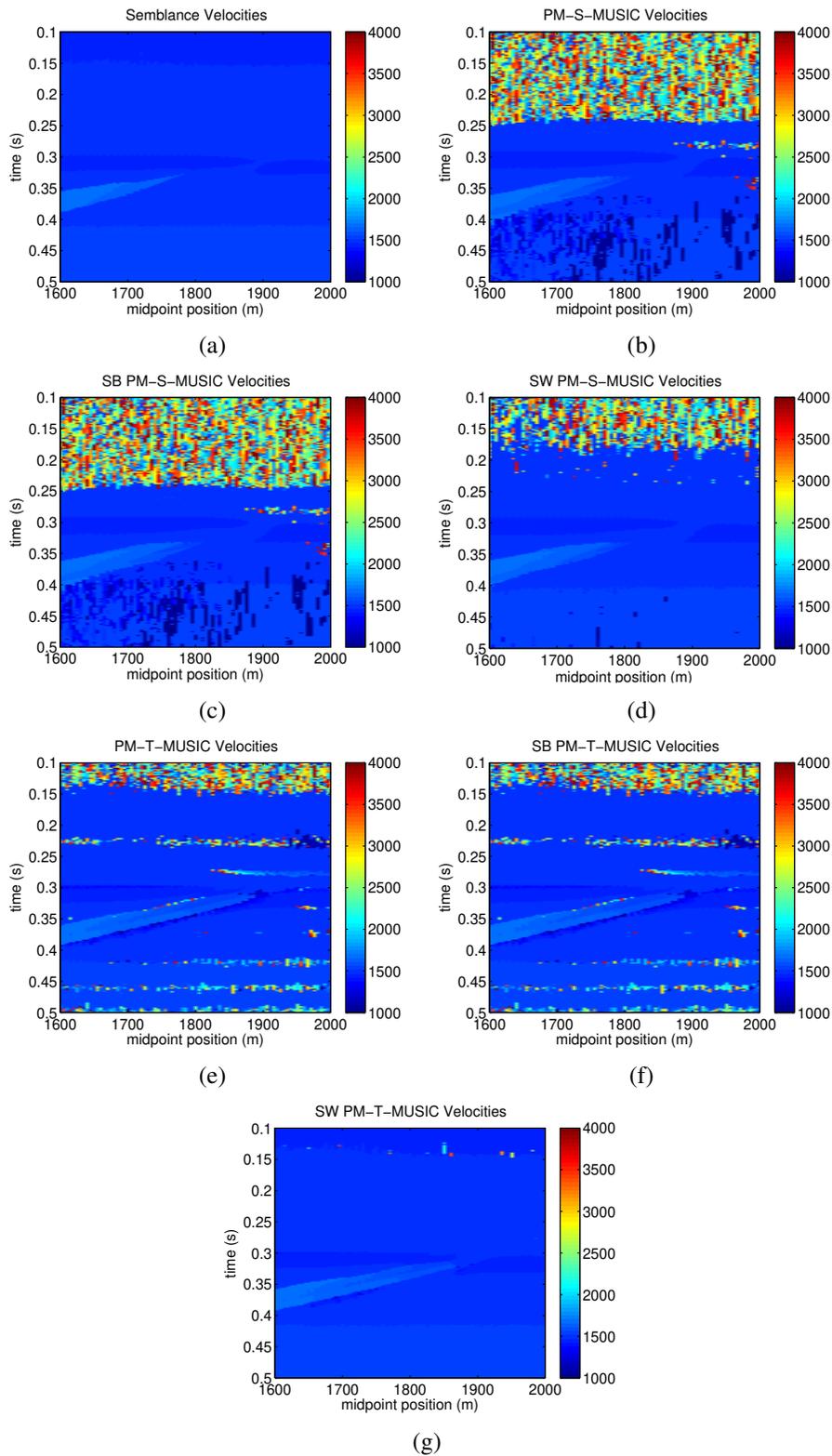


Figure 5.25: Stacking velocities estimated with semblance (a), PM-S-MUSIC (b), SB PM-S-MUSIC (c), SW PM-S-MUSIC (d), PM-T-MUSIC (e), SB PM-T-MUSIC (f) and SB PM-S-MUSIC (g). The criterion was to choose, for each τ_0 , the velocity corresponding to the highest coherence value.

Chapter 6

Conclusions

In this dissertation, we discussed stacking parameters estimation in seismic signal processing. The focus of our discussion was on the high-resolution method known as MUSIC, which produces better estimates of the parameters in comparison with the classical semblance. The standard version of MUSIC is based on the eigendecomposition of the spatial correlation matrix, computed from seismic data, and we refer to it as S-MUSIC. We presented an iterative method to perform this eigendecomposition and coherency calculation, which we named PM-S-MUSIC. When compared to S-MUSIC, PM-S-MUSIC algorithm allows a reduction of complexity, due to the fact that it is based only on the iterative estimation of the eigenvector related to the largest eigenvalue (signal subspace) from spatial correlation matrix, unlike S-MUSIC, which is based on the estimation of all the other remaining eigenvectors (noise subspace). We presented a new way to perform eigenstructure-based velocity spectra calculation, based on eigendecomposition of the temporal correlation matrix. We called this method T-MUSIC. We also presented the temporal variant of PM-S-MUSIC, which we named of PM-T-MUSIC.

The numerical examples, shown in chapter 5, indicated that PM-S-MUSIC outperforms semblance and that its temporal variant, PM-T-MUSIC, can present the same high-resolution as its spatial counterpart. Moreover, PM-T-MUSIC is particularly useful when dealing with correlated signals, as we do not need to use spatial smoothing together with forward-backward averaging for conditioning of the correlation matrix. We have seen, however, that the estimative of the tested eigenvector in PM-T-MUSIC could be improved by taking into account the different amplitudes and shapes of the wavelets along the traces. This appears to be a natural continuation of the method proposed in this work.

The complexity orders for the semblance and for the spatial and temporal versions of MUSIC and PM-MUSIC were presented in chapter 4. From that analysis, we verify that temporal versions of MUSIC and PM-MUSIC leads to computation savings when compared to its spatial counterparts,

mainly due to the fact that usually the window size is smaller than the number of traces in a seismic data.

We have also proposed, in chapter 4, a different normalization function that makes use of semblance coefficient (semblance weighting) in order to deal with high dynamic range in the produced velocity spectra for both S-MUSIC and T-MUSIC coherence functions. We compared that normalization function with a different one, which can be found in the literature (semblance balancing). Based on the numerical examples presented in this dissertation, and others that we did not show here, we observe that semblance weighting tends to generate better coherency spectra in cases that the semblance coherence function already presents a reasonable spectrum. On the other hand, when the spectra obtained from semblance are not good semblance balancing tends to generate better results.

Bibliography

- B. Abbad and B. Ursin. High-resolution bootstrapped differential semblance. *Geophysics*, 77(3): U39–U47, 2012.
- H. Akaike. Information theory and an extension of the maximum likelihood principle. *Proceedings of 2nd International Symposium on Information Theory*, suppl. Problems of Control and Information Theory:267–281, 1973.
- E. G. Asgedon, L. G. Gelius, and M. Tygel. Higher resolution determination of zero-offset common-reflexion-surface stack parameters. *International Journal of Geophysics*, 2011.
- B. Biondi and C. Kostov. High-resolution velocity spectra using eigenstructure methods. *Geophysics*, 54(7):832–842, 1989.
- C. H. Dix. Seismic velocities from surface measurements. *Geophysics*, 20(1):68–86, 1955.
- M. A. Doron and A. J. Weiss. On focusing matrices for wide-band array processing. *IEEE Transactions on Signal Processing*, 40(6):1295–1302, 1992.
- B. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 3rd edition, 1996.
- S. Haykin. *Modern Filters*. Macmillan Publishing Company, New York, 1989.
- H. Hung and M. Kaveh. Focussing matrices for coherent signal-subspace processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36(8):1272–1281, 1988.
- S. C. Key and S. D. Smithson. New approach to seismic-reflection event detection and velocity determination. *Geophysics*, 55(8):1057–1069, 1990.
- S. C. Key, R. L. Kirlin, and S. D. Smithson. Seismic velocity analysis using maximum-likelihood weighted eigenvalue ratios. *57th Annual International Meeting, SEG*, 1987.

- R. L. Kirlin. A note on the effect of narrow band and stationarity signal model assumptions on the covariance matrix of sensor-array data vectors. *IEEE Transactions on Signal Processing*, 39(2):503–506, 1991.
- R. L. Kirlin. The relationship between semblance and eigenstructure velocity estimators. *Geophysics*, 57(8):1027–1033, 1992.
- R. L. Kirlin and W. D. Done. *Covariance Analysis for Seismic Signal Processing*. Society of Exploration Geophysicists, Tulsa, 1st edition, 1999.
- H. Krim and M. Viberg. Two decades of array signal processing research: the parametric approach. *IEEE Signal Processing Magazine*, 13(4):67–94, 1996.
- N. Neidell and M. Taner. Semblance and other coherency measures for multichannel data. *Geophysics*, 36(3):482–497, 1971.
- J. Rissanen. Modeling by shortest data description. *Automatica*, 14(35):465–471, 1978.
- M. Sacchi. A bootstrap procedure for high-resolution velocity analysis. *Geophysics*, 63(5):1716–1725, 1998.
- R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.
- G. Schwartz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.
- T. J. Shan, M. Wax, and T. Kailath. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(4):806–811, 1985.
- H. L. Van Trees. *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory*. John Wiley & Sons Inc., New York, 2002.
- H. Wang and M. Kaveh. Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(4):823–831, 1985.
- Y. Y. Wang, J. D. Chen, and W. H. Fang. TST-MUSIC for joint DOA-delay estimation. *IEEE Transactions on Signal Processing*, 49(4):721–729, 2001.
- M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(2):387–392, 1985.

-
- M. Wax, T. J. Shan, and T. Kailath. Spatio-temporal spectral analysis by eigenstructure methods. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32(4):817–827, 1984.
- R. T. Williams, S. Prasad, A. K. Mahalanabis, and L. H. Sibul. An improved spatial smoothing technique for bearing estimation in a multipath environment. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36(4):425–432, 1988.
- O. Yilmaz. *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data, Volume 1*. Society of Exploration Geophysicists, Tulsa, 2nd edition, 2001.